CHAPTER 1

Introduction

]Rosalind Eyben and Irene Guijt

This chapter introduces the concerns that have guided the Big Push Forward and its culminating conference about the politics of evidence from which originated the case studies in the present book. The book's principal themes that emerge from the case studies are identified and the chapter concludes with a summary outline of the contributing chapters.

In June 2014, the board of the Global Partnership for Education (GPE) adopted a new funding model 'to have a positive leveraging effect on the development of national sector-wide policies, strategies and systems'.[1] Thirty per cent of the funds were to be used for ex-post payments against pre-determined results. The United Kingdom, Australia, and the United States had pushed for a greater percentage of the allocation to be through payment by results (PBR), but the majority of the board members were reluctant to risk a higher proportion of funding to a bold experiment they did not fully understand – and of which the pilot phase was cancelled. One of those objecting to PBR in principle (a civil society representative) emailed Rosalind Eyben:

> [EXT]I don't think the Board really understood the implications of this when it approved this new model – and [recipient] countries certainly were not aware of what this means in practice.[e/EXT]

Despite some discomfort about where the PBR model was taking them, most board members found it difficult to challenge the proposal. Everyone wants results! Eyben's informant, on the other hand, was keenly aware of how PBR risks instrumentalizing education, apart from its possible perverse effects (see Chapter 2). He feared that judging children's performance against measurable learning outcomes would become the sole yardstick for value, crowding out a transformational approach to education as a process of empowerment.

PBR is one of the mechanisms of a 'results and evidence agenda' that seeks to improve and manage development aid through protocols, procedures, and mechanisms for reporting, tracking, disbursing, appraising, and evaluating its effectiveness and impact. As detailed in Chapter 2, the agenda became influential in the first decade of the present century, with an increasing number of development practitioners concerned about what they saw to be the agenda's pernicious effects by around 2010. As in the GPE case, they worried that the results and evidence agenda undermined the potential for development aid to support transformational development: that is, that it leads to changes in power relations and structures that create and reproduce inequality, injustice, and the non-fulfilment of human rights. Concerned practitioners mentioned the time and money wasted in negotiating with funders over the utility and feasibility of imposed protocols and complained of the accountability pressure that forced the generation of 'sausage numbers' (Chapter 3), leaving limited time and energy for adaptive and responsive programming in support of complex change processes. One concerned senior official from a United Nations agency rang the authors of this chapter to explain how his agency had negotiated for several months with a government funder:

[EXT]They themselves knew it was ridiculous what they were asking for but they said it was political. In the end it comes down to money and for X millions of dollars we had to agree. But we rely on you academics working on the flanks to start a conversation about this.[e/EXT]

Where could these and the growing number of concerned voices discuss the trends, consequences, and options?

The initial impetus for what became the Big Push Forward (BPF)[2] was a conference in May 2010 in the Netherlands about evaluation in complex contexts (Guijt et al., 2011). Although it was borne out of a deep disquiet about how the term 'rigour' had been captured by advocates of certain impact approaches, most participants discussed methodological innovations and challenges with little reference to the shrinking political space within which the methods discussed could be used. Knowing that many practitioners were interested in the politics of how impact and results are defined and assessed, later that year the authors organized a meeting in Britain where more than 70 participants shared their experiences and reaction to the results and evidence agenda. Six months later (April 2011) we launched the BPF to make

the results and evidence agenda a legitimate subject for public debate. Our first blog laid out the challenge:

> [EXT]Hard evidence, rigorous data, conclusive proof, value for money, evidence-based policy are tantalizing terms promising clarity about what works and what should be funded in international development. Yet behind these terms lie definitional tussles, vested interests and contested worldviews. For those who hold the purse strings, certain ways of knowing and assessing impact are considered more legitimate than others. Yet increasingly people are recognizing the need for multiple and mixed methods and approaches to better understand complex change and that these compared to imposed standards are more likely to lead to fair assessments helping us learn how to support a fairer world.[3][e/EXT]

In the two years that followed, development practitioners commented on our blogs and debated the issues at meetings and workshops with us, including in Germany, the Netherlands, Nigeria, Sweden, Switzerland, the United States, and the United Kingdom. In Australia, for example, we contributed to a workshop on evaluation methods for staff of AusAid (now part of the Department of Foreign Affairs and Trade), organized by the Development Leadership Programme. Participants discussed the practical challenges staff faced in juggling methodological and political demands in the messiness of development programming in complex contexts (Roche and Kelly, 2012). In these forums, practitioners spoke of managing these demands in ways that would enable them to pursue approaches more aligned with transformational development. We found that the politically alert are subtly playing the game and changing the rules, with some seeing the results agenda as creating opportunities to support transformational development.

