



# IDS WORKING PAPER

Volume **2020** No **536**

## **How Do Donor-led Empowerment and Accountability Activities Take Scale into Account? Evidence from DFID Programmes in Contexts of Fragility, Conflict and Violence**

Colin Anderson, Jonathan Fox and John Gaventa

April 2020

Produced as part of:



Action for Empowerment and Accountability (A4EA) is an international research programme which explores how social and political action can contribute to empowerment and accountability in fragile, conflict, and violent settings, with a particular focus on Egypt, Mozambique, Myanmar, Nigeria, and Pakistan.

Led by the Institute of Development Studies, A4EA is being implemented by a consortium which includes: the Accountability Research Center, the Collective for Social Science Research, the Institute of Development and Economic Alternatives, Itad, Oxfam GB, and the Partnership for African Social and Governance Research. It is funded with UK aid from the UK Government. The views expressed in this publication do not necessarily reflect the official policies of our funder.



How Do Donor-led Empowerment and Accountability Activities Take Scale into Account? Evidence from DFID Programmes in Contexts of Fragility, Conflict and Violence

Colin Anderson, Jonathan Fox and John Gaventa

IDS Working Paper 536

© Institute of Development Studies 2020

ISSN: 2040-0209 ISBN: 978-1-78118-629-9

A catalogue record for this publication is available from the British Library.



This is an Open Access paper distributed under the terms of the Creative Commons Attribution 4.0 International licence (CC BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited and any modifications or adaptations are indicated.

<http://creativecommons.org/licenses/by/4.0/legalcode>

Available from:

Communications and Engagement Unit, Institute of Development Studies, Brighton BN1 9RE, UK

Tel: +44 (0) 1273 915637

Web: [www.ids.ac.uk/publications](http://www.ids.ac.uk/publications)

IDS is a charitable company limited by guarantee and registered in England

Charity Registration Number 306371

Charitable Company Number 877338

# **How Do Donor-led Empowerment and Accountability Activities Take Scale into Account? Evidence from DFID Programmes in Contexts of Fragility, Conflict and Violence**

Colin Anderson, Jonathan Fox and John Gaventa

## **Summary**

Development donors invest significantly in governance reform, including in contexts characterised by conflict and fragility. However, there is relatively little comparative study of their change strategies, and little understanding of what works and why. This paper explores the strategies of six recent DFID-funded programmes in Mozambique, Myanmar, and Pakistan with empowerment and accountability aims. Document review and field interviews are used to analyse the application of multi-scalar or multi-level change strategies, since such approaches are hypothesised to potentially generate more leverage for public accountability reforms. Analysis suggests that these strategies can strengthen citizen ability to navigate governance systems to resolve problems and claim accountability, and can bolster pro-accountability coalitions' internal solidarity and external legitimacy. Multi-level strategies also appear associated with establishing more significant pressure for reform than exclusively local or national approaches. Yet conventional project reporting focuses on counting activities and outputs rather than analysing the dynamic, interactive processes at work in these strategies, and few evaluations are publicly accessible. To fully understand what kinds of action strengthen citizen demands for accountability requires a more transparent and rigorous approach to learning from donor-led governance interventions.

**Keywords:** empowerment, accountability, fragility, donors, multi-level governance.

**Colin Anderson** is a research officer at the Institute of Development Studies. His work focuses on citizen-state relations, including as research officer for the Action for Empowerment and Accountability research programme. He has a particular interest in the development and use of social accountability approaches and their impacts on governance relationships.

**Jonathan Fox** teaches development studies at the School of International Service at American University, where he established and directs the Accountability Research Center (ARC). His work focuses on the relationship between citizen participation, transparency, and accountability from both scholarly and practitioner perspectives. ARC's work involves dialogue with a wide range of public interest groups, grassroots organisations, development agencies, private foundations and government policymakers to understand and inform change strategies.

**John Gaventa** is a political sociologist, educator and civil society practitioner with over 30 years of experience in research, teaching and facilitation, and organisational leadership in North and South. He has written and worked extensively on issues of citizenship and citizen engagement, power and participation, governance and accountability, and leadership for social change in a number of countries around the world. He is the Director of the Action for Empowerment and Accountability research programme, and previously directed the ten-year Development Research Centre on Citizenship, Participation and Accountability.

# Contents

	Summary, keywords and author notes	3
	Acknowledgements	6
	Executive summary	7
<b>1</b>	<b>Introduction</b>	<b>9</b>
<b>2</b>	<b>Situating the study</b>	<b>11</b>
<b>3</b>	<b>The project, methods, sources and limitations of evidence</b>	<b>15</b>
	3.1 Research approach	15
	3.2 Availability of evidence	18
<b>4</b>	<b>The case study programmes</b>	<b>20</b>
	4.1 AAWAZ Voice and Accountability Programme – Pakistan – 2011–18	21
	4.2 Alif Ailaan – Pakistan – 2013–18	21
	4.3 Empowerment Voice and Accountability for Better Health and Nutrition–Pakistan – 2012–19	22
	4.4 Citizen Engagement Programme – Mozambique – 2012–17	22
	4.5 Diálogo (Democratic Governance Support Programme) – Mozambique – 2012–17	23
	4.6 Burma Civil Society Strengthening Programme – Myanmar – 2011–16	23
<b>5</b>	<b>What were the programmes trying to do, and why?</b>	<b>24</b>
	5.1 Diagnoses of accountability conditions	24
	5.2 Multi-scalar and multi-sited governance	25
	5.3 Theories of change and intervention	26
<b>6</b>	<b>What multi-scalar strategies were used?</b>	<b>26</b>
	6.1 Direct vertical integration of citizen demands and claims on authorities	27
	6.2 Building vertical integration of civil society, campaigns and reform coalitions	29
	6.3 Developing horizontal linkages and networks	30
<b>7</b>	<b>What were the effects of multi-scalar strategies?</b>	<b>31</b>
	7.1 Navigation of the governance landscape	32
	7.1.1 Finding and motivating local solutions	32
	7.1.2 Escalation to activate accountability at higher levels	33
	7.1.3 Locating alternative ‘accountability targets’	34
	7.2 Creating solidarity and legitimacy	35
	7.2.1 Solidarity	36
	7.2.2 Legitimacy	37
	7.3 Leveraging evidence and multi-level pressure for policy reform	38
	7.3.1 Reforms to the content of public policy	39
	7.3.2 Reforms to institutionalise citizen participation	40
<b>8</b>	<b>How did the programmes measure their multi-scalar activity?</b>	<b>41</b>
	8.1 Programme monitoring regimes	41
	8.1.1 Measures of empowerment and accountability outcomes	44
	8.1.2 Use of quantitative and qualitative measures	44
	8.1.3 Measurement of multi-level and multi-scalar strategies and effects	45
	8.2 Understanding processes of change and impact	45
	8.3 Implications for learning	48

<b>9</b>	<b>Conclusions and implications</b>	<b>50</b>
	<b>References</b>	<b>54</b>
<b>Boxes</b>		
Box 7.1	Local demands raised and their resolution	33
Box 7.2	Solidarity through multi-scalar networks in Myanmar	37
Box 8.1	DFID's approach to programme evaluation	46
<b>Tables</b>		
Table 3.1	Programme case selection – indicating inclusion in documentary review and case study activities	16
Table 3.2	Availability of DFID programme documentation	19
Table 4.1	Case study programmes	21
Table 6.1	Use of multi-scalar strategies in case study programmes	27
Table 8.1	Summary of output and outcome monitoring measures of studied programmes	42
Table 8.2	Programme evidence – evaluations and other non-logframe sources	47

# Acknowledgements

This paper is the result of the work of an extended research team across the Accountability Research Centre (ARC) and the Institute of Development Studies (IDS). The ARC team comprised Daniel Esser, Jonathan Fox, and Suchi Pande. The IDS team comprised Colin Anderson, John Gaventa, Anuradha Joshi, Erika Lopez Franco, Rosie McGee, and Alex Shankland. Invaluable assistance, input and insights were provided throughout by Annette Fisher, in particular in the facilitation of our key stakeholder meeting. The authors thank all of those involved for their contributions to the research, but take responsibility for interpretations made here. Thanks are due to the many practitioners and observers who engaged in interviews and in stakeholder meetings, provided additional documentation and insights, and helped to shape this enquiry and our analysis.

# Executive summary

Recent years have seen significant development donor investment and activity focused on improving governance relationships in contexts characterised by conflict and fragility. There is relatively little comparative study of many of the strategies that they adopt and their effects, particularly how far they support or challenge a number of research-based propositions about what is more likely to support shifts in empowerment and accountability. In this paper we explore one such proposition – that using multi-level or multi-scalar strategies might be important. Multi-scalar strategies are those that deliberately try to join up action across a number of sites and institutions or organisations. We explore this idea by looking in detail at six programmes designed and funded by the UK Department for International Development that have operated in parallel in Mozambique, Myanmar, and Pakistan during the past five years. The programmes we looked at all attempted to open up more space for the involvement of ordinary citizens in public service decision-making and prioritisation, following a logic that this could support a shift towards more inclusive and responsive governance.

Drawing on interviews and workshops with people involved in these programmes in each country as well as an extensive review of programme documents and impact evidence, we found that:

- All of the programmes were designed to some extent to use multi-level strategies. Some very deliberately responded to evidence that social accountability efforts to mobilise citizens around the quality of local services tend to produce results only at that level, rather than seeding systemic reforms. Others set out ambitions to link up fractured organisations and constituencies.
- In part this responded to understandings of complex governance arrangements and norms in their contexts. Common across these were limited organised citizen demand for service or policy improvements, diffuse sources of formal and informal authority and decision-making, and norms constraining the degree of ‘voice’ that many groups – particularly women – were able to exercise.
- Three distinct multi-scalar strategies were used in practice, with each programme tending to focus largely on one of these:
  - a) **Vertical aggregation of citizen demands and claims on authorities**, conveying these upward to those responsible for public services.
  - b) **Vertical integration of civil society organisations and campaigns**, linking civic organisations at the grassroots with sub-national and national level civil society activism and lobbying.
  - c) **Developing horizontal linkages and networks**, so that organisations and reformists working across geographies or sectors could find common ground and act jointly.
- Using multi-scalar strategies helped some programmes to *navigate the governance and service delivery landscape* to activate accountability at different levels, move around obstacles, and highlight systemic issues.
- Linking up action and organisations both horizontally and vertically was also important in *building and activating solidarity* with loose coalitions for reform, and in actors’ perceptions of *how legitimate their claims and actions seemed* to others.
- Multi-scalar strategies were also involved where programmes argued that they had contributed to systemic changes in policy and practice beyond just their localised sites of operation, partly in *building a weight of evidence*, and partly in *activating multi-level pressure* for reform.

We are also able to make some more general reflection on the ability to learn about strategies for social change from exploring DFID-funded programmes in this way. We found:

- Important gaps and limitations in publicly available documentation on this set of programmes, in particular evaluative information that identifies what changes have taken place and why.
- Patchy practice in formal evaluation, which limits how conclusive we could be about programme success and contributing factors, and very few publicly available evaluations of this kind.
- Significant sources of un-exploited evidence that could inform a stronger understanding of what works, mostly held by implementing organisations rather than DFID itself.
- A gap between what programmes formally measured their success on and what they are actually seeking to do – undermining their contribution to the evidence base as well as producing ongoing tension between practices and measurements of success.
- Significant barriers to the rigorous and open assessment of programme outcomes as a result of incentives in the aid architecture to be seen to succeed and the competitive commercial value of learning and evidence in the sub-contracted delivery model.

Taken together, we can see in the six programmes explored how programme designers and practitioners build in multi-scalar strategies as a contextual response to the complexities of governance in fragile and conflict-affected settings and a tactical response to anti-accountability forces. There is the potential both in these programmes and presumably in similar activities in other contexts to learn more about when and how these tactics work. For this learning to happen, however, there needs to be a shift in incentives towards greater transparency and rigorous assessment of what these donor-funded programmes are achieving, and what this means for those on the ground seeking to build or demand more accountable governance.



# 1 Introduction

Decisions that matter to people's lives are taken in many places within systems of governance. They are taken both at local and national levels, across different sets of actors or institutions, and across different geographies within the same country. They are taken by those with both formal and informal authority, constrained by political systems and arrangements that limit their autonomy and structure their interests. And they are underpinned by the wider, often fluctuating, balance of power between these different sites, institutions, and actors. In contexts characterised and affected by fragility, conflict and violence, we might expect this governance landscape to be particularly complex. Such complexity has significant ramifications for anyone trying to secure accountability from decision makers or introduce pro-accountability reform. Where are the decisions that matter really taken? Which parts of the system need to be targeted to initiate and sustain change?

The answers to these questions matter most to those directly trying to get problems solved or push for reforms to systems they live and contend with on a daily basis. In many countries they also matter to development donors, who in recent years have increasingly set their sights on the development of more accountable governance institutions and public service responsiveness to citizens. In 2014, Thomas Carothers and Saskia Brechenmacher referred to this ambition as a 'new development consensus' (Carothers and Brechenmacher 2014). This trend sees donors funding activities that they hope will build the power of citizens to hold authorities to account, and establish new institutional channels through which they can exercise that power. The UK Department for International Development (DFID) has been part of this consensus. In 2011, DFID developed a specific policy frame on the empowerment and accountability ambitions of its programming (DFID 2011c). Shortly after this policy framing, a macro-evaluation of DFID activities with stated empowerment and accountability aims started with a sample of 2,379 individual projects (e-Pact 2016: 16). This included many activities in contexts affected by conflict and violence, and where governance is considered to be 'fragile' – contexts where we might expect the governance landscape to be particularly complex, liable to flux, and harder to analyse clearly. DFID's approach to these contexts, set out in a number of separate policy statements, position empowerment and accountability activities as part of building institutions that consolidate democratic practices and generate gains in what people can expect of their governments; delivering a 'peace dividend' and rebuilding a social contract (DFID 2016a)

In this paper we explore how far a number of DFID-funded development programmes in conflict-affected contexts take into account the multi-sited nature of governance systems and use strategies that respond to this important feature of governance realities. We look at what kinds of strategies they use, and with what reported effects, exploring what the evidence generated by these programmes tells us about using what we call multi-scalar strategies. We use the notion of scale not in the common sense of 'scaling up' development activities<sup>1</sup> but in the sense of working across different levels and sites in governance landscapes, and distinguish between vertical and horizontal dimensions of scale that might be relevant to achieving change in governance relationships. By vertical scale we refer to the hierarchy of decision-making and spaces for citizen engagement and representation from the local to the national, and supra-national levels. By horizontal scale we refer to connections across geographical areas, service sectors or institutions at a similar level of decision-making or authority.

---

<sup>1</sup> In the development lexicon, 'scaling' has become synonymous with ambitions to take successful activities or approaches to a wider population, or to expand on piloted activities. For examples see (Cooley and Kohl 2006; Cooley and Papoulidis 2017; Hartmann and Linn 2008; Linn 2012; McLean *et al.* 2019; Simmons *et al.* 2007).

The lens of ‘taking scale into account’ allows us to place a focus on strategies of action that link various sites of action towards empowerment and accountability, which has been argued to be a key feature in moving from tactical to more strategic approaches more likely to generate sustained change (Fox 2015, 2016). For development donors, engaging with the realities of governance and accountability environments in this way responds to now-common appeals to ‘think and work politically’ or ‘do development differently’.<sup>2</sup> These appeals sit within a wider literature that is at best equivocal about whether donors and externally driven programmes stand much chance of success in building what might be seen as ‘countervailing power’ to entrenched interests (Blackburn *et al.* 2002; Gibson and Woolcock 2008). We explore these and other debates that lie behind our enquiry in Section 2.

The analysis here is based on research undertaken during 2017 and 2018 as part of the DFID-funded Action for Empowerment and Accountability research programme,<sup>3</sup> and focuses on DFID programmes in Myanmar, Mozambique, and Pakistan. The research involved structured documentary review, fieldwork interviews with practitioners, and a consolidating learning workshop with practitioners and other researchers. Our approach and the assessment of the evidence we were able to gather from these exercises is detailed in Section 3.

We narrowed from an original sample of twenty-one DFID programmes to six programmes to explore in more depth. Section 4 gives a short overview of each of these six case study programmes, explaining their backgrounds and contexts, main activities, and reported achievements. In Section 5 we illustrate how the programmes share a set of common understandings of the core challenge of instituting more accountable governance relationships in fragile and conflict-affected settings. Their contextual analyses had many commonalities, and all recognised the multi-scalar nature of governance systems in the different countries they analysed. They also shared rationales for intervention and drew from a similar repertoire of action and intervention to support claims for accountability.

In Section 6 we look at how the programmes applied their shared understanding that scale matters in achieving governance change. We distinguish three distinct multi-scalar strategies used. One purposefully attempted to surface citizen needs, demands and experiences through vertical hierarchies of decision-making. Another worked explicitly to develop connections between national, sub-national, and local organisations and reform coalitions to strengthen advocacy at each level. A third common approach was to build horizontal networks and relationships to strengthen pro-accountability efforts and countervailing power.

Section 7 explores some of the effects reported of these multi-scalar strategies. We propose three distinguishable categories of effects drawn inductively from the reports of programme activity and outcomes. First, we look at the evidence that multi-scalar approaches allowed the programmes to navigate layers of decision-making and spaces for accountability. We found that several programmes reported direct localised improvements through engaging citizens in deliberation and equipping them with capacities and opportunities to make claims on authorities for better services or for the accountability of local providers and officials. To some degree, these gains appear to have been supported by the existence of hierarchies where unresolved claims could be escalated or networks through which resistance to accountability could be circumnavigated. This appears particularly important given common contextual conditions where informal power and low levels of official delegated authority sit side by side. We then explore effects of multi-scalar approaches on the internal solidarity of

---

<sup>2</sup> These phrases allude to particular schools of thought on how development programmes can be more effective, particularly in engaging with local political contexts. For recent reviews see (Dasandi *et al.* 2019; McCulloch and Piron 2019).

<sup>3</sup> Action for Empowerment and Accountability (A4EA) is a five-year enquiry into how processes of social and political action lead to empowerment and accountability outcomes in contexts of fragility, conflict, and violence. A core research question asks about the role of external actors such as development donors in supporting such processes.

pro-accountability coalitions, and external perceptions of the legitimacy of their claims. We illustrate these with examples of connections developed between otherwise isolated sites of action, and cases where these articulated to respond to challenges in gaining accountability or reform. The final set of effects explored are around the use of multi-scalar strategies to build wider pressure for reform, both to the content of public policy but also to the underlying accountability eco-system. Reports of changes like these are particularly complicated by difficulties in attribution from a limited evidence base, but we explore how multi-scalar strategies might have contributed to reform successes claimed by some programmes.

In Section 8 we explain the limitations of this evidence base in more depth. During our study we sought to understand how programmes measured and evidenced their multi-scalar strategies and effects on the basis that this might contribute to the wider evidence on when, where, and how multi-scalar strategies work to achieve accountability. We looked at official monitoring frameworks and regular performance measurement as well as more reflective and evaluative content found in some programme documentation, and in some cases independent assessments and evaluations. Although the programmes we looked at had very similar aims and used similar strategies, a diversity of approaches to monitoring and evaluation prevents much rigorous comparison on the basis of the available evidence. We conclude this section with some reflections on what this means in terms of learning from the programmes – in particular the incongruities we found between programme aims and measures of success, the obstacles to clear-sighted evaluation of impact, and the limits on informing new programming decisions.

The final section draws together the implications of our analysis. We note the limitations of what we can say about actual effects and interactions but indicate some ways in which this study might motivate and shape further analysis of these and other programmes, and add to the theoretical propositions on why multi-scalar approaches are warranted.

## 2 Situating the study

Establishing accountable governance institutions has become increasingly central to conceptions of what good development looks like in the past two decades (Booth 2012; Carothers and Brechenmacher 2014). Conceived of both as an outcome in itself and as a precursor or requirement for other important outcomes, accountability has become regarded as a key part of supporting sustainable development on the basis that it generates more responsive behaviour from those in authority, holds those in power in check, and forms state-society bargains or social contracts over the provision of goods and entitlements over time. Sustainable Development Goal 16 explicitly links such outcomes with inclusivity, peace and freedom from violence.

Achieving these kinds of institutions and relationships is of course complex and faces many barriers. For one, it is not necessarily in the interests of those in power to make themselves more answerable to the population at large. Such answerability not only comes with risks of individual sanction, but also with disrupting or exposing the networks of power and influence that sustain social and political systems. Even if it is in the interests of those in power to develop more accountable public institutions, it is not necessarily easy for them to do. For example, establishing greater accountability for delivery of public services requires that those with formal authority are able to resource and direct those services, which is perhaps easier said than done. Anti-accountability forces can be strong, pervasive, and persistent. This might be particularly the case in contexts of recent or ongoing conflict and violence, and unresolved issues about who should be holding power.