The BPF convenors (joined in mid-2011 by Chris Roche, Cathy Shutt, and Brendan Whitty) were discovering how diverse and context-specific the effects of the results and evidence agenda were. Findings from a BPF survey (Chapter 3), comments on our website and the 'From Poverty to Power' debate (Green, 2013) revealed that the agenda was not necessarily harmful; it could trigger improvements in design and learning, as well as in accountability. Yet, whether on balance the effects were judged as positive or negative for transformational development depended on the perspective of the judge. That, in turn, begged a political question: 'Whose perspective counts?' We realized that such a fundamental question, along with others about the values and ideology of the results and evidence agenda, risked being

ignored in a context of 'highly constrained resources, crazy time pressure, and the need to deliver some (any!) results to feed the MEL [monitoring, evaluation, learning] machine' (ibid.).

Furthermore, while accounts about the more negative power dynamics of the agenda were circulating over cups of coffee and glasses of beer, detailed case studies were rare. People were frightened of going public about their experiences of distortions and problems. They feared exposing international aid to an often-sceptical press, or being subjected to ridicule – or worse, putting their jobs or organizations at risk. The BPF convenors, therefore, decided to create an opportunity for practitioners and researchers to reflect critically on these issues in a safe space through a conference held in April 2013. We organized it so as to protect participants' identities, with a strict insistence on 'Chatham House rules'.[4] Even so, some participants were reluctant to present case studies (even verbally) and others unwilling to go public in this book. Power was influencing people to opt out of sharing their experience of the reality of the situation. Even the contributors here have exercised care in their decisions about what to include and what to leave out of their chapters – and how to portray choices about results and evidence.

To our chagrin, budget constraints prevented the financing of participants' travel and other costs, limiting the range of perspectives present at the conference.[5] Apart from a video link with Ola Abu Alghaib in Ramallah, experiences of grantee organizations in aid-recipient countries were largely absent (as they are in this book). Participants were either development agency staff (bilateral, multilateral, and international non-governmental organization (NGO)) or managers or technical advisers in projects these agencies funded. The others were consultants and researchers, as were the BPF convenors.

The conference was organized around participants' own case studies of the artefacts – the processes, mechanisms, and tools employed by them and their organizations to assess results and to generate and use evidence (Chapter 2). This book contains a selection of these cases, as well as an assessment of the organizational and power dynamics associated with their use. Conference participants examined whether and how one can reconcile messy, unpredictable, and risky pathways of societal transformation with bureaucracy-driven protocols. Participants stressed the importance of conscious engagement in the politics of evidence (Chapter 11), providing positive examples of how they either successfully resisted the unhelpful demands

of the results and evidence agenda or used these to good effect to challenge myths, promote methodological advances, and gain theoretical insights.

We chose the conference title 'The Politics of Evidence' from Denzin and Giardina's (2008) volume about why and how academic institutions and policy-making bodies regard evidence from qualitative research as being less robust and rigorous than quantitative research. We broadened the 'politics of evidence' term to encompass both how power works to define what is or is not accepted as robust evidence or determined as a result (in both programming and evaluation), and the resistance and contestation that such power generates.

Power and the politics of evidence

We analyse power both as an asset that individuals and organizations control and also as a process to which we are all subject.[6] When conceptualizing 'power' as an asset, the politics concerns who controls the definition of a result or of evidence and what is acceptable to whom (Morse, 2008). This kind of power operates through the formal institutional arrangements for policy-making and implementation, such as the civil service, the legislature, and local government; certain policy actors, such as ministers and parliamentarians, are visibly powerful. Janet Vähämäki (Chapter 8) uses this perspective on power to examine 30 years of recurrent struggle over results-based management between Swedish ministers and the staff of the Swedish International Development Agency (Sida). Power as an asset can also be less formal. Private individuals and organizations, such as philanthropic foundations or academic think tanks, influence meaning and value in development. The financial or intellectual capital they deploy gives them legitimacy in articulating and promoting (or interrogating) the discourse. In 2012, the BPF (itself exercising some modest power in this manner) undertook a preliminary power analysis of the results and evidence agenda from this perspective, as summarized in Chapter 2.