One approach to addressing anti-accountability forces is to seek to build 'countervailing power' outside of governance institutions themselves; sources of demand for reform and accountability that incentivise and motivate the creation of new bargains and build broader social constituencies for change (Agrawal and Ribot 1999; Fung and Wright 2003). Such an approach theorises that establishing these kinds of countervailing power can challenge more extractive, exclusive forms of governance, and hold in check or restrain the actions of those in authority. The underlying intervention logic is to empower relevant groups and individuals to make successful demands for change – whether better services, more equitable allocation of resources, or more inclusive decision-making.

Whilst attempting to build countervailing power might be commonly recognised as the work of activists and movements, it is also well-established in the repertoire of interventions and imagined aims of international development donors in various ways. DFID is one example. Although this framing of donor aims is perhaps not commonplace, it is not unique. Gibson and Woolcock (2008) explore the extent to which large-scale participatory development programmes in Indonesia build countervailing power, and Fox (forthcoming) asks the same questions of a number of historic World Bank projects. It is a lens that we see as productive by focusing analytical attention on how power builds and is exercised, and whether accountability-aimed donor initiatives produce any shifts in these processes and relations (following McGee and Gaventa (2011) in understanding the potential for such shifts taking place as a result of actions targeting public accountability). DFID's theory of change for its development programming on empowerment and accountability posits that increasing inclusivity of decision-making and broadening the political settlement to include wider state-society bargains will lead to a range of political, social and economic benefits. These anticipated shifts in power are hoped to ultimately enable greater voice, choice and control for the poor and marginalised (DFID 2011c). This continues a trend in thinking from a substantial earlier policy paper on governance-related programmes, which set out aims of mutually supportive outcomes of state accountability, capability, and responsiveness in a framing that specifically understands this to involve shifts in political power away from narrow elites and the state (DFID 2007: 6). This latter theme was taken up in DFID's major policy framework for fragile and conflict-affected settings, the Building Stability Framework, which introduced the language of 'fair power structures' and emphasised the contribution of 'helping citizens' voice to be heard, [and] ensuring authorities are accountable to those they serve' (DFID 2016a: 14). Similar language is used in a recently published DFID position paper on governance programming, which also places substantial emphasis on the role of more inclusive and accountable governance in reducing or avoiding conflict (DFID 2019a: 7–8).

Attempting to influence these issues of distribution of power and the behaviour of powerful actors and institutions from the position of an external donor is of course subject to much debate and complexity. There are arguments both around the legitimacy of such ambitions in themselves but also how effective any efforts might actually be. In more complex and conflictual settings, which present less predictable contexts, power relations are harder to read, and there is a greater risk of exacerbating rather than resolving conflicts. Justino (2019) notes that recent substantial donor investment in governance programmes in fragile and conflict-affected contexts is not backed by a strong evidence base. Whilst she finds supportive evidence for a number of common approaches, Justino concludes that the evidence base overall does not allow a fine-grained understanding of what works under different conditions. Christie and Burge (2017) review the evidence on the role of external actors specifically in fragile and conflict-affected contexts, also noting limited evidence behind what appear common practices. They do however find evidence for the efficacy of various approaches to empower civil society organisations to play greater roles in public discourse and social mobilisation, although less in relation to public accountability outcomes. One challenge here is the recognised difficulty in evaluating the impact of governance-focused development programmes, which tend to have aims that are inherently complex to

assess, particularly in the relatively short timescales in which they are funded (Eyben *et al.* 2015; Stern *et al.* 2012).

External actors concerned with governance reform draw on a variety of commonly documented strategies. One is to try to put in place systems that increase local oversight of specific services, gathering evidence on quality of services, and increasing citizen voice on their priorities and experiences as users of these services. These approaches are often framed in terms of the concept of social accountability. Popularised in the early part of the last decade following the publication of the influential 2004 World Development Report (World Bank 2004), social accountability has been the subject of significant policy diffusion globally and has grown to be a key part of externally-funded development programmes (e-Pact 2016; Waddington *et al.* 2019). Joshi (2014) points out that various formations of social accountability appeal to different kinds of outcomes; from the more instrumental immediate fixes to service provision that come from better understanding service user needs, to more transformative changes in how citizens and the state relate to one another in terms of rights and responsibilities.

Social accountability approaches can run alongside other common strategies that we can see as attempting to build countervailing power, which try to build institutions in civil society or within the state that have the potential to hold powerholders to account. Joshi and Moore (2000) distinguish these as attempts to work on 'enabling environments' rather than civic mobilisation. Within the state strategies can include strengthening formal oversight bodies such as ombuds agencies, audit bureaus or independent judicial review processes, or the ability of different levels and sections of government to challenge one another – often described as horizontal or diagonal accountability relationships (Goetz and Jenkins 2005). Within civil society they might involve supporting the growth of civic organisations and collectives that act to represent particular interests or constituencies, advocate for change, and build reformist movements and momentum. Positive outcomes of attempts at these kinds of support to civil society have been reported in independent reviews of DFID programmes (ICAI 2013a), in recent systematic reviews (Waddington *et al.* 2019), and internal reviews of international NGO programmes on civic involvement in governance in conflict-affected contexts (Amakom *et al.* 2018; Fooks 2013: 6).

Although these strategies are well-established in the repertoire, evidence from evaluations and academic study is inconclusive on whether and in what contexts and scenarios they can generate sustained increases in government responsiveness and institutionalised improvements in the local accountability environment. This is in part arguably a result of the ways in which this evidence is gathered (Eyben *et al.* 2015; Holland *et al.* 2009; Stern *et al.* 2012). Reviews of transparency-based and social accountability programmes repeatedly highlight the significance of a multitude of contextual factors (Grandvoinet *et al.* 2015; O'Meally 2013; Tembo 2012) and caution realism in relation to how far such tools in and of themselves should be expected to shift power dynamics that sustain unresponsive governance (Fox 2007b; Gaventa and McGee 2013; Joshi 2013; Rocha Menocal and Sharma 2009). A DFID commissioned macro-evaluation of its empowerment and accountability programming ultimately focused largely on such social accountability initiatives. Using a Qualitative Comparative Analysis method, this concluded that such initiatives almost always led to service-level changes, but that systemic impacts were less observed (e-Pact 2016). The authors took this as evidence of 'low accountability traps' at local level, a term introduced by Fox (2007a) to refer to situations where institutions with weak vertical and horizontal accountability reinforce one another to limit the chances of accountability claims making it through the system. The macro-evaluation suggested that within the programmes reviewed, achieving higher-level policy changes through such approaches required the integration of citizens' experiences within the formal institutionalised spaces of the service delivery and governance system, although this was not sufficient in the examples analysed to generate those changes in isolation. A recent systematic review of

evaluation evidence from a range of citizen engagement modalities, including those associated with social accountability approaches, drew a similar conclusion that localised improvements in services and relationships were common but these were not necessarily linked to evidenced improvements in what they termed ‘wellbeing’ outcomes (Waddington *et al.* 2019).

Fox (2015) highlights that efforts to mobilise citizens or institutions to raise their voice and make louder demands for public accountability often fail to activate institutionalised or formal mechanisms that can effectively sanction decision makers – what he calls ‘voice without teeth’. Reviewing the treatment of evidence for social accountability initiatives, Fox introduces an important distinction between tactical and strategic approaches. Strategic approaches are defined as those that ‘deploy multiple tactics, encourage enabling environments for collective action for accountability, and coordinate citizen voice initiatives with reforms that bolster public sector responsiveness’ (Fox 2015: 346) and thereby have the potential to activate accountability across multiple sites within and outside of the state. This particular element of more strategic approaches was developed further by Fox and colleagues using the notion of ‘vertical integration’ in relation to service monitoring and advocacy efforts to analyse successful pro-accountability action in a number of contexts at multiple levels of government service delivery (Aceron 2018; Fox 2001, 2016; Fox *et al.* 2016). The wider argument is that working in a multi-scalar way enables reformers to ‘address the systemic embeddedness of anti-accountability forces across multiple levels and branches of a state apparatus’ (Fox 2016: 8).

This conception of scale is distinct from the notion of ‘going to scale’ in the sense of doing more or broadening the impact of development activity. In usage here, scale specifically denotes multi-directional relationships between sites of governance or social and political action and decision-making.<sup>4</sup> These connections or linkages might be vertical – connecting between the many layers of action from the very local to the national and international – or horizontal – across multiple physical localities, constituencies or sectors. Working with horizontal scale does not necessarily involve horizontal accountability, which refers more narrowly to the extent to which authorities and institutions at the same level act to hold one another to account (O’Donnell 1998).

Working in this way is shown in the body of literature that emerged around this idea, to have various advantages that overcome the reportedly often isolated impacts of accountability initiatives. A number of reasons for this are explored. For one, it is suggested that approaching accountability deficits in a way that takes scale into account might allow ‘pro-accountability actors... to target or bypass accountability bottlenecks’ to get issues resolved (Fox 2016: 13). It is also suggested, particularly in the cases identified in the Philippines by Aceron (2018), that a critical mass of evidence can be amassed to support reform, and that such connections enable national advocacy to be grounded in local service monitoring. At a wider level the theory suggests that adopting a multi-scalar strategy is important for building pro-reform constituencies and demand for accountability, and that cross-scalar linkages between different kinds of civil society actors and organisations might be important in creating pro-accountability constituencies with the depth and breadth to present a challenge to ingrained anti-accountability forces (Fox 2016). Accountability strategies that take scale into account thus involve coordinated activity at multiple levels, as well as planning how to broaden the social constituencies that can drive institutional change. Fox suggests that disconnects between organisations with ‘eyes and ears’ and felt social needs at a local level, and those undertaking national level policy advocacy hamper pro-accountability reform, and

---

<sup>4</sup> A discussion on these directions and kinds of multi-scalar effects and interactions can be found in a presentation developed during this project by Jonathan Fox, delivered to a project group of the International Budget Partnership in July 2018, and available online here: [https://jonathanfoxucsc.files.wordpress.com/2018/08/taking\\_scale\\_into\\_account\\_ibp\\_spark\\_july\\_24\\_2018.pptx](https://jonathanfoxucsc.files.wordpress.com/2018/08/taking_scale_into_account_ibp_spark_july_24_2018.pptx)

that these disconnects are likely to be ‘especially relevant in fragile and conflict settings’ (Fox 2016: 7).

In later sections we explore how far a particular sample of DFID-funded programmes use strategies like these. First, we explain how we arrived at this sample, and studied the programmes’ activities.

## 3 The project, methods, sources and limitations of evidence

The findings presented here come from an extended period of research undertaken in several stages by a team from the Accountability Research Center and the Institute of Development Studies, as part of the Action for Empowerment and Accountability Research Programme (A4EA). Research activities were motivated by a recognition that in four of the five focus countries of the A4EA programme, DFID had invested substantially in programmes with explicit empowerment and accountability aims. Several of these, such as the SAVI programme in Nigeria, had been considered flagship activities in the governance sector. Under the broad line of enquiry within the A4EA programme exploring the role of external actors in generating empowerment and accountability outcomes, and informed by a 2016 paper written for the programme on the importance of multi-scalar approaches (Fox 2016) we sought to explore:

- The degree to which DFID country-level programmes in FCVAS are taking scale into account in their design, particularly through multi-level and vertically and/or horizontally integrated approaches.
- The ‘theories of change’ the programmes were using or seeking to activate; how they envisaged their actions would lead to increased accountability in the short and long-term.
- How the programmes monitored and evaluated their use of multi-level and vertically integrated approaches.
- What could be learnt about using and strengthening multi-level and vertically and/or horizontally integrated approaches in FCVAS from the delivery of these programmes.

We explored these in a number of ways, including systematic review of documentary evidence, interviews with programme stakeholders in a series of short field visits to three countries, and a stakeholder event bringing together DFID staff, implementers of these kinds of programmes, and our research teams. These are outlined briefly here before a wider discussion on the extent of the evidence this approach has enabled us to gather and its limitations.

### 3.1 Research approach

We used publicly available information on DFID programmes in Mozambique, Myanmar, Nigeria, and Pakistan to identify an initial set of activities for further exploration. We used official programme summaries from DFID’s public disclosure website ‘Development Tracker’ to identify programmes that were reported as seeking to promote outcomes of accountability, empowerment, participation or transparency, that had been active in these countries between 2011 and 2016. This date range was chosen on the assumption that more recently initiated programmes may not be sufficiently established to have generated evidence to review. From the 21 resulting programmes an initial review of documentation and discussions with experts in those contexts narrowed our focus to ten programmes. The most common exclusion factor was for programmes that had public accountability aims but

activities that focused largely on reforms of state institutions with limited activity engaging citizens directly. Two programmes were excluded as they had been completed some time before fieldwork. Various other exclusions were made for relevance; for example, the MUVA programme in Mozambique was excluded as although it sought to promote women’s empowerment it did not also have a focus on accountability.

For this sample of ten programmes we then created a detailed database of programme documentation and evidence. Documents were gathered from the DFID aid transparency portal,<sup>5</sup> the database previously created to support DFID’s macro-evaluation on voice and accountability programming (e-Pact 2016), the websites of those involved in delivering the programmes, and general searches for online material, including references to these programmes in academic publications. These were supplemented over time with the addition of documentation and other evidence from organisations implementing programme activities and contacts within DFID field offices. The research team undertook a structured review of the available documentation using variables based on the research objectives outlined above; for example, identifying expressions of the programme theory of change and noting how multi-scalar activity was included in the monitoring and evaluation schemes. This process of refinement and documentary review informed the selection of six programmes for further focus.

Table 3.1 below illustrates the narrowing from the original sample to the programmes we focus on in this paper.

**Table 3.1 Programme case selection – indicating inclusion in documentary review and case study activities**

Country	Programme (acronym or used name in brackets)	Start year	End year	Budget (DFID)	Doc. review	Case study	Exclusion reason
Pakistan	AAWAZ Voice and Accountability Programme	2011	2018	£39m	Yes	Yes	
	Consolidating Democracy in Pakistan	2015	2019	£27m	No	No	Limited citizen-focused activity
	Empowerment, Voice and Accountability for Better Health and Nutrition (EVA-BHN)	2012	2019	£18.9m	Yes	Yes	
	Strengthening the Rule of Law in Pakistan	2016	2020	£12m	No	No	Limited citizen-focused activity
	Supporting Transparency Accountability and Electoral Processes in Pakistan	2010	2014	£11.7m	No	No	Completed well before study period.
	Transforming Education in Pakistan (Alif Ailaan)	2013	2018	£25.3m	Yes	Yes	
	Khyber Pakhtunkhwa Education Sector Programme	2011	2020	£264m	No	No	Limited citizen-focused activity

(Cont’d.)

<sup>5</sup> [www.devtracker.dfid.gov.uk](http://www.devtracker.dfid.gov.uk)



**Table 3.1 (cont'd.)**

<b>Nigeria</b>	Facility for Oil Sector Transformation II (FOSTER II)	2015	2021	£19.5m	Yes	No	Limited citizen-focused activity
	Facility for Oil Sector Transparency (FOSTER)	2010	2016	£13.2m	Yes	No	Limited citizen-focused activity
	Partnership to Engage, Reform, and Learn (PERL)	2016	2021	£100m	No	No	Studied separately in research programme
	State Accountability and Voice Initiative (SAVI)	2008	2016	£33.8m	Yes	No	Completed well before study period.
	Justice for All	2008	2017	£52m	No	No	Limited citizen-focused activity
<b>Mozambique</b>	Civil Society Support Mechanism	2015	2018	£5m	No	No	Focus on funding national level advocacy.
	Democratic Local Governance Support Programme (Diálogo)	2012	2017	£9.6m	Yes	Yes	
	Muva	2014	2022	£18m	No	No	Focus on economic empowerment only.
	Strengthening Citizen Engagement in Mozambique (CEP)	2012	2017	£10m	Yes	Yes	
<b>Myanmar</b>	Burma Civil Society Support Programme (BCSSP)	2011	2016	£13.6m	Yes	Yes	
	Improving the management of public funds for the benefit of people in Burma	2014	2019	£20m	Yes	No	Limited citizen-focused activity
	Programme for Democratic Change	2014	2019	£24m	No	No	Limited citizen-focused activity
	Supporting Partnerships for Accountability and Civic Engagement in Myanmar (SPACE)	2017	2021	£10.3m	No	No	In inception phase, too early to review
	Supporting Peace-building in Burma	2015	2020	£25.5m	No	No	Largely investments in multi-donor initiatives.

Source: DFID Development Tracker public disclosure website, except EVA-BHN programme, where financial data was drawn from a DFID-supported programme assessment (Taylor and Khan 2018) as EVA-BHN is a sub-component of the DFID programme, 'Provincial Health and Nutrition in Pakistan' and the budget is not disaggregated in programme documentation.

Notes: Budgets indicate DFID contributions, which may be matched with other donor contributions in some cases to increase the overall spend. These budgets include often substantial sets of activity not discussed in this paper given our specific focus on multi-scalar strategies.

Supplementing our documentary review, brief periods of fieldwork on the six case study programmes took place in Pakistan (December 2017), Mozambique (June 2018), and Myanmar (June 2018). Researchers interviewed DFID field office staff, practitioners involved in delivering programme activities, wider stakeholders in the development sector, and in Pakistan some project beneficiaries in the form of local community forums. Interviews sought to understand the context and contextual fit of the programmes and their multi-level approaches, the perceived successes and challenges of these approaches, and views on

their outcomes. These discussions did not seek to evaluate the programme activities in the sense of evidencing outcomes or their attribution to particular activities, but instead to provide additional evidence and insights into how multi-scalar strategies had operated in practice. A report from the first period of fieldwork (Esser 2018), comparing the case study programmes in Pakistan, was developed to inform the subsequent visits and interviews. There were differences in project status in each country. In Pakistan all the programmes explored in fieldwork were currently operational to differing degrees, although two were in their final stages. In Mozambique the two programmes chosen had closed almost six months previously and a new programme based on those experiences was under design. In Myanmar the study focused on activities under a programme that had concluded more than a year earlier, but some components of the work were continuing under a new programme with similar stakeholders and implementing organisations.

Findings from documentary review and fieldwork studies were shared with an audience of DFID staff and programme practitioners in an IDS workshop in July 2018 that aimed to further explore how our selected programmes and others like them worked in practice, and to test findings so far. The audience included several practitioners who had been employed in or involved in the design or evaluation of the specific programmes we studied. Alongside discussion of initial findings, the workshop provided wider opportunities for practitioners to reflect on trends in development programming of this kind and discuss challenges and learning from that programming from their own perspectives.

### **3.2 Availability of evidence**

DFID considers aid transparency a high policy priority (DFID 2018b) and provides a specific online portal designed to increase the transparency of UK development assistance. However, we found significant variation in the quantity and quality of project information that we were able to gather directly through this portal to support our documentary review. We explore here what information we could gather and how informative generally we found that material.

In DFID programme management practice, the business case sets out the strategic reasons for proposed activities, the aims of activities, and the approach to delivery. At the point at which data was originally gathered, two of the ten programmes did not have a published business case, although all became available during the study. Business cases available were on the whole informative and helped understanding of both the rationale and design of the programme concerned. Six of the programmes had some form of original contract documentation published after this initial business case, which often showed how the original aims were translated into a refined set of activities. All programmes had some annual reviews of programme performance published, but for many this did not cover all of the years of operation. There was only one programme that did not have a published results framework or logical framework document setting out how programme outcomes were measured. However, in many cases these were not recently updated. Four of the seven programmes that had finished had a programme completion review available online, two of which became available during the research process, and one was published online a year after the programme had reportedly ended. None of the five completed programmes had final evaluations available on the portal, and only one programme had an independent mid-term review available as of September 2019. During the course of our research we acquired a range of other official documentation from programme managers and implementers. A summary of document availability is presented in Table 3.2, where shaded boxes indicate availability at the time of the document review, with an 'x' indicating where these were shared with us but are not publicly available.

With very few exceptions the transparency portal only includes documentation 'owned' by DFID. Programme implementers had often generated much more extensive documentation,

which DFID field offices had access to. For example, whilst the portal makes available the Annual Reviews of programme activity these contain much less detail than the annual reports that programme providers submit to DFID. This is of particular interest from a research perspective as these annual reports often revealed data and regularly collected metrics on project performance that were not included in the DFID Annual Reviews. Notably, very few of these documents were available on the programme websites, where these existed. In fact, provider organisations generally published even less material about the work of particular programmes online than DFID.

**Table 3.2 Availability of DFID programme documentation**

Programme	Logical framework	Business case	Inception report	Independent mid-term review	Annual review	Project completion reports	End evaluation
AAWAZ	X		X			X	
ALIF-AILAN						X	
EVA-BHN						n/a	n/a
FOSTER							
FOSTER II						n/a	n/a
SAVI							
CEP				X			X
DIÁLOGO			X	X			X
BCSSP				X			
IPFPB						n/a	n/a

Source: Authors' own analysis as at September 2018, building on work by Suchi Pande, based on available documentation on <https://devtracker.dfid.gov.uk>

Note: Shaded boxes indicate documents that we could review. Those marked X were not publicly available. Unshaded boxes indicate documents that could not be reviewed, with n/a noting where this document would not have been expected to be available as programmes were ongoing at the time of study.

We found that documentation varied significantly in detail and quality of explanation and discussion – in particular, Annual Reviews. Project Completion Reviews were more consistent in their quality of reflection on activities, their outcomes, and the evidence for their judgements. The greatest difficulty in using this documentary material came where the activities we wanted to study were part of a larger and more diverse DFID programme. It is fairly common for programmes to have multiple components, but those with fewer and more directly aligned activities were easier to assess from the official documentation. The greatest impact of this was for the EVA-BHN project in Pakistan, which was a component of a much larger health sector support programme and was covered fairly minimally in reports on that programme.

A final point on availability of evidence relates to programme evaluation specifically. DFID practice appears to be to publish programme evaluations not on the pages for each programme on the transparency portal, but in a series of separate web pages that are

published annually.<sup>6</sup> Two final evaluations commissioned by DFID were made available to us by DFID field offices, and at the time of writing these had also not been published on either the transparency portal or the evaluation collection webpage. One programme was reported by the DFID field office to have an evaluation completed but this was not made available (although an independent mid-term assessment of this programme was shared). One DFID-commissioned independent assessment of a programme, not presented as an evaluation, became publicly available during our study (Taylor and Khan 2017), and is available online (although it was not added to the transparency portal or the evaluation collection webpage). Four programmes that had completed or were near completion did not appear to have independent evaluations undertaken or ongoing. This overall coverage of programme level evaluation supports the finding of the comprehensive review of DFID empowerment and accountability programmes undertaken in 2016, which worked with a sample of 50 programmes in depth, and concluded that the availability of evaluations was rare for these programmes (e-Pact 2016: 20). The direct implication of this is that our analysis in the later sections of this paper rely for some programmes on internally produced documentation and reports from local practitioners and stakeholders, rather than on evidence that has been more independently validated.

In contrast to the volume of suitable evidence on programme activities available publicly and in DFID documentation, during our fieldwork we were provided with evidence of and told about significant and sophisticated monitoring systems used by programme implementers. We comment further on this gap between programme internal and external evidence generation in Section 8 of the paper, which also discusses the implications of patchy available evidence and what this means in terms of learning from this significant set of DFID-supported activities. First, however, we introduce the case study programmes, and look at what this evidence does enable us to say about their similarities in design and practice.

## 4 The case study programmes

The programmes we focused on all had explicit empowerment and accountability aims in contexts characterised by fragility, violence, and conflict. Some were explicitly governance-focused programmes whilst others sought to generate improvements in specific sectors or for specific groups. Four programmes were completing their work during the research period, one was completed some time before, and one was continuing. They had a combined investment from DFID of £116 million, and were all planned as at least five year programmes.

Table 4.1 below summarises these factors. Before moving to comparisons, the rest of this section gives a brief introduction to each programme, outlining the context in which they were developed, their general aims, activities, and reported outcomes drawing from programme documentation and available evaluations.

---

<sup>6</sup> See, for example, the following page with evaluations completed in 2017–18: [www.gov.uk/government/collections/evaluations-completed-april-2017-to-march-2018--2](http://www.gov.uk/government/collections/evaluations-completed-april-2017-to-march-2018--2)

**Table 4.1 Case study programmes**

Acronym/ used name	Programme name	Country	Duration	Budget (DFID)	Main sectors
<b>AAWAZ</b>	AAWAZ Voice and Accountability Programme	Pakistan	2011–18	£39m	Governance; Gender; Conflict
<b>Alif Ailaan</b>	Transforming Education in Pakistan	Pakistan	2013–18	£25.3m	Education
<b>EVA-BHN</b>	Empowerment, Voice and Accountability for Better Health and Nutrition	Pakistan	2012–19	£18.9m	Health
<b>CEP</b>	Citizen Engagement Programme	Mozambique	2012–17	£10m	Education; Health
<b>Diálogo</b>	Democratic Local Governance Support Programme	Mozambique	2012–17	£9.6m	Governance
<b>BCSSP</b>	Burma Civil Society Support Programme	Myanmar	2011–16	£13.6m	Governance

#### 4.1 AAWAZ Voice and Accountability Programme – Pakistan – 2011–18

Pakistan’s ‘largest demand-side voice and accountability programme’ according to DFID documentation, (DFID 2018c: 1), AAWAZ operated in the two provinces of Punjab and Khyber Pakhtunkhwa and was nearing completion during our fieldwork. AAWAZ built a multi-level network of spaces and opportunities for citizens to engage together and with officials and service providers. It used a variety of approaches that sought to empower previously marginalised people to raise demands and engage in governance decision-making, with a particular focus on women. Activity included widespread training, information provision, and maintaining local centres where people could connect, including a network of more than 400 ‘Aagahi Centres’, which were reportedly accessed by 100,000 women and girls during the life of the programme (DFID 2018c: 15). These were complemented by a deliberate strategy of establishing spaces for programme participants to engage with officials, including a reported 2,200 open town hall meetings and meetings of district-level committees (DFID 2018c: 16).

The DFID programme completion review reported positively on the outcomes of AAWAZ, suggesting that it ‘provided a robust citizen-state interface where government officials and elected representatives were held accountable to women and the socially excluded’ (DFID 2018c: 1). It reports that more than 27,000 improvements in services were made in response to community level demands (DFID 2018c: 17), that 159 instances of policy and implementation changes that came about could be attributed to the activities (DFID 2018c: 16), and that six pieces of provincial legislation were directly influenced by the programme (DFID 2018c: 17). At completion positive improvements were also reported in terms of citizens’ capacities and willingness to challenge restrictive gender norms and hold the state to account on an ongoing basis (DFID 2018c: 22).

#### 4.2 Alif Ailaan – Pakistan – 2013–18

Alif Ailaan (‘Education Declaration’) was the major component of the DFID programme, Transforming Education in Pakistan. Presented as a political campaign it sought to increase pressure on politicians and service providers to improve education services and outcomes through establishing a broad-based coalition that could ‘galvanise political will’ at various levels of government to see through reforms (DFID 2017b: 1). The explicit focus on incentivising and holding accountable political power-holders was seen as a novel approach in the context of other donor-funded accountability activities in Pakistan focusing on

generating direct community level engagement with service providers, rather than politicians (DFID 2017c: 8). Alif Ailaan identified and provided funding to local organisations to play different roles in campaign-building, and sought to catalyse activism through provision of information and guidance and recruitment of campaign mobilisers in what it called a 'movement building' strategy (DFID 2017b: 1). It also undertook detailed studies of the education system combining existing government data and new research, increasing the evidence of education system performance and disparities. Another key approach was maintaining a high media presence, often sharing these data with a 'naming and shaming' intention in traditional and social media.

The DFID programme completion review claims that the campaign 'catalysed increased government investment in education at the federal and provincial level and ensured reform continued and was prioritised by the next government' (DFID 2018f: 1). Esser found many supporters of its approach amongst practitioners involved in empowerment and accountability programmes in Pakistan, and argued that the programme held 'considerable promise in terms of creating lasting change, not only regarding education outcomes but also more broadly to forge greater accountability' (Esser 2018: 18).

### **4.3 Empowerment Voice and Accountability for Better Health and Nutrition– Pakistan – 2012–19**

EVA-BHN was part of the broader Provincial Health and Nutrition Programme (PHNP) that aimed to improve reproductive, maternal and nutrition health outcomes in the two provinces of Khyber Pakhtunkhwa (KP) and Punjab in Pakistan. EVA-BHN was explicitly presented as the 'demand-side' element of a programme that was otherwise largely acting to build technical and delivery capacity in public services. EVA-BHN sought 'to empower, organise and facilitate citizens and civil society to hold duty-bearers in Punjab and KP to account' for service provision (Palladium 2016a: 4). The programme supported nearly 400 community-level groups at Union Council level in nine districts, that had engaged almost 40,000 people by September 2017 (DFID 2018a: 9). The groups served as the focal point for engaging people in rights education and advocacy training, as well as spaces for collective deliberation and to raise and prioritise issues drawn from citizen-based monitoring activities. New dialogue forums with service providers and other local leaders were established at district and provincial levels to complement this facility-level activity. EVA-BHN also worked on public communication activities intended to improve health-seeking behaviour.

An independent assessment held near the end of the programme found evidence that the activities had contributed to better service provision and issue resolution at a local level, and better 'cross-party dialogue and collaboration' and responsiveness to citizen experience at the district level (Taylor and Khan 2017: 9;13). A separate study of the programme undertaken at the request of the implementing team and published by the A4EA programme, established how it had deployed multiple tactics at different levels in ways that illustrate more strategic approaches to social accountability (Kirk 2017). Impact studies produced by the programme itself highlight diverse pathways of activism and influence that led to issues being resolved beyond the immediate local health centre (Esser 2018). Various broader policy and practice outcomes of the work were reported (Palladium 2017: 18). These were confirmed to a degree by the independent assessment, although this also expressed concerns about how far the activities addressed the systemic challenges in the Pakistan health sector (Taylor and Khan 2017: 24, vi).

### **4.4 Citizen Engagement Programme – Mozambique – 2012–17**

The Citizen Engagement Programme (CEP) was established with the rationale that central technical reforms and additional financing were insufficient to generate improved public services in Mozambique, and that strategies to increase citizen demand were needed. It worked specifically in the education and health sectors promoting the use of social

accountability tools and ‘co-governance’ arrangements that involved community members in the management and oversight of local schools and health facilities. These arrangements already existed in theory, but the programme aimed to demonstrate how they could work more effectively and inclusively through links with participatory community engagement. The activity covered four of the 11 Mozambican provinces and within these: 119 primary schools, 12 secondary schools, 34 health posts, and 2 district hospitals (DFID 2018d: 4). Through a network of district civil society partners, community members were trained as facilitators of citizen engagement processes. These processes were supported by programme management staff at the provincial level, and national civil society advocacy partners in the education and health sectors. Mechanisms were created to aggregate and analyse the kinds of issues raised at the local level with the aim of influencing national policy.

At the end of the programme, 56 per cent of schools and 65 per cent of health facilities engaged were judged to have made changes directly linked to citizen monitoring (DFID 2018d: 7). A detailed programme evaluation found these largely to result from local adjustments, ‘galvanised’ co-management practices, and community contributions rather than systemic changes (Brook *et al.* 2017: vi). Evaluation and programme final reporting also described national level policy gains based on the evidence from the activity and the ‘proof of concept’ of citizen-based monitoring and service co-governance (Brook *et al.* 2017: 28–33).

#### **4.5 Diálogo (Democratic Governance Support Programme) – Mozambique – 2012–17**

The Diálogo programme aimed at improving citizen-state relationships and interactions in Mozambican urban governance, focusing on municipal governance institutions. Part of the rationale for this focus was that efforts to increase accountability in central government institutions had demonstrated little impact (DFID 2012b: 1). Working in five specific urban centres, it worked with local civil society actors, elected politicians, and officials to test out approaches to participatory governance and establish more civic deliberation and independent oversight. Mechanisms adopted included participatory budgeting, citizen score-card processes, participatory approaches to wider municipal planning, and public engagement with particular services and decisions. The programme also supported civil society organisations to build their capacity to participate in decision-making and facilitate wider community participation, and developed rights and governance awareness training for target communities. Alongside these, ran activities supporting independent media outlets to prompt public dialogue on issues of local governance.

The external evaluation of the programme found significant impacts on the capacities of citizens to act on increased understanding of their rights and greater openness towards, and capacity to undertake, participatory governance approaches in the municipalities involved (Shankland *et al.* 2017: 4). The programme completion review reports that the activities ‘strengthened voice and accountability by transforming the institutions and power relations in each of the five municipalities’ (DFID 2018e: 8). The role played in this by more empowered civic organisations, and constructive engagement with reformists in local governance was particularly emphasised.

#### **4.6 Burma Civil Society Strengthening Programme – Myanmar – 2011–16**

The Burma Civil Society Strengthening Programme (BCSSP) was developed in 2010 as Myanmar was undergoing a slow transition from military to elected civilian rule. Responding to a perceived fragmentation and lack of cohesion within civil society (DFID 2011b: 1), the programme funded a number of different kinds of support to build the capacity of civil society groups so that they would be more effective advocates for reform (DFID 2011b: 8). The programme focused largely on establishing the role of civil society organisations within the accountability eco-system, and creating ‘enabling environment’ conditions for their future engagement in policy reform. A number of projects were funded under the programme,

delivered by different organisations and with different contributions to the overall theory of change and action. These included organisational development support and core funding to a group of national-level organisations, and investment in social infrastructure to provide tailored capacity and network-building through the organisation Paung Ku ('Bridging'). One of the projects, Pyoe Pin ('Green Shoots'), worked specifically to form and financially support issue-based coalitions across civil society and government based on analysis of relatively more open policy spaces and tractable reform issues, and has been seen as an exemplar for politically-informed development programmes (Christie and Green 2018).

Programme documentation references many legislative gains and institutional behaviour changes arising from CSO-led policy advocacy (British Council 2016a; DFID 2016b: 2). The programme completion review found that Pyoe Pin activity specifically facilitated stronger engagement between local civil society actors and policy processes (DFID 2016b: 10) and credited the ability of some of the civil society coalitions formed to 'critically engage' with government (DFID 2016b: 2). However this review also noted the difficulty in evidencing the contributions of the activities to policy changes given the rapid and to some degree unanticipated opening of civic space associated with changing political configurations during the programme delivery, and inconsistent outcome measurement (DFID 2016b: 22).

These brief descriptions highlight the different histories, contexts and activities of the programmes chosen as case studies, most of which were operating simultaneously across our three countries of focus. A number of similarities also emerge. The next two sections look at these similarities in more depth, first in terms of their rationales, logics, and contextual analyses, and then in their adoption of similar multi-scalar strategies.

## 5 What were the programmes trying to do, and why?

Whilst the contexts and political economies in which these programmes were working differed both nationally and sub-nationally there were clear similarities in how their underpinning analyses portrayed the background conditions for public accountability. The programmes all acknowledged multi-scalar dynamics and responded to these in part in their design. They also showed similar theories of change and drew from similar repertoires of intervention.

### 5.1 Diagnoses of accountability conditions

The case study programmes shared several common diagnoses of the drivers of low state accountability. Key amongst these is a view of fairly limited organised citizen demand for better services and claims for accountability. For example, the Diálogo business case argues this is a 'major barrier to increasing accountability' (DFID 2012b: 11), whilst the rationale for the EVA-BHN set of activities calls existing citizen demand 'ineffective' (DFID 2012c: 7). The BCSSP business case explains the need for a 'more active citizenry' to take advantage of the transition to democracy in Myanmar (DFID 2011b: 4). Various reasons for this lack of citizen demand are given. One is the exclusion of many people from governance processes. For example the AAWAZ business case argues that 'the social exclusion of the majority of the population from politics and governance must be addressed' in order for Pakistan to achieve global development goals (DFID 2011a: 4). In common with the business case for Alif Ailaan, it argues this results in the absence of a social contract between government and citizens (DFID 2011a: 5, 2012d: 11). All programme designs note a particular exclusion of women from accountability processes. This is explained in part through references to social norms that 'restrict access to the public arena' (DFID 2012a: 6), low priority placed on



women's experiences and wellbeing (DFID 2012c: 5), and underlying gender-based discrimination that is expressed in a lack of access to economic as well as political resources (DFID 2011b: 3).

Another common diagnosis is of distance between citizen, activists, and reformers at different levels. For example the CEP business case noted 'limited linkages between national [civil society] organisations and those representing or comprised of local interest groups, and between 'professionalised' groups and membership organisations' (DFID 2012a: 5). The BCCSP business case goes into some detail on what it calls the 'fragmentation' of existing civic associations, CSOs, and NGOs (DFID 2011b: 4–5). In analysing the extent of civic space, the role of the media in maintaining public scrutiny and deliberation were often covered. For example the Diálogo programme, which made strengthening independent media a significant component of its work, analyses in some depth the need to incentivise investigative journalism linked to citizens' concerns in Mozambique (DFID 2012b: 6). Alif Ailaan, also with a strong media component to its work, notes existing signs that some sections of the media in Pakistan are able to play a campaigning role with some influence (DFID 2012d: 8–9).

Finally, most of the business cases note significant resource and technical constraints on public services. The EVA-BHN business case notes the risk of 'demand outstripping supply' should changes not also be funded and supported within services (DFID 2012c: 12). This is fitting for a mixed intervention working both with service providers and to surface community-level priorities. In some contrast to this, the lack of success of previous reform efforts that focused only on 'supply-side' measures is mentioned in making the case for Diálogo (DFID 2012b: 1) and EVA-BHN's focus on building and facilitating citizen demand (DFID 2012c: 7). Documents relating to the design of AAWAZ state that 'efforts to bring about change by focusing exclusively on supply-side governance have largely failed' (DAI 2015: 3).

## **5.2 Multi-scalar and multi-sited governance**

All of the programmes reviewed acknowledged a multi-scalar governance landscape in some way in their analysis and design. It was common for them to discuss the different 'levels' at which programmes would need to work, often focusing on the political-administrative hierarchy in that context. The business case for AAWAZ notes that '[e]xcluded groups, including women, will need to feel safe enough to become organised and involved in politics, local affairs and governance processes at all levels (community, district, province, nation)' (DFID 2011a: 10). Design options for CEP were assessed for their ability to empower citizens 'to dialogue with service providers and the state at different levels..' (DFID 2012a: 6). The BCCSP business case placed a particular focus on the need for activity to build both vertical and horizontal linkages amongst civil society actors, with these being key criteria for selecting a programme design (DFID 2011b: 12).

There is some discussion in the programme documents on where power lies across these levels and different actors. This involves some apparent contradictions; emphasising both centralisation and concentration of official decision-making, and simultaneously the informally important role of a wide range of fairly dispersed actors and institutions. Highly concentrated formal decision-making at the national government level is highlighted in business cases for Mozambique (DFID 2012a: 11, 2012b: 8), in the mid-term evaluation of the BCSSP in Myanmar (DFID 2013a: 46), and in documentation of the EVA-BHN approach in Pakistan (Palladium 2016a: 7), although in this case power is seen as centralised at provincial level in Punjab. The business case for Diálogo not only notes the dominantly centralised decision-making in Mozambique but emphasises this as part of a rationale for attempting to develop better governance practices at municipal level (DFID 2012b: 8). At the same time, a common diagnosis is that low citizen engagement with the state is driven by patronage-based governance (DFID 2011a: 9, 2012a: 2, 2012d: 9), informal power-holders and institutions

such as local elites and political intermediaries such as party brokers (DFID 2011a: 9, 2012a: 4; Palladium 2016b: 6), and the role of the military (DFID 2011a: 9, 2011b: 1). The Alif Ailaan business case particularly highlights the role of military elites in resource control (DFID 2012d: 7). Both Mozambican programmes note in particular the blurred boundaries between the ruling party and the state, which is seen to undermine opportunities for citizen participation, and the sensitivity of the party to any criticism, which is typically framed as support for the political opposition (DFID 2012a: 4, 2012b: 6-7).

### 5.3 Theories of change and intervention

Programme design documents consistently view more inclusive governance developing over time through the building of countervailing power outside of existing institutions and elites. In keeping with the language of the DFID over-arching theory of change for governance programmes (DFID 2011c), programme analyses often use the notion of political settlements existing in which ordinary people hold little sway over governance decisions. Programme rationales posit the possibility for donor-supported action that can support a broadening of the political settlement, most notably through a more empowered citizenry that might shift incentives or challenge the status quo. The BCCSP business case argues that whilst change in the political settlement will ultimately 'come from within', the programme can 'equip organisations and coalitions with the capacity to work towards a more plural and democratic society' where opportunities arise (DFID 2011b: 8). The programme completion review of the AAWAZ programme for example credits it with 'breaking state monopoly of 'official executive oversight'' (DFID 2018c: 23).

A number of strategies for citizen empowerment are covered in the common repertoires of the programmes. These include activities designed to raise awareness of rights and entitlements (AAWAZ, CEP, Diálogo and EVA-BHN in particular), development of spaces and structures for deliberation and discussion of preferences and to agree collective strategies, and creation of structures or opportunities to channel and legitimate expressions of individual and collective agency. This latter point is important given the histories of repression and violence by state actors, and the need to provide 'safe' avenues through which grievances and demands might be raised. The outcomes anticipated of this more empowered citizenry are described in terms of improved services, increased access to services and entitlements, and in terms of better citizen-state relationships. For example, the DFID Terms of Reference for the EVA-BHN programme posits that 'through greater pressure and demand from... citizens' the programme would be able to 'influence health legislation, policy, practices, and service delivery' (DFID 2014: 2). The Diálogo business case imagines that through support to local groups and municipalities 'women, young people, and men will increasingly participate in a meaningful way in the decision-making and political processes at a local level' (DFID 2012b: 3). Most ambitiously perhaps, the BCSSP business case posits that 'the more that people and organisations work together, in constructive ways, to achieve tangible changes, the more they will strengthen the role of citizens in building a functioning democracy' (DFID 2011b: 3-4).