Power as an asset can be related to the idea of 'power over', when A (with more power) controls how B (with less power) behaves. The international aid system is often portrayed as a power chain – official aid agencies at the top, international NGOs and recipient governments in the middle, and local community organizations at the bottom. Yet many rules and procedures cannot be clearly attributed to specific decisions made by identifiable actors. Much of the results and evidence agenda is an effect of more diffuse power dynamics. Even when someone 'in charge' makes a decision, they may be unaware of – or indeed unable to

prevent – what actually happens, since this is an outcome of multiple interactions by myriads of interconnected actors. Those in authority – chief executives and government ministers – are as much subjects of power as the most junior staff member; and multilateral aid agencies as much as grass-roots organizations.

This kind of invisible power is at work in all our relationships – each time we walk into a room, make a suggestion, or participate in a workshop. It is the process of socializing and embedding that shapes what we think, say, and do. An example from the present book is how societal norms render children marginalized and their voices disregarded (Johnson, Chapter 9). Participatory and visual evaluation methodologies designed to elicit children's perspectives and knowledge fail to influence adult decisions, unless there are complementary efforts to change the social norms. Because it shapes habitual patterns of relations, this kind of power is rarely noticed – and therefore remains unchallenged. 'Power relations were never discussed during the term of the project: nobody brought it up,' observe Causemann and Gohl (Chapter 10). Only by becoming aware of such power can anyone seek to change inequities in the relationship, an act of conscious agency.

Both of these approaches to power are drawn upon in *Power and Organizations*, which identifies 'significant and subtle power practices' (Clegg et al., 2006: 176) at work in institutional life. We find six approaches particularly relevant to the international development sector. Awareness of these can help people understand the nature of their experience with the results and evidence agenda:

- locking members inside and keeping outsiders out and systematically misrepresenting other realities;

- the division of labour into complex chains of power, enabling those at the top to maintain a distance from the effects of power;

- staff staying obedient through a ceaseless round of activities with little time for reflection;

- delegation to intermediaries, obliging them to implement decisions that have been made higher up the system;

- making those who are the subjects of power complicit in its exercise;

- applying instrumental and value-free science.

Reference is made to these power practices in the thematic analysis that follows.

[A]Principal themes

The case studies in this book take up two substantive issues of concern to those interested in transformational development, namely the impact of the evidence and results agenda on the ability to pursue rights-based approaches and whether the growing emphasis on upward accountability is trumping mutual learning. Perspectives on these issues are informed by personal values and professional formation, and influenced by the position – or positionality – of individuals in an organization or of an organization within the aid system. These interact to affect people's judgement as to whether, on balance, the effects of the results and evidence agenda are positive or negative and to influence how people respond to the changes induced by that agenda. The dynamics of perspective and of push and response are thus examined first, followed by a discussion of how these play out with respect to the substantive issues of rights/results and learning/accountability.

[B]Position and perspective

We launched the BPF to 'push back' against the bureaucracy-driven protocols we believed were shrinking the possibilities of development organizations supporting messy, unpredictable, and risky pathways of societal transformation. Not everyone shared our initial views! The survey of practitioners conducted prior to the conference (Whitty, Chapter 3) suggests that whether the agenda's effects are experienced as largely positive or negative depends on a person's organizational role and on their organization's location in the aid nexus. BPF convenors fell into that category of development professionals, mainly academics and consultants, with the most negative views of the results and evidence agenda. In contrast, people with M&E responsibilities had more positive experiences than, for example, programme officers, and senior staff were more positive than technical advisers or those in middle management. The importance of position and perspective is also evident in Chapter 5, where Chris Roche looks at the arguments between managers responsible for finance and fundraising, who want standardized, organization-wide performance measures, and those in programming and advocacy, who seek a variety of forms of assessment and reporting depending on the context and activity.

However, the survey was far from comprehensive. Very few recipient governments and national NGOs responded, with even fewer responses from smaller civil society organizations in the South. Abu Alghaib's experience in leading a network of disabled people's organizations (DPOs) and her research findings into the experience of other DPOs in the region (Chapter 7) indicate that such organizations have been impacted negatively by the results agenda. Her case study of what happened to her own organization after winning a grant from a multilateral agency is a devastating account of the impunity and arrogance of a bureaucratic machine that 'makes those who are subjects of power complicit in its exercise' (Clegg et al., 2006: 179). In the end, writes Abu Alghaib, 'the organization became a stranger in its own project'.

Although Abu Alghaib had no access to the internal workings of the bureaucratic machine that had triggered such effects, other chapters written from an inside perspective show such bureaucracy to be more heterogeneous than might appear when viewed from the bottom end of the chain. Vähämäki's account of the evolution of the results agenda in Sida tells how staff were aware of the possible negative effects of what they were asking from grantees. In the 1970s, Sida field offices were already complaining: 'Talks and observations on how results valuation affect the Kenyans has given me goose bumps' (Chapter 8).

As Roche discusses (Chapter 5), it is difficult to imagine what the world might look like if one were positioned differently. Causemann and Gohl (Chapter 10) observe that Southern NGO staff had little understanding of the realities and needs of the German NGOs, intermediaries in the funding chain:

> [EXT]They had hardly ever travelled to Germany to meet NGO staff there, and as such did not have an opportunity to experience the realities of a donor desk officer … Donor desk officers told us they lacked the time for such explanation, but we did not see that they recognized the need, not appreciating how hard it was for Southern NGOs to put donor requirements in context. [e/EXT]

Two practices of power are at work in the above example: maintaining a distance between the designated exercisers and subjects of power, and ensuring NGO obedience to power through 'a ceaseless round of activity with little room for reflection' (Clegg et al., 2006: 178).

Another subtle power practice is to delegate to intermediaries decisions that have been made higher up the system (ibid.: 177). Intermediary organizations, for example a UN agency or a

Northern development NGO, as well as individuals in middle-management positions, can have a 'squeezed middle' experience (Chapter 2) when they are aware of the potential negative effects of what is asked of them. These individuals are responsible for reporting results while at the same time trying to protect the front line from the perceived excessive demands of those higher up the chain. Van Es and Guijt (Chapter 6) discuss how the programme officers in a Dutch NGO, Hivos, were 'balancing between insisting on justified requirements for accountability while maintaining a flexible attitude to enable responsive programming by partners'. Intermediaries are potential brokers because their position provides them with more than one perspective; this enables them to adopt creative responses to this particular power practice, and these are examined in this book's conclusion (Chapter 11.)

[B]The dynamics of pressure and response

Some of the chapters usefully take a historical perspective on pressure from and reactions to the results and evidence agenda, illustrating the dynamics of thinking and practice. Vähämäki analyses her case with reference to the literature on management reform cycles in public sector bodies that highlight politicians' desire to control by institutionalizing new management technologies. These technologies inevitably fail to meet expectations, yet subsequent governments ignore history and try again. Her case study – an account of top-down pressure and of staff resistance – appears to support the theory that every 'failure allows the surge of new optimism for yet another round of reforms' (Chapter 8). Eyben (Chapter 2) notes how 'performance measurement systems' operate in a near-continual cycle of reform, but rather than history just repeating itself, each failure results in the introduction of more stringent and controlling practices, echoing Vähämäki's account of the most recent phase in Sida. Vähämäki cites senior managers' belief that this time – at last – the reforms might trigger their intended effects. One reason she cites for this situation is that long-serving staff have retired – or, as Eyben discovered when participating in a BPF-initiated public meeting in Stockholm, they have resigned because of their unhappiness with the new organizational culture of compliance and silence (driven by the results agenda) – to be replaced by new staff who share management's vision of a results-oriented Sida.