## 6 What multi-scalar strategies were used?

Having established a core of similar diagnoses and underlying logics, we now move to explore how programme activities responded to the recognition of multi-level governance and the importance of scale. Three distinct uses of multi-scalar strategies were distinguishable in the case study programmes:

- **Direct vertical aggregation of citizen demands and claims on authorities.** Establishing mechanisms to formulate and gather citizens' preferences, priorities and experiences and convey these upward to those responsible for public services.

- **Building vertical integration of civil society organisations and campaigns.** Undertaking activities that linked civic action and organisations at the grassroots with sub-national and national level civil society activism and lobbying.
- **Developing horizontal linkages and networks.** Incentivising and creating conditions for organisations and reformists working across geographies or sectors to find common ground and act jointly.

These have different logics that may be complementary but may also involve some trade-offs in terms of where to direct effort. Some programmes involved aspects of all three, whilst some focused more on one than others. Table 6.1 below illustrates our analysis of this, which suggests that programmes each had a dominant focus on one of the three strategies.

The remainder of this section discusses each strategy in turn, illustrating their uses within the case study programmes, and highlighting some key characteristics. It is important to note that all of the programmes we looked at include significant other components or repertoires of intervention – either integrated into the work that we focus on in this paper or in distinct projects. They all sought to work to some degree on wider aspects of the enabling environment for accountability claims from citizens, for instance. These include seeking to hold or expand civil society space, incentivise public media reporting that is pro-accountability, in one case working with religious leaders, and also developing knowledge and skills amongst citizens in terms of rights and rights-claiming. In many cases they sought to marry their work on citizen empowerment, mobilisation, and action with action to build state or service capacity to respond more fully to the citizen demands. In the following analysis we refer to these only where these activities were multi-scalar in nature.

**Table 6.1 Use of multi-scalar strategies in case study programmes**

Programme	Role of identified multi-scalar strategies		
	Vertical aggregation of citizen demands	Vertical integration of CSOs and campaigns	Horizontal linkages and networks
AAWAZ	Major	Minor	Minor
Alif Ailaan		Major	Minor
BCSSP		Major	Minor
CEP	Major	Minor	Minor
Diálogo		Minor	Major
EVA-BHN	Major		Minor

Source: Authors' assessment

### 6.1 Direct vertical integration of citizen demands and claims on authorities

In three of the programmes we reviewed a defining characteristic of their design involved strategies of mobilising people around specific local needs and public service delivery quality and aggregating these issues or demands vertically. To different degrees and in different combinations the AAWAZ, EVA-BHN, and CEP programmes all attempted to vertically integrate versions of common social accountability tools: community score cards, citizen-based monitoring, public hearings or 'town hall' meetings, and 'co-management' mechanisms like school management committees. The programmes saw these structured multi-level processes as a way to get beyond the most local of issues and bring evidence and challenge to decision makers elsewhere in the system. This is important given the whole range of decisions that are beyond the power of frontline service delivery units and personnel, the systemic causes that lie behind many service deficiencies, and problems that have their root

in overall public investment in services. They designed and operated their multi-level processes in different ways.

The activities of EVA-BHN responded deliberately to the scholarly evidence base on social accountability (Kirk 2017; Palladium 2016b) and structured vertical integration of demands and citizen-based monitoring evidence was a 'defining component' of the programme (Esser 2018: 16). The vertical structure of citizen-led groups it built mirrored local governance decision-making structures in Pakistan, and the direct involvement of citizen representatives at multiple levels was a distinctive part of the programme implementation. Union-council level groups were used to surface and explore experiences of using primary health services and highlight issues to be addressed or specific grievances, with people from these groups taking these to local health providers. District-level advocacy forums brought those participating together with higher level health sector decision makers and wider civil society representatives, and were used to both review the general patterns of issues coming up at the Union Council level and to raise specific problems or demands that were not able to be resolved locally. Sometimes these were issues that were not in the power of local services to resolve – for example allocation of equipment or staff. In other cases, they were issues with local providers and staff who were not being responsive to community feedback – for example staff not following procedures or providing poor quality care. This 'demand escalation' model was found in an independent assessment to have had good effect in many cases, with district authorities able to step in to support changes (Taylor and Khan 2017: 13–14). The programme also convened provincial-level advocacy forums that engaged authorities and influential political figures. These discussed the pattern of demands arising from communities, attempted to draw attention to both the structural issues in the health system that underpinned these, and promoted the value of citizen or service user feedback in generating greater service responsiveness.

AAWAZ worked in the same provinces of Pakistan. Facilitated by local NGOs the programme established a multi-level set of groups that in design took a 'deliberately cross-scalar' approach (Esser 2018:14). A broad base of women's community centres and youth groups provided spaces for citizens 'to formulate priorities and voice resulting demands' (Esser 2018: 14). Issues and demands from these groups were raised and discussed in district and provincial level fora with government representatives including in 'town hall' meetings, in a model reportedly designed to 'aggregate and escalate' issues to 'higher tiers of government' (DAI 2015: 6). DFID programme documentation reports that nearly 30,000 'demands' were raised with public officials as a result of these activities and that the programme used provincial forums to 'tangibly influence' provincial legislation (DFID 2018c: 17). There is little illustration or evidence in programme documentation however of how these changes and the demand aggregation system were linked, and how and by whom representations to decision makers were made. In our fieldwork some local practitioners expressed scepticism about the ways that these volumes of activity were measured, documented and incentivised (Esser 2018: 16), but confirmed that the programme operated in a clear vertically articulated way – noting that the programme was 'trying to form a vertical structure... by creating local councils, then sending people upward' (Esser 2018: 14).

CEP focused on citizen demands regarding health and education services in rural Mozambique. Activities at a 'facility' level – mostly primary schools and primary health centres – used community score cards and to a lesser extent citizen report cards to feed into what the programme termed 'co-governance' arrangements; school and health centre committees that were mandated in some way to have community involvement. The findings of community level scorecards and prioritisation were collated and aggregated upwards by the CSOs that led on delivering the programme at a local level to district and provincial levels. Through engagement and funding of established national civil society advocacy groups in each sector, with access to a database established of issues raised in facility-level scorecards, this evidence was then fed into national policy debates. An independent

evaluation of the programme confirmed that the programme generated numerous 'facility-level improvements' but attributed these largely to the connections between community scorecards processes and the co-governance mechanisms that allowed the outcomes to be prioritised in facility-level planning (Brook *et al.* 2017: vi). It concluded that the programme had struggled with a 'lack of vertical responsiveness' to demands taken to the district and provincial levels (Brook *et al.* 2017: v).

## **6.2 Building vertical integration of civil society, campaigns and reform coalitions**

In addition to working directly with state or other decision-makers, several of the case study programmes used multi-level approaches in establishing or strengthening vertical networks of civil society organisations and advocacy campaigns. This was an important strategy in the activities of the BCSSP in Myanmar and the Alif Ailaan programme in Pakistan. It was less central but also used in the approach of CEP in Mozambique and AWAAZ in Pakistan. These activities aligned with programme diagnoses of 'fragmented' or 'piece-meal' initiatives and activism by civil society organisations (DFID 2011b: 8,10, 2012a: 11, 2012d: 27).

Positioned overtly as a campaign with movement building and advocacy logics at its core, the design of the Alif Ailaan programme recognised the need to connect constituencies vertically to generate the kinds of pressure across the education system that might stimulate change (DFID 2012d: 7, 21, 27). The campaign approach sought to catalyse popular demand for better public education, rather than specific localised changes in particular schools noted to be achieved by previous targeted social accountability activities (DFID 2017c: 8). Alongside support to activists and direct lobbying activities, the programme worked extensively on generation and packaging of spatially disaggregated data on the education system. Whilst much of this work used government produced data, which Esser (2018: 17) argues 'shielded' it from political criticism to a degree, the campaign also funded new research to counterpoise official data. For example, a 2015 study of one province showed that a high number of officially listed schools in one district were not in fact open (Alif Ailaan 2015). Evidence from this data analysis was shared with a network of activists engaged at various levels and used in media campaigns and lobbying with politicians at provincial and national levels. The campaign had a specific focus on appealing to politicians to overcome a perceived lack of political commitment to investing in and reforming public education. This included attempts to achieve commitments and undertakings from politicians and political parties that could be the basis for later accountability claims (DFID 2017b: 5). The programme completion review argues that this vertically connected multi-level approach was instrumental in the activities achieving greater political prioritisation and resourcing of the education system (DFID 2018f: 26).

Unlike Alif Ailaan, the multi-level strategy of the BCSSP did not focus on a single sector, but on the ability of civil society organisations in Myanmar to act collectively and form multi-stakeholder and multi-level coalitions around a number of salient issues. One of the strategies employed as part of this broader approach, largely led by the Pyoe Pin project, was the instigation, funding, or practical support and brokerage for cross-scale civil society networks of various levels of formality. They included networks on legal aid, education, land, extractive industries, sex workers, and gender, and 875 CSOs were reported to have been part of supported networks over six years (DFID 2016b: 10). The networks connected national level organisations more focused on advocacy and policy with more locally focused and diverse community-based organisations. Several of the networks established sub-national level platforms to engage with authorities at those levels, for example on the extractives industries. Pyoe Pin deliberately incubated and supported new national networks with sub-national links, and provided hands-on support in the practices of network building, with one review by the implementing organisation arguing this learning was a significant outcome in itself (Ward 2016: 14). Whilst acknowledging the complexity of proving attribution

in a complex and fast-changing context, the DFID programme completion review considers the 'reach and clout' of civil society coalitions developed during the programme a key achievement (DFID 2016b: 2).

The formation of vertically linked coalitions was also part of the strategy of several other programmes. CEP in Mozambique aimed to develop vertical links between grassroots organisations funded to facilitate social accountability processes and two national advocacy NGOs. This was anticipated to support the aggregation and presentation of issues raised through community monitoring and co-governance mechanisms, but also build and thicken vertical relationships between the policy-oriented NGOs and the diversity of smaller organisations engaged at district levels (Brook *et al.* 2017: iii). The Diálogo programme, also in Mozambique, undertook activity to gather its municipal level stakeholders together from across the country partly in the hope of generating policy action from national level authorities and building vertical connections between municipal and national officials, including through the national association of municipalities and a 'national urban forum' (DFID 2018e: 10). The final evaluation found positive contributions to the national association, but limited policy traction of these efforts (Shankland *et al.* 2017: 66). Whilst the AAWAZ programme concentrated largely on the surfacing and resolution of issues at more local levels, the delivery of the programme through established civil society organisations was also reported to support these organisations' wider advocacy efforts at provincial and national levels through providing them greater evidence of ordinary citizens' needs and experiences (DFID 2018c: 23) and through multi-sector state-level provincial forums (DFID 2018c: 18).

### **6.3 Developing horizontal linkages and networks**

In addition to these other approaches, most of the programmes clearly set out to build horizontal linkages – connecting people, institutions, and actions across different geographies or sectors. Documentary evidence suggests this strategy aimed to address existing disconnections between isolated sites of action (DFID 2012b: 2), develop social capital (DFID 2011b: 12), and develop communities of practice across different sites at the same administrative level (DAI 2015: 9; DFID 2018d: 8). The lack of connection is referred to in several ways as a result of prior political strategy – with the BCSSP business case referring to the need to overcome 'divisions of survival' that responded to authoritarian rule (DFID 2011b: 4). Horizontal links like this form part of the proposition of Fox (2016), who argues that they are likely to strengthen countervailing power and the ability of reformists to navigate existing power systems.

Diálogo activities were specifically designed to improve citizen-state relationships and interactions in Mozambican municipalities. Working in five major urban centres, the programme worked across horizontal scale in a number of ways. One was the linking up of government, independent media, and civil society actors across these municipalities, looking to spread and create mutual support across geographical areas and develop commonality of experience and transfer of skills and experience in participatory governance. Another was a significant focus on building networks of civic organisations and groups within these municipalities to promote coalitions for action on locally identified priority issues. People involved in the programme often described it as having a 'bridging' effect, including linking civil society actors, officials, elected local politicians, and the media (Lopez Franco 2018). The programme also sought to have a 'spill-over' effect to other municipalities by engaging those trialling participatory governance with other local governments, and through the national association of municipal government. Ongoing programme evaluation analysed these through exploring coalitions and communities of practices, and their effects, finding many positive outcomes of the strategy (IDS 2016; Shankland *et al.* 2017).

The BCSSP explicitly included building horizontal linkages as part of its strategy of connecting up civil society organisations and reformers (DFID 2011b: 10). It was particularly concerned that fragmentation of civic associations at a local level limited collective voice and hoped that developing better links between groups would lead to a more ‘empowered and cohesive civil society’ (DFID 2011b: 1). The mid-term review in 2013 found that the programme had made ‘truly impressive progress’, including in forging such connections (DFID 2013a: 5), although the later Programme Completion Report notes the fast-changing context of expanding civic space that the activities operated in at the time (DFID 2016b: 2). The Pyoe Pin component of the programme worked at different levels through explicit coalition-building around issues that emerged from political economy analysis as salient and presenting opportunities for reform. The Paung Ku component worked with small civic groups, including farmers’ and neighbourhood associations, to establish connections between them in similar geographies and between geographical areas. The programme emphasis was on doing this amongst ‘groups that would not normally associate’ (DFID 2013a: 12). The mid-term review emphasises that the logic of this was developing ‘bridging’ social capital within and between civil society groupings<sup>7</sup> (DFID 2013a: 15).

Horizontal strategies were also part of other programmes’ approaches. EVA-BHN deliberately created horizontal networks of its local level community groups, providing peer learning and mutual support between volunteers in different areas (DFID 2018a: 9), and supported exchange visits between those involved in district level advocacy forums (Taylor and Khan 2017). AAWAZ also connected its local centres and groups in horizontal networks across districts (DFID 2017a: 4). The focus on particular sub-national areas meant that these networks were fairly dense, and activities such as town hall meetings became commonplace in those districts. The CEP business case set out ambitions that CSOs involved in the programme would form coalitions together, including across different forms of civil society group (DFID 2012a: 8,10). Networks of civic organisations in different geographies were intentionally developed as ‘communities of practice’ using face-to-face meetings, email lists and social media groups. These were attributed a role in the programme’s deepening and broadening of civil society knowledge both of social accountability tools and strategies, and of health and education sector policy debates (Brook *et al.* 2017: 37).

We can see then that multi-scalar strategies were common across the programmes, with most using more than one to some degree. In the next section we move from exploring similarities of analysis and strategy to considering how these played out in practice, looking at what kinds of effects can be associated with these strategies.

## 7 What were the effects of multi-scalar strategies?

Three distinct multi-scalar strategies could be identified in the programmes we reviewed. How did they impact on their outcomes? We propose three types of effects:

- **Navigation of the governance landscape** – where multi-scalar strategies enabled responsiveness at different levels, allowed different layers of accountability to be activated, and helped people to focus on accountability ‘targets’.
- **Creating solidarity and legitimacy** – where connections between social action both vertically and horizontally increased their power to demand accountability either by building internal strengths and support, or external legitimacy.

---

<sup>7</sup> Referring to productive connections forged between individuals and groups and others who are different or distant from them in some way.

- **Leveraging evidence and multi-level pressure for policy reform** – where networks and coalitions between actors at different levels supported advocacy on policy reforms, including the institutionalisation of greater citizen engagement in governance.

We discuss each of these in turn in the remainder of this section, with illustrations from within the case study programmes.

## 7.1 Navigation of the governance landscape

Connecting action and accountability claims across levels, geographies, and sectors were reported to have generated important gains in several programmes. Strategies and practices that attempted to ‘cover the bases’ of decision-making and use this to find the right accountability targets were seen to generate clear successes in some cases. However, several reports also highlight the challenges of doing this consistently given the nature of power and decision-making in the programme contexts.

### 7.1.1 Finding and motivating local solutions

Programmes with a vertical scaling approach, and particularly those with a social accountability logic, confirmed the existing evidence from a range of contexts that citizen-based monitoring and support to formulate demands can generate changes at a ‘point of access’ level. Box 7.1 below illustrates this within the case study programmes. Many issues were reported to be resolved at the level of the school, health centre, or community where they were raised. The evaluation of CEP notes that improvements reported at school and health centre level tended to be those that required little new spending and thus could be fairly easily resolved (Brook *et al.* 2017: vi). A similar reflection can be made on the EVA-BHN programme, in which health centre staff attitude and behaviour were amongst the most common concerns raised and the great majority of the thousands of ‘issues’ raised locally were not seen as requiring specific higher level discussion or resolution (Taylor and Khan 2017). In both cases the kinds of issues resolved were often reported to be about holding officials accountable for their service delivery responsibilities or improving local facilities and infrastructure.

Whilst many of the effects involved changes on the part of service providers or officials, there are also clear examples where solutions were brought about through community contributions of some kind. An EVA-BHN programme impact case study describes the role of local fundraising for ultrasound equipment for the health centre (EVA-BHN n.d.). The CEP evaluation notes the importance of community donations of time and materials to improve facilities (Brook *et al.* 2017: 6). This outcome of social accountability programming has been noted elsewhere, including in relation to DFID supported activities in Malawi (ICAI 2013b: 11). Community contributions might represent efficient ways of solving certain problems, and embody positive ‘co-production’ (Joshi and Moore 2004) of public services against a background of public resource constraints. Youngs (2019: 12) also argues that ‘self-help’ is in itself a form of civic activism. However, these co-production or self-help outcomes may also function to excuse, rather than exact, accountability from those in higher levels of authority. Dias and Tomé (2018) argue that health service community scorecard activity in Mozambique ‘inverted’ responsibilities by encouraging community contribution and resolution – minimising accountability of those with duties to provide those services.

The key question here is whether the deliberate multi-level design of these sets of activity made local resolution and responsiveness more likely or effective. Plausibly, the existence of vertical links to higher level bodies that issues could be escalated to, might provide a degree of pressure or threat of sanction that motivates lower level solutions. Horizontal links might also plausibly build some momentum or generate a shift in norms towards responsiveness. The evidence available from the case study programmes isn’t able to confirm either of these possibilities, however, partly because of limitations in data available from their outcome



measurement schemes. There is some suggestion in the evaluation of CEP that community participants felt that without external facilitation from an independent organisation they wouldn't have been listened to (Brook *et al.* 2017: 14), but it is not clear how far this relates to the multi-scalar reach of those organisations. It is also very difficult to judge whether they were more or less successful than approaches that didn't use multi-level design.

### **Box 7.1 Local demands raised and their resolution**

- **EVA-BHN:** By September 2017, three years into the programme, more than 5,000 'issues' had been recorded as raised at a facility or community level, of which more than 2,000 had been resolved (Taylor and Khan 2017: 19). By the end of that year, 1,770 of the issues were reported to have been raised for discussion at the district level (DFID 2018b: 2), which reviews take to suggest a high rate of local resolution.
- **CEP:** More than 10,000 people were reported to have participated in community monitoring and prioritisation processes within CEP. At the end of the programme, 56 per cent of schools and 65 per cent of health facilities engaged were judged to have 'introduced changes as a result of citizen monitoring' (DFID 2018e: 7).
- **AAWAZ:** By the end of the AAWAZ programme nearly 30,000 'demands' were reported by the project as being raised with public officials or political representatives individually, in town hall meetings, or other coordination mechanisms (DFID 2018d). The DFID programme completion report claims that more than 27,000 improvements in services were made in response (DFID 2018d: 17).

#### **7.1.2 Escalation to activate accountability at higher levels**

Several of the case study programmes provide evidence of vertical integration allowing for issues that can't be resolved at a local level to be escalated to those with greater authority. Within the social accountability oriented programmes, the first level of aggregation in their vertical systems seem to have offered the most space for activating higher level responses, allowing them in some cases to overcome local resistance to change or reach officials with the power or resource to support solutions. Independent assessments of both EVA-BHN and CEP give particular examples in relation to service staffing. In these cases appeals channelled through district-level forums gained responses of allocating new staff to under-resourced services (Brook *et al.* 2017: 42; Taylor and Khan 2017: 9). In both programmes there are also reports of sanctioning or removal of poorly performing staff (Brook *et al.* 2017: 15; EVA-BHN n.d.). The independent assessment of EVA-BHN notes that relationships built between citizen representatives and health service officials at the district level developed acceptance of citizen monitoring data and feedback processes, and supported prioritisation of service improvements (Taylor and Khan 2017: 14). The Myanmar Alliance for Transparency and Accountability (MATA), supported by the BCSSP, has used its national presence to draw attention to specific local issues that otherwise, they argue, would have gone under the radar (MATA, interview, Yangon, June, 2018). A separate study claims that jade mining conditions and impacts in one particular area in the north of Myanmar were taken up nationally as a result of this multi-scalar approach (Christian Aid 2017: 23).

Two observations seem interesting from these experiences in terms of understanding what might determine how effective multi-level approaches are for escalation. First, there are questions of what kinds of spaces or structures to use. The macro-evaluation based on DFID programmes found that using existing state structures and 'invited spaces' was a necessary condition for sustained change to emerge from donor-led social accountability activities (e-Pact 2016). The CEP approach to escalation fits this model. Existing multi-sectoral forums were used by the implementing CSOs to present advocacy stories from local scorecard and co-governance processes. Additionally, local health and education staff communicated requests drawn from these processes upward through their management hierarchy. However EVA-BHN's reported success was achieved through new, externally convened groups, which Esser (2018) characterises as 'additive' to existing management nature structures. The EVA-BHN District and Provincial Advocacy Forums brought health officials, politicians, journalists,

and local influencers together with locally nominated representatives of community-based monitoring groups. Programme documentation argues that these new groups were necessary as suitable existing structures for citizen interface were unavailable – where suitability also included a view of how far they were effected by elite capture and represented genuine opportunities for engagement, including by women (Palladium 2016b: 1). Evidence suggests that the construction of these spaces enabled the programme to set the agenda based on citizen issues, bring in the weight of relatively powerful actors (such as journalists) to support citizens' demands, and build trust with officials outside of formal decision-making spaces (Taylor and Khan 2017: 14). The approach of BCSSP, through the Pyoe Pin project, was to establish new 'issue coalitions' with carefully selected members who often had existing influence (British Council 2016a). It may be that where spaces for citizen participation are particularly rare or closed, and informal power is important, 'additive' spaces are necessary.

Second, the programme experiences confirm what might seem a fairly obvious point; the success of escalating to the 'supervisory' level of authority is dependent on that level in the vertical hierarchy having the ability or resources (as well as the willingness) to respond. The evaluation of CEP notes that the lack of vertical responsiveness to issues raised in the education system was connected to the limited discretion of district-level officials and extreme resource pressures (Brook *et al.* 2017). Examples are given of officials making requests of their superiors to fix issues prioritised by communities, but these not being prioritised, with the evaluation suggesting that this kind of procedural aggregation of priorities to district level is insufficient, and needs the active advocacy of those involved to get redress or reform (Brook *et al.* 2017: 43). The independent assessment of the EVA-BHN programme noted that responsiveness to community-generated demands increased once the Punjab level government delegated more resources to the district level (Taylor and Khan 2017: v), suggesting that this meant that prioritised issues could actually be addressed. Esser (2018) recommends to the programmes in Pakistan that the district level should be the focus of any future similar programmes, given the concentration of resources and decision-making power at this level. Such patterns of delegation are liable to change, however, perhaps more so in fragile and conflict-affected settings, and subsequent changes in local governance in Pakistan may well have changed the grounds on which this recommendation was made. At a broader level, these experiences suggest that analysis of what levels need to be included in a multi-level approach might be key to success; with a focus on where decisions might actually be taken.

### **7.1.3 Locating alternative 'accountability targets'**

Working across scales might also present more options where bottlenecks and resistance to change present a problem. A fundamental question behind the vertical integration of accountability mechanisms is whether such integration and externally induced strategies can get to where the relevant decision-making power lies for different issues. Given that it might be located in various parts of the system in different moments and specific local contexts, and that we might expect this to be less legible and static with background conditions of conflict and fragility, a truly vertically integrated system to channel citizen demand needs to run effectively across this system. This includes outside of the immediate sector or service delivery hierarchy, as well as outside of formal sources of authority and decision-making. Evidence from the programmes gives instances where this kind of navigation of the accountability landscape has been possible, but also where the nature of these systems of authority have limited success.

There are several examples where the programme activities 'skipped' levels or went to alternative authorities to circumvent blockages and address more systemic issues. A case study of impact written by the EVA-BHN programme illustrates one way to do this by making claims of different authorities. In this example, the poor condition of health centres was

addressed not by activating health service decision-making, but by including a number of renovations and improvements in the separate local government budget (EVA-BHN n.d.). The external mid-term evaluation of the Diálogo programme highlighted that the focus of activities were on services or decisions that lay solely with the municipal authorities, despite a number of important services being delivered in cities by national government departments and state-owned companies, or there being mixed accountabilities for these (IDS 2016). The programme subsequently introduced activities to address this through upward advocacy to national levels on decentralisation policy and the specific issues affecting municipalities (Shankland *et al.* 2017: 11). This included working through the national association of municipalities, ANAMM, in order to find allies from other municipalities. However one respondent central to the programme management emphasised the difficulty of this given a 'huge gap' between the municipal and national level and the lack of a driving central issue to campaign on (Lopez Franco 2018). Another described this national advocacy as a significant 'leap' from the original programme design (Lopez Franco 2018). CEP also placed a focus on linking facility and community level advocacy and civic engagement with national policy given a lack of discretion and response at the district and provincial levels, with the evaluation suggesting that the programme had surfaced a 'missing middle between local accountability and national policy' in Mozambique (Brook *et al.* 2017: iii). The BCSSP activities went beyond the national level in several cases and worked with international actors and partnerships as part of their influencing strategies. Examples include developing a regional South East Asia Legal Aid Network to provide wider support for the Myanmar Legal Aid Network (British Council 2016b: 33), and reporting issues directly to the international secretariat of the Extractives Industry Transparency Initiative when the national government was unresponsive (MATA, interview, Yangon, June, 2018; Vijge 2018). These illustrate what Keck and Sikkink (1998) have called the 'boomerang strategy' to circumventing obstructions at the national level.

The programmes also give us some illustration and understanding of why these navigations are necessary in these contexts. First, the formal decision-making systems tackled by the programmes were, on the whole, relatively centralised with limited discretion at lower levels, as identified in programme contextual analyses. In Mozambique political dynamics meant that national level resistance was reportedly linked to the fact that some municipal areas involved in the Diálogo programme were run by the opposition party. Second, decision-making is not only formal, but informed by personal and political relationships and informal powerholders. EVA-BHN tried to address informal power structures by tactically choosing to engage influential people in key roles in the programme and its fora, alongside more marginalised people, 'protecting and amplifying voices' (Palladium 2016b: 11). However according to practitioners, it also experienced cases where decisions made in those fora, even by those with designated formal power, were later overturned as a result of informal power structures. In one example, a member of staff was assigned for transfer after repeated complaints from the community-based monitoring group, but managed to use political contacts to have three separate transfer orders overturned (Taylor and Khan 2017: 12). CEP was able to leverage the personal networks and credibility of the directors of its sector lead CSOs (N'weti and CESC) as well as the programme director herself (as a former Ministry of Health official). However according to some interviewees tensions between the central management team and the sector lead NGOs limited the overall effectiveness of this leveraging strategy. The CEP programme completion review notes that despite successes the programme made clear the extent to which 'national policy frameworks are implemented according to local politics' (DFID 2018d: 8), alluding to similar informal dynamics.

## **7.2 Creating solidarity and legitimacy**

The second set of effects of multi-scalar strategies we found in reviewing DFID programmes were in building solidarity between social actors seeking changes to services, policy reforms, and greater civic engagement in governance, and the perceived legitimacy of their claims

and efforts. Both of these are related to coherence and strength of countervailing forces, but they come from different directions. Solidarity, as a form of 'power with' (VeneKlasen *et al.* 2002: 55) might result in greater collectivity in action, and confidence to take action. Legitimacy is more outward facing and a factor in how action is perceived by others – in particular formal authorities but also international actors. The two factors are connected in the extent to which increased collective power also drives a sense of legitimacy amongst actors. They are also connected in that they are both functions or creations of particular kinds of coalition-building, understood as a process that goes beyond less densely formed networks or alliances of interests to bring together differently positioned actors in common cause (Fox 2010).

Both of these support a common logic in programme design that tries to establish 'strength in numbers' through multi-scalar activity, and to develop the 'critical mass' required for reform that is discussed in several programme design statements (Brook *et al.* 2017: 65; DAI 2015: 15; Shankland *et al.* 2017: 11). Review of other DFID empowerment and accountability programmes notes that building this critical mass was also a common rationale for activities funded in other contexts (ICAI 2013b: 10, 2018: 26). Similar effects have been highlighted as important in international NGO-led accountability programmes (Amakom *et al.* 2018: 19; Fooks 2013: 6). As well as potentially having direct empowerment outcomes for those that benefit, these effects might also do some groundwork for future accountability claims and increase the chances that these claims have sufficient 'clout', the term used to note the effects of civil society coalition building in the final report on the BCSSP (DFID 2016b: 2).

### **7.2.1 Solidarity**

Effects that we describe as increasing solidarity are about building the internal strength of reformist coalitions, campaigns and civic action. Programmes variously reported building collective identity amongst previously dispersed groups and sets of activity, giving them greater shared purpose, leveraging cross group support, and allowing them to explore the common causes of their relative weakness in relation to more powerful actors. Programmes adopted this language to greater and lesser degrees. Alif Ailaan, for example, was explicitly presented as building a 'social movement' around education reform (DFID 2018f: 1) that recruited 'mobilisers' at community level. Several programmes explicitly use language of coalition-building.

Interviews with key actors in the Diálogo and CEP programmes in Mozambique identified that structured exchange of experiences, and the opportunity to generate informal networks of reformists were key in the programmes' successes. The Diálogo programme's strategy emphasised horizontal coalition-building which was based on 'connecting and promoting synergies between actors' within and across municipalities and purposefully 'amplifying the networks' (Lopez Franco 2018). However a number of interviewees on the programme noted challenges in building these coalitions, particularly as NGOs and CSOs were reluctant to engage openly with one another in a broader environment of competition for donor funding, and other donors such as USAID encouraging the same organisations to focus on service delivery rather than advocacy (Lopez Franco 2018). The independent assessment of the EVA-BHN programme found efforts to link up community group coordinators horizontally to build 'a sense of collective identity' between these volunteers which included 'an understanding of their common problems' (Taylor and Khan 2017: 13). It also highlighted the role of women-only community groups established under the programme, noting 'a strong sense of solidarity' emerging from these spaces (Taylor and Khan 2017: 30).

### **Box 7.2 Solidarity through multi-scalar networks in Myanmar**

Amongst several multi-level civil society networks supported by the BCSSP programme in Myanmar was the Myanmar Alliance for Transparency and Accountability (MATA). MATA was initially formed to drive civil society voice when the Government of Myanmar announced an intention to join the Extractive Industries Transparency Initiative (EITI) in 2011. Whilst the government initially wanted to choose 'friendly' NGOs to join the steering group, donors supported local activists to create a viable alternative in MATA (Christian Aid 2017:19).

MATA members and staff gave several examples in interviews where internal connections within the alliance gave rise to solidarity actions. They reported how members across the country 'mobilised' when one sub-national area pursued what they called a 'clamp-down on CSOs' through new regulations – threatening national protests that saw the regulations cancelled (MATA, Interview, Yangon, June 2018). In another example they claimed that the announcement that they planned a national protest march drove the release from prison of member local activists who had been part of demonstrations aimed at the Kayah State (provincial) government (MATA, Interview, Yangon, June 2018). In 2014 MATA threatened to boycott the EITI process in response to the police shooting of an activist protesting the environmental impact of the Letpadaung copper mine, drawing international attention to both the impact of the extractive industries in Myanmar but also the state response to challenge (Vijge 2018).

One of the effects of increased solidarity that relates specifically to multi-scalar strategies is the ability of connected actors and groups to act in protection or support of others. This was evident particularly in the work of the BCSSP in Myanmar. Examples from the work of a programme-supported national civil society alliance are given in Box 7.2. These effects are also reported by external reviews of MATA's work. Vijge (2018) suggests that the alliance formed gave individual organisations 'a sense of security... against the Myanmar government and army' (Vijge 2018: 23). A separate study of the nature of CSO networks in Myanmar that covered many of those supported by BCSSP reported that network participation was spoken of by members as a way to sustain their motivations, arguing that these networks make solidarity 'visible' (Christian Aid 2017: 18, 39).

#### **7.2.2 Legitimacy**

In several programmes multi-scalar links also seem connected with the sense of external legitimacy of those taking action (whether citizen-based monitoring groups, or CSOs making representative claims), the nature of the social action, and the accountability claims stimulated and supported. The focus appears to have been on the perceived legitimacy in the eyes of decision makers or power-holders – most notably governments and officials – but this also connects with the work of several programmes to promote the role of new social actors and mobilised citizens in independent media coverage.

The Alif Ailaan campaign seemingly derived legitimacy from its attempts to stimulate a broad-based and country-wide coalition for education service reform and prioritisation, validating local action under the umbrella of a larger campaign. Esser (2018: 17) notes this reach as particularly valuable in the Alif Ailaan programme in contrast to the two other programmes he explored in Pakistan. Programme documentation refers to the importance of building 'demand at the constituency level' for reform (DFID 2017b: 15) and this work entailed localisation of the campaign in 50 districts involving more than 7,000 volunteers (DFID 2018f: 14).

According to our interviewees, Dialogo's local legitimacy came from brokering links between CSOs, media and local government, and also between municipalities run by the opposition and by the national ruling party – exercising convening power across scales. CEP used national networks to enhance the legitimacy of service monitoring, aligning this with national policies on strengthening local co-management initiatives and conveying these to local facilitators. In interviews for the evaluation, several facilitators or community members noted that they did not feel they would have been taken seriously by teachers or health professionals without external support, with the evaluation concluding that the programme 'lent weight and authority' to the community engagement processes (Brook *et al.* 2017: 14). In a case identified in an impact tracing study by the EVA-BHN programme, the denial of a

vaccination for a child by a local health worker was successfully challenged when activists within the programme used this kind of external legitimacy to stage a campaign, including a protest outside the district health office (EVA-BHN n.d.).

Within BCSSP the coverage and depth of networks with national representation was given as a significant source of legitimacy by interviewees. The Land Core Group (LCG) is a research and advocacy network of land activists and CSOs in different areas across Myanmar. Representatives from LCG suggested that the recognised strength of their network was behind their ability to call a significant public consultation on new land laws announced by the government, which meant that the laws were subsequently amended (Land Core Group, interview, Yangon, 2018). A leader of MATA spoke of the vertical and horizontal connections within the alliance as a source of 'legitimate power' in relation to authorities (MATA, interview, Yangon, June, 2018). Others involved in the coalition noted how this collective power meant that whilst individual members often failed to get information from government, requests through MATA were more successful as the coalition was seen as more able to 'make a problem' if they didn't get the information (MATA, interview, Yangon, June, 2018). This effect was not only seen when requests were escalated up to the national level, but also for those flowed downward to support local level action. Members of MATA reported various examples where their actions to monitor extractive industries were validated or permitted through their connection to the national network, including a case where citizen monitoring had led to badly managed mines being closed (MATA, interview, Yangon, June 2018). During our stakeholder workshop with programme practitioners from different contexts this was referred to in terms of donor programmes providing a 'protective umbrella'; using the relative political strength of donors and some national NGOs to decrease the risks to local challenges of power structures and decision makers.

There were however challenges to the perceived legitimacy of some of the programmes and the activities they support, connected to issues of scale. In EVA-BHN and CEP the evidence generated by citizen monitoring was not seen by some as sufficiently weighty to drive reforms. The independent assessment of EVA-BHN found that responsiveness from provincial level authorities was affected by the relatively small geographical scope of the programme and thus a view of the evidence generated as of 'limited scale' (Taylor and Khan 2017: v). Esser reports a view from one development sector stakeholder that a focus on a small number of areas made it 'more of an irritant than a game changer' (2018: 22). In CEP, national advocacy organisations could use grassroots data in combination with other evidence from the provinces and districts covered, but one member of staff told us 'having a weight of argument based on these small samples was unrealistic' (Lopez Franco 2018). It is notable that other case evidence on vertically integrated citizen monitoring and advocacy involves much more significant geographical coverage (Aceron 2018; Aceron and Isaac 2016; Fox 2016).

### **7.3 Leveraging evidence and multi-level pressure for policy reform**

All of the programmes we explored aimed to have wider levels of impact on service delivery and citizen-state relations through stimulating high-level policy change, using their multi-level approaches as part of this. Two kinds of impacts were hoped for. One was that evidence conveying the realities of citizen experiences and preferences informed new policy or practice. The other was that citizen engagement in service monitoring, co-governance, or policy processes became more widely recognised and institutionalised. Both were envisaged to have impacts wider than the specific geographies or sectors the programmes worked in and longer-run dividends in terms of state-society relations. The programmes were in various ways trying to connect the different 'institutional geographies' (Fox 2016: 19) of service delivery, policies, and governing frameworks, whereby those closest to the problem become the 'eyes and ears' of national advocates (Fox 2016: 18). The ability to attribute policy

impacts of donor-led interventions like the ones we've looked at is limited even in the detailed evaluations available, but a number of examples are given by the programmes.

### **7.3.1 Reforms to the content of public policy**

All programmes we looked at targeted and reported impacts on specific public policies at the national or sub-national level. These were linked to multi-scalar activities either by drawing up evidence to higher levels of decision-making or establishing pressures at multiple points in the system. As discussed in the next section, achieving policy impacts like these was commonly part of official programme measurement frameworks.

The two programmes with the strongest social accountability logic, CEP and EVA-BHN, both give illustrations of taking experiences from geographically-targeted citizen-based monitoring and prioritisation, and engaging in policy that affects a wider population. The CEP evaluation gives examples of how the national level CSOs involved in the programme used the grassroots experience data generated by the programme as part of their national campaigning. It notes that citizen-monitoring data was used as supportive evidence in national level reports by the education advocacy organisation (CESC) and informed their campaigning on teacher and student absenteeism (Brook *et al.* 2017: 29). The cross-scalar links gave what a CESC member of staff called a 'clear unifying goal' (Lopez Franco 2018). In the health sector the evaluation reports that the national advocacy partner organisation, N'weti, 'was able to play a strategic role linking local realities with national policy' using CEP generated data – for example emphasising that national level indicators on medicine availability were incongruent with 'demand-side' evidence from citizens that medicines were still in short supply (Brook *et al.* 2017: 50). Within the timeframe of the project, however, neither the evaluation nor the programme completion review were able to point to specific policy wins based on or responding to the aggregation of citizen voice through the programme. In addition to its reportedly greater success at activating reform and re-prioritisation at the district level, EVA-BHN claims some more systemic success at the provincial level in Pakistan, where the majority of responsibility for health service financing and delivery rests. One example was the establishment of a Patient's Rights Charter in one of the provinces the programme worked in, which its reports argue reflected the kinds of service provision issues that emerged as particularly important from aggregated citizen experiences and demands (EVA-BHN n.d.). Produced in partnership with the Provincial Health Care Commission, a formal oversight body, this may lay the ground for future accountability claims, with standards of fair treatment and access to services now formally recognised by authorities. The independent assessment of EVA-BHN also suggests that the programme played a role in making the case for delegation of budgets to lower levels in order to resolve issues more effectively, using evidence from the programme to illustrate how greater local discretion could improve services and experiences (Taylor and Khan 2017: 24).

Other programmes used connections across scales in ways that might be seen to pressure different parts of the system to prioritise reform. Alif Ailaan sought to mobilise local campaigners behind a set of core asks and advocacy priorities at a national level, which they were tasked with localising at district and provincial levels within the overall umbrella of the campaign (DFID 2018f: 14). The programme created new presentations of data and conducted multi-level research to support these local asks (DFID 2018f: 8). It also used these connections and advocacy at different levels to reinforce a sense of widespread prioritisation. The Programme Completion Report for the programme argues that this multi-level strategy was a key factor in what it sees as the success of the campaign – pressuring politicians by creating a 'constant buzz around them at all levels' (DFID 2018f: 15). The Programme Completion Report also considers the Alif Ailaan campaign to have made a 'significantly positive contribution' to education sector reforms at different levels, including a reduction in the number of children out of school, increased provincial budgets for maths and

science education, and infrastructure improvements in more than 800 schools (DFID 2018f: 16).

The AAWAZ programme reported successes in influencing six specific pieces of provincial and national level legislation on gender equality based in part on the credibility of expertise developed from local women's groups, which included contributions to drafting as well as lobbying and advocacy on issues like violence against women, although little detail on these processes is given (DFID 2018c: 21). Amongst the policy influencing reported by activities in the BCSSP there are also examples of the programme enabling localised concerns and struggles to be represented in national advocacy and in the policy frameworks. For example, MATA representatives reported that grassroots evidence on State Owned Enterprises' involvement in local disputes about mining was used to pressure government to include these organisations in the scope of the next EITI report (MATA, interview, Yangon, June, 2018). Whilst warning about the difficulty of attributing policy changes to a single programme in this context, the final assessment of the BCSSP reports 'compelling narratives' on how multi-level coalitions supported more inclusive and pro-poor policy (DFID 2016b: 4).