Describing the experience of Hivos, van Es and Guijt note that: 'Although the need for more systematic and better quality result information of partners for organizational learning was internally felt and discussed … there was little external pressure to change existing M&E

practices' (Chapter 6). But from the early 2000s, growing internal and external pressures triggered a creative process of improving the organization's and its partners' M&E capacities in ways that sought to stay faithful to Hivos' transformational objectives and partnership values. Some partners were critical of the new request for indicators, despite Hivos giving them the freedom to choose, asking 'why social change investments and agencies are being so closely scrutinized for effectiveness and impact, when private sector actors are not similarly held to account' (van Es and Guijt, Chapter 6).

Causemann and Gohl tell a similar story of how in the early 2000s a network of German NGOs came under increasing pressure from the German government and the European Union to improve their reporting of results. Moreover, 'NGOs themselves, many of whom financed themselves by public donations or by foundation funds rather than through government funding, also believed that the public expected enhanced accountability on the outcomes and impacts achieved with a results-based mindset' (Chapter 10). A group of NGOs took the initiative in designing an approach 'before the government imposed more restrictive requirements'.

In contrast to these historical analyses that take us up to the present day, Cathy Shutt focuses on the current dynamics of push and response in the UK. She examines how a network of individuals from foundations and development NGOs – as well as the Department for International Development (DFID) itself – responded to DFID's emphasis on value for money (VfM) from 2010 onwards by simultaneously adapting their practice to respond to the VfM discourse while keeping open the space for continued DFID funding of transformational development that many believed was under threat (Chapter 4). Ex post VfM analyses provided grant-makers with opportunities for learning:

> [EXT]The main reason an outstanding programme that began with a few women getting together to reflect on their lives led to a successful political and social movement anticipated to provide economic empowerment for several thousands living in their community was the extreme flexibility the grant-maker allowed the women in their use of funds. [e/EXT]

This is in striking contrast to the stringent requirements placed on the women's DPO that Abu Alghaib describes in an account of the organization's response to the promotion of the results agenda (Chapter 7).

As further discussed in the book's conclusion, collectively the case studies reveal a wide variance in organizational responses to the demands of the results and evidence agenda. We now turn to discuss how this diversity of positionality and organizational dynamics with respect to the results and evidence agenda plays out with respect to the challenge of pursuing rights-based approaches, and when and how upward accountability trumps organizational learning.

[B]Rights and results

Rights-based, transformational approaches (where 'how' matters as much as 'what') and transactional results-based management (focused on the 'what') may be uncomfortable bedfellows (Hulme, 2010). The number of agencies, particularly international NGOs, using rights language has continued to increase, but the contradiction between rights-based approaches - and their political and process approach to intangible goals such as empowerment - and the increasing popularity of results-based management has become very apparent. It is harder to manage support for transformational approaches when one is required to report tangible, easy-to-measure changes, as discussed, for example, by van Es and Guijt (Chapter 6) with respect to the experience of Hivos, a Dutch NGO committed to transformational development.

The BPF convenors, such as Hivos, understand transformational development as involving processes of political change that secure a greater realization of rights and social justice. We understand poverty as a product of structured relationships of unequal power, and, as David Mosse puts it, 'the consequence of normal economic and political relations' (2007: 6). Hence, from this perspective, transformation is a process that changes the status quo. However, because the international development sector is part of 'normal economic and social relations', when it seeks to support transformational development it encounters fundamental contradictions concerning its legitimacy of action, its practice of power, and its lines of accountability, all of which are recurrent themes in the politics of evidence and results, as explored in this volume's chapters. It is only by addressing these contradictions that the sector has the potential to make a modest contribution to processes of transformational change. A transformational approach is thus about both means – how and with whom the sector behaves – and values. In Table 4.1 in Chapter 4, Cathy Shutt contrasts aid as transactional (value-free, technical, focusing on the 'what') with aid that supports transformational change. In the same chapter (Table 4.2), she demonstrates how a managerial

concept such as the practice of VfM can be adapted to serve transformational change, provided its users are conscious of their values and of the power relations at play. And in Chapter 5, Chris Roche explores how the meaning and practice of accountability – another central theme in current international development practice – change when associated with a transformational development approach.

We equate a transformational with a rights-based approach, and we understand this approach as providing a normative framework (human rights) that not only guides development practice but also – and importantly – is one in which:

> [EXT]People are placed at the centre of development processes, no longer seen as beneficiaries of development projects with needs, but as active citizens with rights and entitlements. As a result, aid can be seen as contributing to the transformation of state–society relations [that empower] all citizens to claim their rights … Politics and power relations are thus put at the centre of programming analysis and interventions rather than seen as negative 'risk' factors attached to projects. (Piron, 2005: 22–3) [e/EXT]

Not everyone at the Politics of Evidence conference agreed with us. In conference evaluation feedback, one participant disliked 'transformational development' because it was highly normative and made assumptions about what aid could or could not achieve. Rather, the participant wrote, the purpose of aid is to support local development 'which may or may not be transformative depending on the context and need'. We believe, however, that when aid is indifferent to the power relations of the process, is regarded as objectively technical (non-normative), and concentrates on outcomes, then by default it risks contributing to the maintenance of inequitable power relations that prevent greater social justice. For this reason, we fear that the results agenda may have a negative effect when framed in terms of efficiency and effectiveness – a power practice that applies the 'instrumental and value-free science' (Clegg et al., 2006: 177) that is common to transactional approaches that treat rights as a measurable project outcome (Ackerman, 2005).

In Chapter 2, Eyben discusses how, in 1997, the New Labour government in the UK introduced the discourse of technical, value-free 'evidence' to escape accusations that its policies were driven by old-fashioned socialist ideology. In Sweden (Chapter 8), results-based approaches were promoted by both centre-right and centre-left governments.

Historically, the results and evidence agenda can be seen as part of the 'new public management' paradigm, modelled on corporate sector practices designed to maximize shareholder profit and eschewing any explicit ideological commitment (which, of course, does not preclude such practices being informed by tacit values and ideology).

Even a development agency still subscribing to a normative human rights framework may prejudice transformational development if it focuses solely on outcomes and ignores process. Abu Alghaib's chapter (Chapter 7) provides such a case of how the funder, focused on the delivery of rights outcomes, undermined the grantee's rights-based approach of recognizing and respecting the situation of marginalized groups and providing space and time for them to gradually build their self-confidence as a precursor to claiming their rights. By assuming and requiring that the small and doubly marginalized network of women's DPOs should be capable of organizing themselves into national advocacy networks within a two-year funding period, the donor agency was adopting a transactional results-based approach despite framing its support in terms of realizing rights. Similarly, Oxfam Australia (Chapter 5) risked undermining its commitment to a rights-based approach through its head office's desire for transformational change in impossibly short time frames.

Johnson (Chapter 9) also shows how talking about 'rights' does not necessarily imply a rights-based approach. In revisiting projects she had evaluated some 10 years earlier, she found that, while everyone agreed on the importance of children's rights, participatory evaluation approaches enabling children's rights to be heard were under threat. NGO staff were under pressure to deliver a certain kind of evidence they described as 'hard', in contrast to the 'soft' children's evidence from participatory visual methods. Vähämäki (Chapter 8) noticed that even the mere reference to transformational development (let alone support for transformative processes) was disappearing. Sida staff feel that the increasing emphasis on results and reporting systems has made them lose sight of the agency's purpose. 'Results, transparency, and accountability' have become the primary reform objectives, prioritized above content themes such as gender and poverty reduction'.

Shutt, on the other hand, describes how some British development NGOs are managing to square this circle. These NGOs initially believed that DFID's VfM requirements threatened transformational development projects:

[EXT]And, even though there was no evidence to imply long-term social change programmes were receiving less money, some working on such initiatives were frustrated that the focus on quantitative metrics denied adequate discussion of qualitative dimensions of value.[e/EXT]

Yet, enabled by their relationships with 'politically savvy DFID staff equally frustrated with DFID's focus on economy and efficiency', over time the NGOs are succeeding in interpreting VfM in ways consistent with transformational development. Nevertheless, only some grantees have succeeded in playing the game to change the rules. The particularly difficult challenge for those engaged in transformational development is to avoid diluting their efforts and ideas of what counts as evidence of change.