### **7.3.2 Reforms to institutionalise citizen participation**

There are also ways in which the programmes tried to use their multi-scalar strategies to create greater commitment for citizen engagement in governance issues generally and open up the policy-making process.

In CEP these kinds of outcomes were prioritised through the national advocacy partners. It worked in particular to strengthen commitment to the policy regarding 'co-governance' mechanisms of joint management committees of citizens and providers at the service level. However the programme also undertook advocacy, pushing national ministries to include more performance indicators that reflect citizen experiences on areas such as medicine availability (Brook *et al.* 2017: 50). In education, the national partner CESC reported that CEP experiences of co-governance allowed them to galvanise national support for regional and national fora of school management committees, as well as share positive examples of local problem resolution through these mechanisms to counter perceptions that CSOs were only presenting problems rather than solutions (Brook *et al.* 2017: 31). Whilst the Government of Mozambique didn't subsequently adopt new citizen engagement policies, our discussions with local stakeholders revealed some felt that the programme's successes had opened the eyes of some officials, and especially donors, to the potential for citizen-based monitoring in the country, and led to discussions about using external influence to institutionalise these approaches (Lopez Franco 2018). Subsequently CEP-inspired citizen participation and monitoring mechanisms were built into a new World Bank health sector programme through the advocacy of CEP's national health partner, N'weti. Another impact in terms of national advocacy and programme learning was the recognition of the need to aggregate and analyse facility-level evidence more effectively. This informed the development of a database to capture citizen experiences and demands that included social accountability initiatives funded by several other donors and run by a range of NGOs, potentially creating a longer-term link between national advocacy and grounded evidence. Similarly, the end evaluation of the Diálogo programme noted increased openness to participatory governance approaches in the target municipalities, as well as substantial interest in other urban areas in the participatory budgeting approach in particular, which was actively promoted in 'lesson learning' sessions for other areas (Shankland *et al.* 2017: 8).

In Pakistan, EVA-BHN worked to build citizen monitoring and ongoing spaces for dialogue into the district and provincial levels, seeing the work as establishing 'proof of concept' for the benefits of citizen engagement. Towards the end of the programme both provincial governments issued instructions for officials to establish or reinvigorate district-level dialogue platforms with citizens' representatives. The independent assessment found that it was likely



that this was driven by the ‘pressure’ created by EVA-BHN’s demonstration of how these could work, and the expectations raised amongst communities of ongoing involvement in decisions (Taylor and Khan 2017: 24). EVA-BHN was also regarded as having been instrumental in the incorporation of ‘demand-side’ indicators on citizen experiences and ‘client satisfaction’ into the measures used by the independent oversight commission on the health sector in KP province (Taylor and Khan 2017: 24). The AAWAZ programme completion review argues that the programme ‘broke the monopoly of executive oversight’ (DFID 2018c: 21) and had an important and likely lasting impact on social norms of women and young people’s participation. However there is little detail on any institutionalisation of these effects, although evidence is presented that suggests that attitudes to women’s political and civic participation had become more positive during the programme (DFID 2018c: 22). Towards the end of the Alif Ailaan programme new outcome measures were added to identify how far it was bringing about changes that might outlast the campaign, for example how far ‘popular demand’ for education sector reform was mobilised (DFID 2017b: 15). The programme also formed issue and geographical alliances of CSOs to adopt and take forward the campaign messages after its end (DFID 2018f: 14).

An interesting reflection to emerge from our engagement with practitioners in these programmes was how these longer-run outcomes of institutionalising citizen engagement in governance practices, and changing ‘the rules of the game’ were prioritised over the more immediate policy changes emphasised in the expected programme results. In the next section we explore this and other tensions and issues that emerge from the ways that programmes officially measured their impact.

## 8 How did the programmes measure their multi-scalar activity?

The previous sections have explored how multi-scalar accountability strategies have been used in DFID-funded programmes, and what kinds of effects we can see associated with these. In our study we also set out to understand how far DFID-funded programmes contribute to the evidence base for adopting strategies of this kind. We explored the ways in which programme activity was monitored and evaluated and evidence and learning on changes gathered, analysed, and shared. This included detailed review of official measurement frameworks as well as internal and external evaluations as far as they were available, and discussions with programme stakeholders both during fieldwork and in a wider practitioner workshop in the UK about these issues.

In this section we first summarise the kinds of monitoring and outcome measurement approaches taken in our sample of programmes and discuss the choices made. We note differences in conceptualising measures of empowerment and accountability, a focus on quantitative measures, and a lack of measurement specifically on multi-scalar effects. We then look at the broader evidence available outside of these frameworks, including provider-produced evidence, independent evaluations, and evaluative content and reflections in other documentation. Finally, we draw out some implications of the limitations of the evidence in terms of a wider ambition to learn from these programmes and what our study suggests drives these limitations. We explore differences in the quality and robustness, and indications of a potentially rich but inaccessible set of relevant data held by programme implementers.

### 8.1 Programme monitoring regimes

The programmes we looked at all used a ‘logical framework’ as a way to record the extent and results of activities. Current DFID guidance does not require programmes to use a

logical framework approach but notes that this is commonly how programmes meet the requirement to record expected and actual results (DFID 2019c: 69). The logical framework monitoring tools we saw followed the standard structure used within DFID where statements of impact, outcome, and outputs are linked hierarchically along with indicators or measures for each. The ‘logical’ element of this approach is an assumption that these are causally linked; that lower level activities measured contribute to higher level and more significant overall changes. All programmes had far more extensive monitoring regimes and evidence collation practices than were expressed in these logical frameworks. However, the logical frameworks are important because the standard structure of annual and completion reviews of programmes tends to tightly follow indicators and targets from these. The success or failure of a programme is largely judged in relation to the chosen indicators. This produces incentives in terms of where programmes concentrate their resources both in implementation and in generating evidence and analysis. These incentives and effects are widely discussed in the literature on measuring change in donor programmes (Eyben *et al.* 2015; Holland *et al.* 2009; Shutt 2016) and in terms of DFID management practice specifically (Valters and Whitty 2017). Their implications for programme learning have also been explored by the UK Independent Commission on Aid Impact<sup>8</sup> (ICAI 2014, 2019).

Each programme set out some initial expectations of results and measurement in their business cases, which were then developed further as programmes progressed. We found evidence of regular changes to logical frameworks, expected results, and measurement approaches both in iterations of the logical frameworks through different versions we saw, and in the reports of programme implementers. In some cases these changes were explained in annual review documentation. Where logical frameworks were publicly available on the Development Tracker website these typically reported results against the included measures for the previous years, although it was often difficult to find out how measures were defined and measurement undertaken.

Table 8.1 below summarises the kinds of measurement used in these logical frameworks at output and outcome level. We then explore the frameworks in relation to their choices in three areas – indicators of empowerment and accountability, the use of qualitative and quantitative approaches, and explicit measurement of multi-level and multi-scalar interactions.

**Table 8.1 Summary of output and outcome monitoring measures of studied programmes**

Programme	Outputs	Outcomes
AAWAZ  PCR	<ul style="list-style-type: none"> <li>• Numbers of participants in training, activities, forums, and action at several levels and by sub-national location.</li> <li>• Changes in knowledge, behaviour and attitudes amongst participants</li> <li>• Volume of citizen ‘demands’ raised in different fora</li> <li>• Number of policy and implementation changes made with some attribution to programme activities</li> <li>• Number of research and evidencing projects and associated advocacy campaigns completed</li> </ul>	<ul style="list-style-type: none"> <li>• Impacts on legislation supporting rights of women and excluded groups</li> <li>• Knowledge, behaviour and attitudes survey on women’s political participation and violence against women</li> <li>• Reported use and satisfaction of excluded groups with public services</li> <li>• Direct and indirect beneficiaries of AAWAZ forums and centres</li> </ul>

(Cont’d.)

<sup>8</sup> ICAI was established in 2011 to provide independent scrutiny of UK Overseas Development Assistance, reporting to parliament through the International Development Select Committee. As part of its mandate to explore the effectiveness of UK development programming it conducts learning reviews on particular areas of practice.

**Table 8.1 (cont'd.)**

Programme	Outputs	Outcomes
Alif Ailaan <i>PCR</i>	<ul style="list-style-type: none"> <li>• Reflection of programme advocacy goals in political party, politician, and opinion leader knowledge and actions</li> <li>• Policy influencing activities; research, individual briefings,</li> <li>• Number of government policy actions in line with campaign priorities</li> <li>• Increases in quantity and quality of media coverage of education issues</li> <li>• Extent of campaign membership from CSOs and activists and alliance-building</li> <li>• Volume of localised campaigning by activists</li> </ul>	<ul style="list-style-type: none"> <li>• Budget allocations to education</li> <li>• Improved learning outcomes</li> <li>• Decreases in out-of-school children</li> <li>• Adoption of campaign goals by CSOs and campaigners</li> <li>• Increased demand for and understanding of quality education amongst parents and politicians</li> <li>• Increased political will to prioritise education reform</li> </ul>
BCSSP <i>PCR, MTR</i>	<ul style="list-style-type: none"> <li>• Qualitative scoring of increases in collaboration, shared analysis, and policy influencing skills amongst CSOs*</li> <li>• Issues identified for advocacy, and reform coalitions built around them.</li> <li>• Reach and inclusivity of coalitions and advocacy campaigns</li> <li>• Qualitative scoring of increases in horizontal and vertical linkages between groups*</li> <li>• Qualitative scoring of CSO organisational capacities</li> <li>• Media initiatives and reporting of issues relevant to reform coalitions</li> </ul>	<ul style="list-style-type: none"> <li>• Quality of civil society advocacy cooperation, communication, and inclusiveness*</li> <li>• Number of coalitions using evidence-based advocacy strategies</li> <li>• Evidence of increased social capital</li> </ul>
CEP <i>PCR</i>	<ul style="list-style-type: none"> <li>• Reach and understanding of citizen awareness activities on rights and entitlements</li> <li>• Reach and volume of citizen-based monitoring/ community scorecard activities</li> <li>• Development of action plans by facilities/ services to address priorities</li> <li>• Volume of operating co-governance mechanisms (school/health centre committees)</li> <li>• Volume of advocacy activity to higher term authorities using programme evidence and on specific issues that need higher level resolution</li> <li>• Dissemination of learning and tools developed</li> </ul>	<ul style="list-style-type: none"> <li>• Number of public service facilities introducing changes as a result of citizen monitoring</li> <li>• Service users' perceptions of influence over service provision</li> </ul>
Diálogo <i>PCR</i>	<ul style="list-style-type: none"> <li>• Civic groups and community leaders supported and subsequently engaging with municipal institutions</li> <li>• Number of coalitions developed</li> <li>• Volume of media coverage of municipal governance issues and marginalised populations</li> <li>• Number of journalists trained</li> <li>• Provision of technical support to Municipal Institutions</li> <li>• Behaviour change of Municipal Institutions</li> <li>• Dissemination/advocacy based on lessons learnt</li> </ul>	<ul style="list-style-type: none"> <li>• Citizen perceptions of their inclusion in municipal governance*</li> <li>• CSO capacity to engage in municipal governance (qualitative score)</li> <li>• Increased quantity and quality of media coverage of municipal governance issues</li> <li>• Municipal Institutions capacity to engage other social actors</li> <li>• Volume of participatory events run by Municipal Institutions</li> <li>• Advocacy to influence national policy and influence other municipalities</li> </ul>
EVA-BHN <sup>9</sup> <i>Logframe</i>	<ul style="list-style-type: none"> <li>• Number of citizens reached through public health/rights and entitlement communications</li> <li>• Number of demands raised from local groups to district level, and their outcome</li> <li>• Level of relevant coverage of health and nutrition issues in print and television media</li> </ul>	<ul style="list-style-type: none"> <li>• Number of changes in government policies and/or practices where EVA-BHN contributed</li> </ul>

Source: Authors' summarisation from programme completion reviews and logical frameworks. Asterisks mark measures included in final frameworks but ultimately not measured over the course of the programme.

<sup>9</sup> The comparatively lower number of measures and indicators for EVA-BHN in the formal logical framework is explained by its place as a sub-component of a much larger programme of investment.

### **8.1.1 Measures of empowerment and accountability outcomes**

We looked at how empowerment and accountability outcomes were measured in programme logical frameworks. The growth of development programmes seeking these kinds of outcomes took place in parallel with the shift in programme measurement and evaluation towards more ‘results-based’ models (Valters and Whitty 2017), and the consequence of this has been a period of trialling and testing ways in which these outcomes can be identified and counted. There have also been a number of contributions exploring how impact of empowerment and accountability programmes can be evaluated, and providing advice on measuring change (Eyben *et al.* 2015; Holland *et al.* 2009; Rocha Menocal and Sharma 2009). Practitioners made it clear during our discussions that they don’t yet feel they have sufficient solutions to these measurement challenges.

Within our focus programme logical frameworks, accountability outcomes were largely framed in terms of national policy changes or changes in practices in particular service sectors, or as the voicing of and response to demands from mobilised citizens. Programmes commonly looked to count the number of instances or examples where policy or practice change resulted from the expression of citizen voice or civic engagement activities, whether through individual demands or the aggregation or representation of these in advocacy processes. This policy level impact was looked for at different levels of authority and with different sectoral policy targets. Three programmes specifically counted the volume of citizen ‘demands’ raised on public authorities (AAWAZ, CEP, EVA-BHN). The EVA-BHN monitoring framework was unusual amongst these in disaggregating how many of these demands were considered to be met (or ‘issues resolved’), although the CEP counted how many schools and health centres had responded to community scorecard prioritisation.

Empowerment outcomes were rarely monitored in logical frameworks. Despite many of the programmes involving citizen education and rights awareness initiatives, often reaching high numbers of people, their monitoring frameworks tended to judge the activities only on the reported number of participants rather than any changes supported. Only one programme logical framework (AAWAZ) explicitly measured empowerment outcomes in terms of increased agency and capacity to act, through regular tracking of skills and behaviours of trained community mobilisers or activists.

### **8.1.2 Use of quantitative and qualitative measures**

Writing for DFID, Holland *et al.* (2009) recommend that empowerment and accountability aims should be measures with a focus on behavioural and institutional change in order to bridge the gap between ‘inputs’ and higher level outcomes, and by evidence of the quality of processes as well as the quantity of activities or participants. Later DFID guidance on designing programme measurement emphasises that qualitative process-focused measures might be more appropriate for ‘challenging environments’ and where long-term changes are the aim (DFID 2019b: 62). In the logical framework measures used by our focus programmes, however, there was a general tendency towards measures of quantity over interrogation of quality.

There were a number of exceptions. The BCCSP was initially designed to use a set of qualitative scales of change on issues such as CSO capacity and multi-scalar linkages, using a qualitative assessment scorecard approach. Material from one component illustrates how stories of change were gathered to support this (British Council 2016a, 2016b). However implementation of the approach proved difficult and the final review of the programme queries the strength of the evidence gathered (DFID 2016b: 20). The Diálogo programme used ‘outcome mapping’ approaches to gather and judge evidence for higher level measures, and citizen perception studies at the start and mid-point. It also undertook a qualitatively based ‘ripple effect’ study on spill-over effects, although this didn’t inform the logical framework measures (Brook *et al.* 2017). The EVA-BHN programme also

commissioned more qualitative studies (EVA-BHN 2017a, 2017b, 2017c, 2017d), although again these didn't feature in the official measurement of results.

Associated with this, programme measurement frameworks also tended towards indicators that illustrated significant volumes of activity rather than the role of these activities in processes of change. Again, there were notable exceptions; the Diálogo framework theorised its output measures as steps in a process of capacity-building and reform, and there was a focus in higher level monitoring of the AAWAZ programme on changes in behaviour facilitated by the parallel processes of community-based empowerment and new spaces in which to make claims of authorities. A number of the frameworks included indicators which recorded very high numbers of beneficiaries or anticipated reach – for example potential audiences of media coverage or wider populations that may have been indirectly impacted by the programme activity. There were suggestions from our discussions with programme practitioners that these aspirational indicators were added to monitoring schemes largely to supply data for global indicators that DFID has used, for example on the aggregate number of people DFID programmes have supported to have increased 'choice and control'.

### **8.1.3 Measurement of multi-level and multi-scalar strategies and effects**

We analysed the logical frameworks for indications of measurement across vertical and horizontal scales and of multi-level interactions. The programmes incorporating a citizen-based monitoring or demand aggregation logic (AAWAZ, CEP, EVA-BHN) all included indicators related to the volume of activity at different levels. The Alif Ailaan and BCSSP programmes measured the extent to which coalitions and networks built were multi-level, both in terms of geographic spread and at different vertical levels. The Diálogo programme's logical framework had outcome measurements on the horizontal spread of participatory governance practices across municipalities, and advocacy efforts at a national level on the basis of programme evidence and experience at the municipal level.

However, we found little logical framework measurement capturing articulation or interactions between sites or levels of action – how the work at each level influenced other levels. Both EVA-BHN and CEP specifically counted instances of 'escalation' of issues to an upper authority. In the case of CEP, a target was set specifically for the proportion of issues identified as needing higher level resolution which were then taken to a higher-level decision maker. The most concerted attempt to measure the interactions of levels was in the BCSSP programme, which had elements of its monitoring framework explicitly aimed at measuring the development of vertical and horizontal linkages. The programme planned to use a qualitative scale judged by a review panel, but the programme completion report suggests that in practice this did not work well, with methodological difficulties with the qualitative assessment scorecard approach hindering results (DFID 2013a, 2016b). In the absence of this the 'output' on the development of these linkages was measured by the membership of multi-scalar networks.

## **8.2 Understanding processes of change and impact**

Although important in terms of institutional incentives and definitions of success, logical frameworks and official monitoring schemes might not be the best place to look for evidence that helps us understand complex change processes. Some argue these frameworks primarily perform internal and upward accountability functions in DFID programme management, rather than being contributions to evidence (Eyben *et al.* 2015; Valters and Whitty 2017). This was a view that also emerged during our practitioner interviews and workshop. Eyben refers to them as 'disciplining artefacts' that tend towards easily measurable outcomes (Eyben 2015) whilst the UK Independent Commission for Aid Impact has expressed concern about the 'distortive effect' they have on learning by narrowing the scope of evidence collected (ICAI 2014: 29). Taking this into account, we researched other forms of programme evidence that might be considered more likely to explore longer-run

outcomes and processes of change. We looked in these for what we characterise as more 'evaluative' material and judgements.

One of the places that we looked for this evidence was in standard DFID programme documentation. Overall, we found annually completed review documentation to reflect little on the contribution of programmes to processes of change or to judge the relative effects of different activities. They were often limited by a standard paperwork format that only included commentary and measurement at the 'output' level of the logical framework. Many of these were indicators of volume and lacked discussion of the quality or meaningfulness of the demands and responses they counted – for example what kinds of demands were being raised and the nature and extent of their resolution. One hope might be that review paperwork 'joins the dots' between indicators in monitoring frameworks to indicate linkages and processes of change. This, however, wasn't often the case. For example, where programmes involved outputs on both civic education/rights awareness activities and counting demands or issues raised with authorities, there was no discussion of whether the same people were involved in both kinds of activities.

More recent versions of the standard DFID paperwork for annual reviews include additional sections to assess progress towards higher level outcomes and impacts, but for several of the programmes reviewed consideration of these wider effects were notably absent. This was less the case in mid-term annual reviews, which were in places more detailed on contribution and change – either based on independent evaluation or led by DFID staff external to the operating context. Annual review documentation largely exhibited similar upward accountability and project control characteristics to logical frameworks, and was spoken about in similar terms by programme practitioners. The function of the annual review to agree an annual performance 'grade' for programmes, affecting payments to implementers and reputations of those involved, was reported to overtake the desire for a broader understanding of changes taking place. We found programme completion reviews, which follow a format that is designed to prompt analysis on programme contributions to pre-identified outcomes, more likely to engage in dynamics of change and interactions of activities with context, and a wider set of evidence. However, they did not often address multi-scalar or multi-level project designs in great depth or explicitly.

We were also able to review a number of independent programme evaluations, and a volume of evidence produced by the programme implementer themselves. Table 8.2 summarises these sources, although it is important to say that not all of this material was available for us to review. As noted earlier evaluations did not exist for all programmes. This is in line with DFID's overall strategy on programme evaluation, which emphasises programme-level decisions on evaluation approaches (summarised in Box 8.1).