[B]Learning and accountability

While there is no doubt that the results and evidence agenda is a veritable battlefield between 'accountability' and 'learning', the issue is more nuanced than the stereotypical frustration about upward accountability always trumping learning (Wallace and Porter, 2013).

As a positive, accountability pressures can stimulate deep organizational discussions on core values and what success looks like, along with heightened appreciation of the contribution of M&E as an organizational practice. Half of the 153 respondents in the BPF survey reported that the push for results had led to some or considerable improvement in their organization's ability to learn (Chapter 3). The head of Oxfam Great Britain's M&E unit commented that 'five years ago [we] struggled to get the ear of senior managers (let alone ministers). But the results agenda has increased the stakes … encouraging organizations not only to increase investment, but also to listen to the findings coming from our own data gathering and analysis'.[7] The Hivos experience echoes this (Chapter 6). Until external pressures accelerated the search for improvements, Hivos' existing M&E practice with partners contributed little to their learning about social change, knowledge sharing, or building new theory and practice. In response to growing demands, the M&E unit at Hivos developed a theory of change (ToC) approach that encourages partners to be explicit, self-critical, and adaptive. In the UK, the VfM debate has also stimulated internal debate and learning around what is valued and how to represent that (Chapter 3).

Perhaps the clue lies in how accountability is defined (Guijt, 2010). If accountability is relational – being answerable to others within a relationship of power (Goetz and Jenkins,

2005), it requires clarity about 'who has the power to call for an account and who is obligated to give an explanation for their actions' (Newell and Bellour, 2002: 2). For transformational development, in which the 'how' is as important as the 'what', being accountable to oneself is essential to understand what one has undertaken, to explain strategic decisions and the ToC, and, of course, to show how money was spent. This form of 'strategic accountability' seeks to answer the question: 'Did we act as effectively as possible with the means we had?' In this sense, accountability is intrinsically about commitment to one's ideas and strategies (Fry, 1995).

Roche (Chapter 5) recounts how the accountability debate opened up space in Oxfam Australia to explore, in particular, more bottom-up processes of feedback. There has been a lot of experimentation in recent times (Guijt, 2014); however, these experiments have often lacked adequate political or power analysis in their design (Fox, 2014). Several chapters allude to this perspective on accountability in which learning plays a critical role in ensuring that organizations are not mere 'delivery agents' but rather 'drivers of change' (Reeler, 2007). Van Es and Guijt explain how Hivos pursued ToC as one learning mechanism to strengthen the strategic accountability of partners and Hivos alike. Shutt cites a similar example using a VfM perspective: 'The HIV/Aids Alliance, one of the earliest adopters, found that comparisons across their portfolio raised questions about sustainability that they could integrate in future planning.' Johnson refers to the way in which deep listening to children can remind decision-makers about what they need to be accountable for – and to whom (Chapter 9).

But tensions remain. And there is a constant need to ask who is learning and for what. Being held accountable means having 'respond-ability', according to Ebrahim (2005). This becomes impossible when organizations are viewed by funders as 'merely … contractors for implementing the ideas of others' (Abu Alghaib, Chapter 7). Interestingly, even when couched in terms of 'more critical thinking to get better results', some Hivos partner organizations still perceived ToC as 'a donor-driven invention' for upward accountability. This was aggravated in the first years of ToC use, as Southern partners of Hivos noted that while *they* were required to use a ToC to enhance reflective learning, Hivos itself had not done so. Abu Alghaib's survey of 25 DPOs found that only five of these Southern NGOs considered that M&E approaches helped promote organizational learning. And although German NGOs valued regular reporting as an indication that 'something meaningful was

happening on the ground' (Chapter 10), reports 'did not contribute to anyone's learning or conceptual development with respect to impact assessment', as reports were not read and were written only to meet accountability requirements. Rigid protocols for reporting back to donors are choking the space to learn and adapt, aggravated by budget pressures.

Consultants hired by Sida observed that the new reporting requirements had not enabled Sida staff to think more 'but rather less' (Chapter 8). Eventually, the push for upward accountability – which generated specific results matrices – created such dire difficulties for those grantees who had to provide the data that reporting became solely the responsibility of Sida staff, a death blow to the original intention of mutual learning. Currently, grantees are not involved – upward accountability trumped learning in this case. The drive is more 'to appear transparent and communicate results than to actually work for results in practice', a Sida director told Vähämäki. A possible irony of the drive for accountability is that, in detaching itself from any reality other than Swedish politics, Sida may have secured what some of its staff have been seeking for the last 30 years: that is, allowing its grantees considerable flexibility to carry out their work free from interference.