### **Box 8.1 DFID's approach to programme evaluation**

DFID guidance uses the OECD Development Assistance Committee definition of evaluation as the 'systematic and objective assessment of an on-going and completed project, programme or policy, its design, implementation and results in relation to specified evaluation criteria', and emphasises that this happens within programmes, not only through external assessments (DFID 2013b). Programme evaluation in DFID largely follows what it calls a 'decentralised' model in which those responsible for designing and managing programmes also have the responsibility for building in evaluation activities (DFID 2013b). This approach reportedly 'helps increase programme team ownership of the evaluation findings and helps facilitate the learning being applied to improve programme implementation' (DFID 2019a: 7). The 2017/18 report from the central team responsible for supporting evaluation practice reports that 15 per cent of current programmes were being evaluated, and 30 per cent of bilateral spending was subject to evaluation processes (DFID 2019a: 6).

**Table 8.2 Programme evidence – evaluations and other non-logframe sources**

Programme	Evaluation approach	Other evidence sources
AAWAZ	Initially independent monitoring and evaluation component, but absorbed within programme management during implementation and no evidence of programme evaluation	Large number of published reports, although little exploration of programme theory.
Alif Ailaan	No evidence of programme evaluation outside of logframe.	Reportedly, a significant monitoring database recording individual advocacy engagements. Independently commissioned ‘tracker studies’ on public and politician attitudes to inform logical framework.
BCSSP	Independent mid-term and final impact evaluation.	Detailed impact assessment documents from separate projects.
CEP	Independent ‘Monitoring, Evaluation and Learning’ organisation contracted throughout. Baseline study, mid-term, and final evaluations completed.	Detailed database recording outcomes of community scorecard processes.
Diálogo	Independent ‘Monitoring, Evaluation and Learning’ organisations contracted throughout. Baseline study, mid-term, and final evaluations completed.	Separate ‘ripple effect’ study on spill-over benefits.
EVA-BHN	DFID-funded independent assessment completed following request during delivery from programme implementer.	Tracer studies on impact completed by programme management. Detailed database recording issues raised at community level and their resolution. Academic working paper on programme approach to Social Accountability (Kirk 2017)

The exemplar programmes in terms of evaluation evidence were those in Mozambique, although it should also be noted that these were not publicly available during our research. Both Mozambique programmes had baseline, mid-term and final evaluations undertaken by reviewers with some independence from the activity, appointed at the start of the programmes, and able to build up evidence and recommendations and engage in programme learning and adjustment over time. Both final evaluations (Brook *et al.* 2017; Shankland *et al.* 2017) engaged with the theory of change that underpinned the programmes and sought to assess evidence for the relationships and causality assumed in the programme design. This entailed actively looking for multi-level dynamics and the outcomes of multi-scalar strategies. The BCSSP final evaluation was not available for review, but the mid-term evaluation did look in depth at processes of change – for example the creation of social capital of different kinds across scales (DFID 2013a). It was limited however by the extent and kind of evidence collected from the monitoring scheme, and did not involve the degree of independent research and analysis that the Mozambique programmes did.

The programmes in Pakistan had the most limited independent or external evaluation. The AAWAZ programme was originally designed to have an independent monitoring component, but this was later subsumed into the programme management role (DFID 2018c; Esser 2018). The EVA-BHN programme was reviewed in a detailed independent assessment (Taylor and Khan 2017) that programme implementers argued was funded as a result of direct requests by the programme to DFID to help them evidence the importance of their contribution (Esser 2018). The assessment explored in some depth the kinds of issues raised at community level and where in the system these were resolved. It also looked substantially at the impact on wider system reform that might be attributed to the programme. It is notable however that this piece of work was largely framed in terms of health and health systems outcomes, rather than the governance outcomes we discuss here.

As a consequence, it re-figures 'citizen demands' from the original intention of citizens making claims on public services to 'service demand' in the sense of increased health centre use. This study was not an impact evaluation, although the content and analysis have evaluative characteristics.

As indicated in Table 8.2, programmes reported other substantial sources of potential evidence outside of DFID programme monitoring and commissioned evaluation. We found that programme implementers had far more extensive evidence from their activities than were used in standard reporting or indeed in external evaluations where that took place. This supports a finding from a 2014 review on DFID learning processes, which argued that the lack of ability of DFID to use the depth of evidence generated by third-party programme implementers was a 'key deficit' (ICAI 2014: 27). The result of this is that whilst DFID documentation often only focused on one part of the multi-level system or a volume indicator (for example, number of issues raised), implementer reports and databases sometimes took a broader view – identifying patterns of issues raised and what issues were resolved (see for example Palladium (2017)). The Pyoe Pin component of the BCSSP produced impact case studies that addressed multi-level effects (British Council 2016b). The EVA-BHN programme developed a set of 'tracer studies' that explicitly traced the links and explored processes between activities at different levels, gathering qualitative evidence to show these interactions (EVA-BHN 2017a, 2017b, 2017c, 2017d). In both cases this content was not referred to substantially in DFID documentation.

The programmes based on aggregating citizen demands – AAWAZ, CEP and EVA-BHN – all reported using specially developed databases tracking issues and their resolution across levels and scales. As noted, in the CEP programme the internal database was an important part of the vertical integration system itself, used by national partners to identify patterns and provide evidence for advocacy (Brook *et al.* 2017). In the case of Alif Ailaan, programme implementers reported a record of hundreds of meetings with officials, the commitments they made, and evidence that they had followed up on these (Esser 2018: 20). We were not able to interrogate these databases, or find significant evidence of them being used analytically to understand change pathways. However, they indicate a potential future source of analysis. Esser argues that the EVA-BHN and Alif Ailaan programmes have monitoring data that could hold 'significant analytical potential' for understanding how multi-level processes do and don't lead to change (2018: 20). The same is arguably true of the CEP-created database, which was expanded to also include the results of other social accountability programming in Mozambique.

### **8.3 Implications for learning**

DFID programmes are encouraged to develop contextually and programmatically suitable measures of progress and change (DFID 2019b: 62). The discussion in this section so far illustrates that the approaches to monitoring and evaluation within the programmes we looked at were accordingly reasonably diverse. Whilst the choice and comparability of measurement indicators and evaluation approaches might seem a fairly technical concern it has important ramifications.

First, feedback we gathered from those designing and working within the programmes reflected a concern about whether, in fact, the activities were being judged against the right kinds of impacts and changes, or simply those things that were easiest to measure. Esser reports frustration amongst practitioners at what they saw as 'the donor's' quantitative measurement emphasis (2018: 22). DFID participants in our discussions noted the difficulties of instituting the kinds of measurement that they felt were more appropriate for these programme activities, and the realities that that what appears in official measurement often



responds to upward accountability and project management needs over learning<sup>10</sup>. A DFID case study on the monitoring and evaluation approach of the SAVI programme in Nigeria, included in our document review but not a case study programme, notes that over time the programme was pushed to develop indicators that might be simpler to understand than its sophisticated qualitative scales of behaviour change, such as the passage of specific pieces of legislation (Derbyshire and Donovan 2016: 7). This opens up a potential gap between what programmes are measured on and what they are actually seeking to do – one that undermines their contribution to the evidence base as well as producing ongoing tension between practices and measurements of success.

Second, the devolved nature of programme measurement and learning and the lack of interrogation of causal pathways and change strategies in programme documentation limits how far the resemblances we have established between the programmes allows real comparison. The greatest potential for more comparative learning might lie in the detailed programme monitoring systems and databases. However, these are not considered 'products' of the programme and are owned by competing implementing organisations. Undertaking this kind of analysis would require new levels of collaboration that go beyond the immediate needs of those running projects. This has ramifications for the design of any research seeking to learn from these activities and significant investments. Whilst we might not expect comparability of programmes in terms of precisely the same indicators or measures, it does seem reasonable that evidence could be available to shape into a comparative framework, which we didn't find to be the case.

Third, and associated with this last point, our review of the measurement of change in these programmes and discussions with practitioners suggests barriers to sharing and interrogating evidence that emerge from organisational politics and commercial competition. Various contributions to the study of international development programmes have highlighted the macro and micro 'politics of evaluation' (O'Brien *et al.* 2010; Roche and Kelly 2012). Rigorous evaluative consideration of programme strategies comes with the risk that the planned effects aren't found, and the risk of appearing to 'fail' has been noted elsewhere in terms of constraints to DFID evaluation practice (ICAI 2014). In a recent contribution on the politics of development assistance, Yanguas (2018: 163) notes the tight control of political economy analyses and other internal knowledge within DFID field offices. Whilst DFID's reputation amongst donors and with national governments can be affected by negative evaluation results, arguably the strongest dynamic against transparent evaluation that emerged from our work was the extent of commercial competition between implementing organisations. Our case study work in Mozambique and Pakistan highlighted the level of competition for resources between not only the large main contractors of DFID, but also amongst the collection of local NGOs and CSOs that are then their sub-contractors. This environment means that learning how to do this work well has a commercial value. It also produces a bias towards 'success stories' that may limit learning from when similar strategies don't work.

It is important to say that we found an appetite for learning and examples of positive collaboration amongst practitioners across the programmes despite these larger organisational or institutional barriers. This was particularly apparent when we brought together practitioners from programmes run by different providers, along with DFID programme managers as part of our research. This suggested that a form of community of practice does exist, transferring learning across programmes, and in particular as practitioners move between them. In these discussions, practitioners were keen and happy to put the competitive pressures aside in this kind of facilitated and closed-door discussion.

---

<sup>10</sup> Eyben, based on experience working within DFID, is of the view that these pressures and the logical framework approach taken 'normalise inappropriate means for designing and assessing development programmes' (Eyben 2015: 34).

However, this enthusiasm does relatively little to support a grounded, open and systematic analysis of what kinds of strategies work to address governance challenges in complex settings.

## 9 Conclusions and implications

Have these programmes shown the adoption and effectiveness of multi-scalar strategies in donor-supported empowerment and accountability activities? How far do they contribute to an evidence base for working in this way as a means to build countervailing power and increase accountability of decision makers to the wider population? In concluding, we assess our findings against the driving questions of our study, and draw out some implications both for further study and the field more widely.

We found many of the programmes adopted multi-scalar strategies. This should not be taken for granted and indicates an awareness of those involved of the need to be more strategic and work across the governance landscape to achieve change. In the case of more social accountability focused activities in particular it represents a welcome acknowledgement of the existing evidence base on 'low accountability traps' and an attempt to overcome the limitations of earlier social accountability programming and donor activity that assumed that stimulating citizen voice was either easy or sufficient in itself to lead to change. In other programmes the multi-scalar approaches taken also seem important to their logic and practice although these might not have been so overtly stated. This serves as a reminder that the 'how' of building pro-accountability reform deserves more detailed attention than the broad appeals in the existing literature and guidance to 'work politically' and take 'adaptive' approaches. The programmes remind us that building countervailing power, in particular of previously excluded or marginalised groups, needs to be viewed not in terms of technical fixes but in terms of what strategies work best in campaigns and mobilisations to challenge inertia and embedded narrow interests. Within this, the kinds of multi-level strategies recommended by a part of the existing literature clearly deserve more attention and bringing to the fore, and our study of a limited number of programmes suggests that the practice – if not the evidence – is out there.

Exploring these programmes has also helped us to begin to clarify some different ways in which multi-scalar strategies might be incorporated into local campaigns, whether donor-funded or otherwise. Although potentially complementary, there is a distinct difference between using vertically-integrated strategies to empower people to make appeals upward and activate accountability in various places in a bureaucratic and policy hierarchy on the one hand, and joining local observers and campaigners with experienced national policy advocates on the other. It is interesting that only one programme we reviewed – the Citizen's Engagement Programme in Mozambique – actually tried to do both of these in substantial ways, rather than focusing on one over the other. The distinction between these and horizontal scale is more obvious, but it is also interesting that horizontal scale was the least prioritised in programme designs as far as we could see, and limited in some cases to looking for replication of successful activities. This disaggregation could helpfully be applied to other cases of action or donor programming. Given what we've seen, there may be a particular advantage in bringing the literature on the horizontal spread and diffusion of movements and spatialities of activism to the fore in the analysis of accountability reform strategies. There could also be more done to understand the longer-run impacts through 'ripple effects'.

Looking at these programmes also gives us some further ideas on why multi-scalar strategies might be worth trying in promoting greater accountability, and where we should look to trace how they work. By synthesising from programme reports to disaggregate three

key areas of 'effects' we have been able to draw attention to a number of ways that they seem beneficial. This highlights a number of framings and propositions that might call for further examination. Whilst we don't have an evidence base from these programmes to test the proposition of Fox (2016, Fox *et al.* 2016) that vertically integrated approaches offer more purchase than less integrated alternatives, we can see within recent practice some examples emerge that could be added to a larger set of cases to start this testing. We can also add to the proposition with some ideas and further evidence on why this might be the case; the importance of higher level authority oversight and its existence as an enforcer of local responsiveness, the availability of recourse where this isn't possible, and the provision of multiple channels through which blockages and resistance might be circumvented. The effects we found on solidarity and legitimacy – aspects of the strength of citizen's collective actions – are relatively absent from the language used to assess and understand donor-led programmes. There are probably good reasons not to use these more overtly political terms in official programme documents, but it might be fruitful to take this lens across a wider set of cases, recognising that these two factors may not be explicit. Finally, there seems scope to look more precisely at the different ways in which policy reform can be leveraged and have impact. The programmes used multi-level strategies not only to attempt reforms to the content of policy, but also to the nature of state-citizen relations. This offers an important reminder that the effects of these programmes and sets of activities are likely to be much longer run than their periods of funding and the tracing of immediate impact. In the case of Pakistan, we also see how likely it is that these sets of activity have some interaction themselves, although we weren't able to explore this.

There is a substantial caveat to these conclusions, however. We hoped that in the monitoring and evaluation schemes and evidence generated from a sample of programmes we could find sufficiently robust and comparable evidence to make some broad statements about what is more effective, and how this interacts with difficult and unpredictable contexts. This wasn't the case. The available monitoring and evaluation of the programmes was insufficient to make sound judgements on how well multi-scalar strategies work and their impacts on accountability relationships in these externally supported initiatives. There were three main limitations here.

The first is the simple availability of evidence, with the detail of what went on in these initiatives hard to come by and assess, despite DFID's commitment to transparency and desire to learn from practice. Whilst some of this is driven by inconsistent practices in publishing relevant material, some certainly appears to be driven by the nature of competitive incentives. Doing this work well gives commercial advantage for DFID sub-contractors and is part of how DFID programme managers and country offices are judged. Independent interrogation of evidence is not necessarily in many people's interests, and evidence of success or otherwise becomes commercially sensitive.

The second limitation is the disjointed nature of most programmes' monitoring and evaluation schemes, and incongruities between what practitioners say they are setting out to achieve and what is measured and published. These are sharpest in relation to the outcomes of the programmes. Implementers say that they are seeking longer-run changes in state-society relations and balances of power and that sustainability comes about through ongoing patterns of pro-accountability behaviours, conditions, and re-shaped norms, rather than the continuation of mechanisms that they have used in the short-run to motivate these behaviours. Those external to the programmes, including sometimes DFID managers and others involved judging programme success, often see sustainability in terms of those mechanisms being institutionalised so that they no longer need to be donor-funded and 'scaled up' to reach more people. This is a wider issue in the politics and incentives of development donor management, and what became referred to in our practitioner workshop as 'the invisible power of the logframe'.

Partly driven by these first two factors, the final limitation here is that we found little scope for robust comparison between programmes' evidence. Thus, the similarity of these programmes in design and intent does not result in comparability in research terms. Whilst DFID has invested in other important ways to develop the evidence base for empowerment and accountability programmes, we see missed opportunities to learn from the programmes we've explored and use their activities as part of that evidence base. It is doubtless important that programmes develop contextually suitable schemes to generate evidence of what is working and not, and DFID's decentralised approach to programme design and evaluation activity supports this. However, the latitude provided in this approach has important implications for the wider field of study and also for DFID's ambitions to engage in evidence-informed programming and to ensure learning from its work. From our study there are clear and interesting examples of applying what has been learnt from previous systematic study of empowerment and accountability programmes, but several of these didn't seem to provide the kinds of evidence that could expand this systematisation of knowledge and practice.

All of the programmes that we explored identified in their documentation conditions associated with fragility, conflict and violence. This was particularly the case in evidence related to programme design. We see relatively little direct engagement with these conditions in more retrospective analysis, although arguably they were attempting to address key characteristics of fragility in developing new state-society relationships and (in line with the DFID theory of change) widening the political settlement. This is perhaps not surprising given the focus of the activity on particular countries, localities, and sectors, and that addressing these broader conditions were not often identified in programme monitoring and evaluation frameworks. However, a number of points emerge from our analysis that might indicate what is especially important in considering multi-level and multi-scalar empowerment and accountability in conditions of fragility, conflict and violence.

First there are a set of implications for programme design and implementation. One is that multi-level strategies and designs need to be based on where decision-making and authority lie. The examples we have looked at highlight how far this is likely to be in flux as a result of the contestations going on in political and civil society. This identification of sites of action also highlights the complexities of seeking to use existing governance spaces – which may be subject to elite capture or co-option, to be taking a particular 'side' in a conflict or may maintain exclusionary norms – versus creating new spaces. In both the Pyoe Pin and EVA-BHN strategies new spaces with newly inclusive norms were overtly part of the strategy. Strategies in Mozambique tended to focus on expanding practice in parts of the governance structure seen as most malleable – co-management committees of local services and in municipal governance. Associated with these points, several of the programmes provide examples of the importance of informal power and powerholders, and how they can be targeted as part of a multi-level strategy.

Second, our analysis of the effects of multi-scalar strategies emphasises coalition-building practices and outcomes that might be particularly important where actors have been fragmented and the distribution of legitimate authority and representation is unsettled. The practical implications of this are less clear, but the importance that emerges here suggests some question about how deliberately programmes approach the issues of building power within and through coalitions. This touches on issues of the underpinning theories of change in complex contexts, and through what mechanisms programmes might achieve 'scale' in the sense of an impact wider than their particular sub-national sites or sectors. In contexts where processes of social change might be less predictable, or the channels through which they operate more fluid, we see some challenges to declared ambitions that geographically limited pilots might be enthusiastically embraced, and in practice a perhaps more realistic view of programme intervention as a series of 'nudges' to the (assumed) social contract.

Taken together, the results of our study suggest that in important ways DFID empowerment and accountability programmes are taking scale into account – even in contexts with particularly complex and conflictual governance arrangements. Incorporating multi-scalar strategies may be producing important effects. They may also in the future offer lessons for other processes of social change and establishment and mobilisation of countervailing power. There is a clear case for applying the lens of scale, as we have used it here, to other cases, and finding ways to get the evidence that allows a more robustly comparative study. However, the evidence needed to make the kinds of judgements required on dynamics of change seems either not gathered or not shared in current DFID and programme implementer practice. Advancing the field therefore requires concerted collaboration of practitioners and researchers or significant interventions on the part of funders to make a more secure contribution to evidence. Only if monitoring and learning investments are more aligned with tracking effects of these strategies can we answer important and pressing questions of how donors can support citizens in navigating governance systems to make them work in their favour, and contribute to the significant shifts of power needed to overcome anti-accountability forces.

# References

- Aceron, J. (ed.) (2018) *Going Vertical: Citizen-led Reform Campaigns in the Philippines*, 2nd ed., Manila, Philippines: G-Watch, [www.g-watch.org/resources/vertical-integration-research/going-vertical-citizen-led-reform-campaigns-philippines-2nd](http://www.g-watch.org/resources/vertical-integration-research/going-vertical-citizen-led-reform-campaigns-philippines-2nd) (accessed 10 June 2018)
- Aceron, J. and Isaac, F. (2016) *Going Vertical: Citizen-led Reform Campaigns in the Philippines*, Manila, Philippines: G-Watch, <https://opendocs.ids.ac.uk/opendocs/handle/123456789/12718> (accessed 27 August 2019)
- Agrawal, A. and Ribot, J. (1999) 'Accountability in Decentralization: A Framework with South Asian and West African Cases', *The Journal of Developing Areas* 33.4: 473–502
- Alif Ailaan (2015) *Going Local for Data: Union Council-led Education Statistics and How They can Inform Policy*, Islamabad: Alif Ailaan
- Amakom, U.; Fashola, T.; Gay, C. and Shutt, C. (2018) 'Doing Accountability Differently: A "Vertically Integrated" Approach. Exploring Practice and Results of Christian Aid's Voice to the People Programme in Anamabra State, Nigeria', January, Abuja: Christian Aid, [www.christianaid.org.uk/sites/default/files/2018-03/Doing-Accountability-Differently-V2P-Governance-January2018.pdf](http://www.christianaid.org.uk/sites/default/files/2018-03/Doing-Accountability-Differently-V2P-Governance-January2018.pdf) (accessed 27 August 2019)
- Blackburn, J.; Chambers, R. and Gaventa, J. (2002) 'Mainstreaming Participation in Development', in N. Hanna and R. Picciotto (eds), *Making Development Work: Development Learning in a World of Poverty and Wealth*, New Jersey: Transaction Publishers
- Booth, D. (2012) 'Development as a Collective Action Problem', *Africa Power and Politics Programme Policy Brief 9*, London: Overseas Development Institute, <https://assets.publishing.service.gov.uk/media/57a08a7040f0b652dd000722/app-synthesis-report-development-as-a-collective-action-problem.pdf> (accessed 3 March 2019)
- British Council (2016a) *Pyoe Pin Phase 2 Programme Completion Report: October 2011 to December 2016*, London: British Council
- British Council (2016b) 'Pyoe Pin Phase II Programme Completion Report', Annex 1 Issue Based Projects Case Stories, December, London: British Council
- Brook, S.; Holland, J. and Macuane, J. (2017) *Mozambique Citizen Engagement Programme (CEP): Final Evaluation Report*, Oxford: Oxford Policy Management
- Carothers, T. and Brechenmacher, S. (2014) *Accountability, Transparency, Participation, and Inclusion: A New Development Consensus*, Washington DC: Carnegie Endowment for International Peace, <http://carnegieendowment.org/2014/10/20/accountability-transparency-participation-and-inclusion-new-development-consensus-pub-56968> (accessed 3 March 2019)
- Christian Aid (2017) 'The Art of Networking: A Study of Civil Society Networks in Myanmar', September, Christian Aid, <https://www.christianaid.org.uk/resources/about-us/art-networking-study-civil-society-networks-myanmar> (accessed 22 May 2018)

- Christie, A. and Burge, R. (2017) *The Role of External Actors in Supporting Social and Political Action towards Empowerment and Accountability with a Focus on Fragile, Conflict-and Violence-Affected Settings*, IDS Working Paper 503, Brighton: IDS
- Christie, A. and Green, D. (2018) 'Adaptive Programming in Fragile, Conflict and Violence-Affected Settings, What Works and Under What Conditions?: The Case of Pyoe Pin, Myanmar', Itad and Oxfam in association with A4EA, <https://opendocs.ids.ac.uk/opendocs/handle/123456789/13888> (accessed 24 August 2019)
- Cooley, L. and Kohl, R. (2006) *Scaling Up – From Vision to Large-scale Change: A Management Framework for Practitioners*, Washington DC: Management Systems International
- Cooley, L. and Papoulidis, J. (2017) 'Tipping the Scales: Shifting from Projects to Scalable Solutions in Fragile States', *Development* 60.3–4: 190–6
- DAI (2015) 'AAWAZ Theory of Change', February, Bethesda, Maryland: Development Alternatives Inc.
- Dasandi, N.; Laws, E.; Marquette H. and Robinson, M. (2019) 'What Does the Evidence Tell Us about 'Thinking and Working Politically' in Development Assistance?', *Politics and Governance* 7.2: 155–68
- Derbyshire, H. and Donovan, E. (2016) *Adaptive Programming in Practice: Shared Lessons from the DFID-funded LASER and SAVI programmes*, SAVI; LASER, [http://savi-nigeria.org/wp-content/uploads/2016/08/Laser\\_Savi\\_Report-online-version-final-120816pdf.pdf](http://savi-nigeria.org/wp-content/uploads/2016/08/Laser_Savi_Report-online-version-final-120816pdf.pdf) (accessed 3 March 2020)
- DFID (2007) *Governance, Development and Democratic Politics: DFID's Work in Building More Effective States*, London: Department for International Development
- DFID (2011a) *Business Case: AAWAZ Voice and Accountability Programme*, August, London: Department for International Development
- DFID (2011b) *Business Case: Civil Society Strengthening Programme*, London: Department for International Development
- DFID (2011c) *DFID Empowerment and Accountability Theory of Change and Background*, London: Department for International Development
- DFID (2012a) *Business Case: Mozambique Citizen Engagement Programme*, London: Department for International Development, [http://iati.dfid.gov.uk/iati\\_documents/3719299.odt](http://iati.dfid.gov.uk/iati_documents/3719299.odt) (accessed 3 March 2020)
- DFID (2012b) *Business Case: Mozambique Democratic Governance Support Programme*, November, London: Department for International Development
- DFID (2012c) *Business Case: Provincial Health and Nutrition Programme*, December, London: Department for International Development
- DFID (2012d) *Business Case: Transforming Education in Pakistan*, London: Department for International Development, [http://iati.dfid.gov.uk/iati\\_documents/3719222.odt](http://iati.dfid.gov.uk/iati_documents/3719222.odt) (accessed 3 March 2020)
- DFID (2013a) *Burma Civil Society Strengthening Programme Mid-Term Review*, July, London: Department for International Development

- DFID (2013b) *International Development Evaluation Policy*, London: Department for International Development
- DFID (2014) *Terms of Reference: Empowerment, Voice and Accountability for Better Health and Nutrition (EVA-BHN)*, London: Department for International Development
- DFID (2016a) *Building Stability Framework*, London: Department for International Development,  
<https://assets.publishing.service.gov.uk/media/5968990ded915d0baf00019e/UK-Aid-Connect-Stability-Framework.pdf> (accessed 1 August 2019)
- DFID (2016b) *Programme Completion Review: Burma Civil Society Strengthening Programme*, December, London: Department for International Development,  
[http://iati.dfid.gov.uk/iati\\_documents/7294937.odt](http://iati.dfid.gov.uk/iati_documents/7294937.odt) (accessed 3 March 2020)
- DFID (2017a) *Annual Review: AAWAZ Voice and Accountability Programme*, London: Department for International Development
- DFID (2017b) *Annual Review: Transforming Education in Pakistan*, June, London: Department for International Development,  
[http://iati.dfid.gov.uk/iati\\_documents/15250669.odt](http://iati.dfid.gov.uk/iati_documents/15250669.odt) (accessed 3 March 2020)
- DFID (2017c) *Transforming Education in Pakistan: Addendum to Business Case for Bridge Funding and Scale-Up Cost Extensions*, London: Department for International Development,  
[http://iati.dfid.gov.uk/iati\\_documents/20382378.odt](http://iati.dfid.gov.uk/iati_documents/20382378.odt)
- DFID (2018a) *Annual Review: Provincial Health and Nutrition Programme*, March, London: Department for International Development
- DFID (2018b) *Logical Framework – Pakistan Provincial Health and Nutrition Programme*, London: Department for International Development
- DFID (2018c) *Open Aid, Open Societies: A Vision for a Transparent World*, February, London: Department for International Development,  
[www.gov.uk/government/publications/open-aid-open-societies-a-vision-for-a-transparent-world](http://www.gov.uk/government/publications/open-aid-open-societies-a-vision-for-a-transparent-world) (accessed 2 August 2019)
- DFID (2018d) *Programme Completion Review: AAWAZ Voice and Accountability Programme*, June, London: Department for International Development,  
[http://iati.dfid.gov.uk/iati\\_documents/34328231.odt](http://iati.dfid.gov.uk/iati_documents/34328231.odt) (accessed 3 March 2020)
- DFID (2018e) *Programme Completion Review: Citizen Engagement Programme*, March, London: Department for International Development
- DFID (2018f) *Programme Completion Review: Democratic Local Governance Support Programme (Diálogo)*, March, London: Department for International Development,  
[http://iati.dfid.gov.uk/iati\\_documents/28432092.odt](http://iati.dfid.gov.uk/iati_documents/28432092.odt) (accessed 3 March 2020)
- DFID (2019a) *Evaluation Annual Report 2017-18*, London: Department for International Development
- DFID (2019b) *Governance for Growth, Stability and Inclusive Development*, DFID Position Paper, March, London: Department for International Development
- DFID (2019c) *Smart Rules: Better Programme Delivery – Version X Effective January 2019 until 1st April 2019*, London: Department for International Development



- Dias, J. and Tomé, T. (2018) 'Inverted State and Citizens' Roles in the Mozambican Health Sector', *IDS Bulletin* 49.2: 35–48
- e-Pact (2016) *Macro Evaluation of DFID's Policy Frame for Empowerment: Empowerment and Accountability Annual Technical Report 2016: What Works for Social Accountability*, Oxford: e-Pact Consortium, <https://itad.com/wp-content/uploads/2017/06/EA-Macro-Evaluation-Technical-report-Dec16-FINAL.pdf> (accessed 1 August 2019)
- Esser, D. (2018) *How do DFID E&A Projects in Pakistan Work Across Scales? Report on A4EA Pilot Research in Pakistan*, Washington DC: Accountability Research Center
- EVA-BHN (2017a) *Finding Alternatives: Operationalization of Basic Health Unit (BHU) in Koragh, Mardan through Local Government Funds*, 1, Tracer Study, Karachi: Palladium
- EVA-BHN (2017b) *Improving Citizens Access to their Right to Health*, 5, Tracer Study, Karachi: Palladium
- EVA-BHN (2017c) *Raising Voices, Challenging Discrimination! Denial of Vaccination at Basic Health Unit Shahpur in Layyah*, 2, Tracer Study, Karachi: Palladium
- EVA-BHN (2017d) *The Change Within: Community Efforts for Ultrasound Facility in Shamkay Bhattian*, 3, Tracer Study, Karachi: Palladium
- Eyben, R. (2015) 'Uncovering the Politics of Evidence and Results', in R. Eyben, I. Guijt, C. Roche and C. Shutt (eds), *The Politics of Evidence and Results in International Development: Playing the Game to Change the Rules?*, Rugby: Practical Action Publishing
- Eyben, R.; Guijt, I.; Roche, C. and Shutt, C. (2015) *The Politics of Evidence and Results in International Development: Playing the Game to Change the Rules?*, Rugby: Practical Action Publishing, [www.ids.ac.uk/publications/the-politics-of-evidence-and-results-in-international-development-playing-the-game-to-change-the-rules/](http://www.ids.ac.uk/publications/the-politics-of-evidence-and-results-in-international-development-playing-the-game-to-change-the-rules/) (accessed 3 March 2020)
- Fooks, L. (2013) *Within and Without the State: Governance and Fragility: What We Know about Effective Governance Programming in Fragile Contexts*, Oxford: Oxfam, <http://policy-practice.oxfam.org.uk/publications/governance-and-fragility-what-we-know-about-effective-governance-programming-in-306683> (accessed 18 April 2018)
- Fox, J. (2001) 'Vertically Integrated Policy Monitoring: A Tool for Civil Society Policy Advocacy', *Nonprofit and Voluntary Sector Quarterly* 30.3: 616–27
- Fox, J. (2007a) *Accountability Politics: Power and Voice in Rural Mexico*, Oxford Studies in Democratization, Oxford: Oxford University Press
- Fox, J. (2007b) 'The Uncertain Relationship between Transparency and Accountability', *Development in Practice*, 17.4–5: 663–71
- Fox, J. (2015) 'Social Accountability: What Does the Evidence Really Say?', *World Development*, 72: 346–61
- Fox, J. (2016) *Scaling Accountability through Vertically Integrated Civil Society Policy Monitoring and Advocacy*, Making All Voices Count Research Report, Brighton: IDS, <https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/12683> (accessed 3 April 2020)

- Fox, J.; Acheron, J. and Montero, A.G. (2016) 'Doing Accountability Differently', August, Chr. Michelsen Institute, [www.u4.no/publications/doing-accountability-differently-a-proposal-for-the-vertical-integration-of-civil-society-monitoring-and-advocacy/](http://www.u4.no/publications/doing-accountability-differently-a-proposal-for-the-vertical-integration-of-civil-society-monitoring-and-advocacy/) (accessed 3 April 2020)
- Fox, J. (2010) *Coalitions and Networks*, UC Santa Cruz Working Paper, <https://escholarship.org/uc/item/1x05031j> (accessed 4 December 2019)
- Fox, J. (forthcoming) 'Can "Participatory" International Development Projects Contribute to Building Countervailing Power for the Excluded? Lessons from Outlier World Bank Projects', *World Development*
- Fung, A. and Wright, E.O. (2003) *Deepening Democracy: Institutional Innovations in Empowered Participatory Governance*, The Real Utopias Project 4, London: Verso
- Gaventa, J. and McGee, R. (2013) 'The Impact of Transparency and Accountability Initiatives', *Development Policy Review*, 31.s1, <http://onlinelibrary.wiley.com/doi/10.1111/dpr.12017/full> (accessed 3 April 2020)
- Gibson, C. and Woolcock, M. (2008) 'Empowerment, Deliberative Development, and Local-Level Politics in Indonesia: Participatory Projects as a Source of Countervailing Power', *Studies in Comparative International Development* 43.2: 151
- Goetz, A.M. and Jenkins, R. (2005) *Reinventing Accountability: Making Democracy Work for Human Development*, London: Palgrave Macmillan, [www.palgrave.com/gp/book/9781403906243](http://www.palgrave.com/gp/book/9781403906243) (accessed 3 April 2020)
- Grandvoinet, H.; Aslam, G. and Raha, S. (2015) *Opening the Black Box: The Contextual Drivers of Social Accountability*, Washington DC: World Bank Publications
- Hartmann, A. and Linn, J.F. (2008) *Scaling Up: A Framework and Lessons for Development Effectiveness from Literature and Practice*, Wolfensohn Center for Development Working Paper 5, Washington DC: The Brookings Institution
- Holland, J.; Thirkell, A.; Trepanier, E. and Earle, L. (2009) *Measuring Change and Results in Voice and Accountability Work*, DFID Working Paper 34, <https://gsdrc.org/document-library/measuring-change-and-results-in-voice-and-accountability-work/> (accessed 3 April 2020)
- ICAI (2013a) *DFID's Empowerment and Accountability Programming in Ghana and Malawi*, October, London: Independent Commission on Aid Impact, <https://icai.independent.gov.uk/wp-content/uploads/Empowerment-and-Accountability-081013-FINAL.pdf> (accessed 3 April 2020)
- ICAI (2013b) *DFID's Empowerment and Accountability Programming in Ghana and Malawi*, October, London: Independent Commission on Aid Impact
- ICAI (2014) *How DFID Learns*, ICAI Report 34, London: Independent Commission on Aid Impact
- ICAI (2018) *DFID's Governance Programming in Nepal and Uganda*, London: Independent Commission on Aid Impact
- ICAI (2019) *How UK Aid Learns*, September, London: Independent Commission on Aid Impact, <https://icai.independent.gov.uk/report/how-uk-aid-learns/> (accessed 12 October 2019)

- IDS (2016) *Mid-Term Review: Diálogo*, February, Brighton: IDS
- Joshi, A. (2013) 'Do They Work? Assessing the Impact of Transparency and Accountability Initiatives in Service Delivery', *Development Policy Review* 31.1: 29–48
- Joshi, A. (2014) 'On Social Accountability: An Issues Paper', unpublished manuscript
- Joshi, A. and Moore, M. (2000) 'Enabling Environments: Do Anti-Poverty Programmes Mobilise the Poor?', *Journal of Development Studies* 37.1: 25–56
- Joshi, A. and Moore, M. (2004) 'Institutionalised Co-production: Unorthodox Public Service Delivery in Challenging Environments', *Journal of Development Studies* 40.4: 31–49
- Justino, P. (2019) 'Governance Interventions in Conflict-Affected Countries', *Journal of Development Studies* 55.7: 1364–78
- Keck, M.E. and Sikkink, K. (1998) 'Transnational Advocacy Networks in the Movement Society', in D.S. Meyer and S. Tarrow (eds), *The Social Movement Society: Contentious Politics for a New Century*, Boulder CO: Rowman & Littlefield Publishers
- Kirk, T. (2017) *A Strategic Approach to Social Accountability in Pakistan*, IDS Working Paper 497, Brighton: IDS, <https://opendocs.ids.ac.uk/opendocs/handle/123456789/13275> (accessed 2 August 2019)
- Linn, J.F. (2012) *Scaling up in Agriculture, Rural Development, and Nutrition*, Washington DC: International Food Policy Research Institute
- Lopez Franco, E. (2018) *Interview Analysis – Taking Scale into Account Mozambique Case Study*, Brighton: IDS
- McCulloch, N. and Piron, L-H. (2019) 'Thinking and Working Politically: Learning from Practice', Overview to Special Issue, *Development Policy Review* 37.S1: O1–O15
- McGee, R. and Gaventa, J. (2011) *Shifting Power? Assessing the Impact of Transparency and Accountability Initiatives*, IDS Working Paper 383, Brighton: IDS
- McLean, R.; Gargani, J. and Gargani, J. (2019) *Scaling Impact : Innovation for the Public Good*, Abingdon: Routledge
- O'Brien, T.; Payne, S.; Nolan, M. and Ingleton, C. (2010) 'Unpacking the Politics of Evaluation: A Dramaturgical Analysis', *Evaluation*, 16.4: 431–44
- O'Donnell, G.A. (1998) 'Horizontal Accountability in New Democracies', *Journal of Democracy* 9.3: 112–26
- O'Meally, S.C. (2013) *Mapping Context for Social Accountability: A Resource Paper*, Washington DC: World Bank
- Palladium (2016a) *Pakistan's Health System: Understanding the Politics and Institutionalising a Culture of Accountability*, EVA Technical Paper 3, London: Palladium Group
- Palladium (2016b) *Reframing Citizen-State Relations from the Bottom-up: EVA-BHN's Model of Social Accountability in Pakistan*, EVA-BHN Technical Paper 1, August, London: Palladium Group

- Palladium (2017) *Year 3 Annual Report: Empowerment, Voice & Accountability for Better Health & Nutrition*, January, London: Palladium
- Rocha Menocal, A. and Sharma, B. (2009) *Joint Evaluation of Citizens' Voice and Accountability*, London: DFID, [www.odi.org/publications/2560-joint-evaluation-citizens-voice-and-accountability](http://www.odi.org/publications/2560-joint-evaluation-citizens-voice-and-accountability) (accessed 3 April 2020)
- Roche, C. and Kelly, L. (2012) *The Evaluation of Politics and the Politics of Evaluation, DLP Background Paper 11*, Developmental Leadership Programme, <https://www.dlprog.org/publications/research-papers/the-evaluation-of-politics-and-the-politics-of-evaluation> (accessed 3 April 2020)
- Shankland, A.; McGee, R.; Taela, K. *et al.* (2017) *Diálogo Democratic Governance Support Programme, Mozambique: Final Evaluation*, Brighton: IDS
- Shutt, C. (2016) *Towards an Alternative Development Management Paradigm?*, July, Stockholm: Expertgruppen för Biståndsanalys, <https://openaid.se/app/uploads/2015/03/2016-07-Towards-an-Alternative-Development-Management-Paradigm.pdf> (accessed 3 April 2020)
- Simmons, R.; Fajans, P. and Ghiron, L. (2007) *Scaling up Health Service Delivery: From Pilot Innovations to Policies and Programmes*, Geneva: World Health Organization
- Stern, E. *et al.* (2012) *Broadening the Range of Designs and Methods for Impact Evaluations*, DFID Working Paper 38, London: DFID
- Taylor, G. and Khan, S. (2017) *Empowerment, Voice and Accountability for Better Health and Nutrition (EVA-BHN) – Independent Assessment*, Oxford: HEART Consortium, <https://www.heart-resources.org/wp-content/uploads/2018/04/EVA-BHN-Independent-Assessment.pdf> (accessed 3 April 2020)
- Tembo, F. (2012) *Citizen Voice and State Accountability: Towards Theories of Change that Embrace Contextual Dynamics*, London: Overseas Development Institute
- Valters, C. and Whitty, B. (2017) *The Politics of the Results Agenda in DFID*, London: Overseas Development Institute, [www.odi.org/publications/10902-politics-results-agenda-dfid-1997-2017](http://www.odi.org/publications/10902-politics-results-agenda-dfid-1997-2017) (accessed 29 August 2019)
- VeneKlasen, L.; Miller, V.; Budlender D. and Clark, C. (2002) *A New Weave of Power, People & Politics: The Action Guide for Advocacy and Citizen Participation*, Oklahoma City: World Neighbors
- Vijge, M.J. (2018) 'The (Dis)empowering Effects of Transparency Beyond Information Disclosure: The Extractive Industries Transparency Initiative in Myanmar', *Global Environmental Politics* 18.1: 13–32
- Waddington, H. *et al.* (2019) *Does Incorporating Participation and Accountability Improve Development Outcomes? Meta-analysis and Framework Synthesis*, Systematic Review 43, London: International Initiative for Impact Evaluation (3ie)
- Ward, B. (2016) *Learning Together: Nurturing the Agency of Local Actors – Pyoe Pin*, London: British Council
- World Bank (2004) *Making Services Work for Poor People. World Development Report*, Washington DC: World Bank

Yanguas, P. (2018) *Why We Lie about Aid: Development and the Messy Politics of Change*,  
London: Zed Books

Youngs, R. (2019) *Civic Activism Unleashed: New Hope or False Dawn for Democracy?*,  
Oxford: Oxford University Press