Overall, while the results agenda in theory opens up space for bottom-up, people-centred accountability (Melamed, 2011), the narratives in this book are not optimistic about the possibility of leveraging accountability for learning. The particular interpretation of accountability as an upward, financial 'holding to account' on promises obstructs learning but also fails in its raison d'être of communicating to citizens and politicians what has been achieved with taxpayers' money. We need to understand better what makes it possible for what Whitty notes is the potential of accountability: 'encouraging a more disciplined reflection on programming that will achieve results, strengthen knowledge management, curb the tendency to "over claim" and improve broader learning processes'.

[A]Structure of the book

Earlier versions of Chapters 2 and 3 were written as framing papers for discussion at the conference. Updated with insights from the conference, they provide crucial context for the nuanced experiences of the results agenda discussed in the case study chapters, which analyse the politics of results and evidence in a variety of organizational contexts and sets of relationships. Two of the cases included in the book are from the UK, with one each from Australia, Germany, the Netherlands, Palestine, and Sweden. They illuminate different sets of

relationships in the aid chain, including those between a network of grass-roots organizations and a multilateral agency, within a government agency, between a government agency and its own country's development NGOs, and within and between Northern NGOs and their Southern partners.

Following this introduction, Chapter 2 aims to alert the reader to the way in which power informs the apparently technical and value-free language of efficiency and effectiveness. Rosalind Eyben starts with the tools and protocols – the 'artefacts' – that are the concrete expression of the results and evidence discourses and analyses their diffuse power effects. 'Through a mutually constituted process of production, consumption, and resistance, artefacts evolve and mutate, beyond the control of any individual or organization.' She then traces the interlinked genealogies of the discourses of 'results' and 'evidence' in public sector management and looks at how and why they have become so influential in the international development sector. Armed with these insights, practitioner readers should be better placed to analyse and confront power that works invisibly in order to normalize inappropriate means of designing and assessing programmes with multiple pathways of change and a transformational development objective. At the same time, any such power analysis requires reflexivity, a self-awareness to avoid disempowering others.

The other framing chapter (Chapter 3) is Brendan Whitty's analysis of the findings from the crowdsourcing survey. At the Politics of Evidence conference, he brought home the extent to which opinions differ on the pros and cons of the results agenda. His chapter offers insights from over 100 experiences of people's responses to results artefacts, noting how respondents' positive or negative experiences are shaped by their positionality (as discussed above) and – as part of that positionality – by their specific relationship to the production and use of numbers: people in senior management were most positive about how simplification and reporting against common formats provide a useful big picture. Respondents' perspectives are also shaped by their experience of how results-oriented reforms have been implemented in their own part of the sector and by their own organization's capacities for and history of monitoring, evaluation, and reporting. Whitty provides an organizational typology that can particularly help funders engage more productively and sensitively with grantees with low administrative capacity to translate their 'nuanced contexts and programming to standardized administrative formats'. He points out the grave error of funders supporting only those who

are good at showing what they can do. Failure to demonstrate this because of limited administrative capacity 'does not necessarily indicate poor work'.

The next three chapters deal with key concepts: VfM, accountability, and ToCs. In Chapter 4, Cathy Shutt makes use of her recent experience in advising variously positioned British development organizations in order to address readers who want to engage politically with the increasingly dominant VfM concept. A useful table summarizes the differences between transformational and transactional development as these relate to VfM, and the chapter helps the reader further by unpacking the different technical elements of assessing VfM. Shutt then discusses the tactics she has observed being deployed to reduce the negative effects of VfM and identifies the possibilities for making the concept more supportive of transformational development. She argues that, positively, VfM can offer power-aware practitioners opportunities to increase accountability to citizens and to each other, and it also has the potential to influence links between VfM and what is valued in development.

In Chapter 5, Chris Roche critiques the upward accountability that is informed by principal-agent theory (as discussed in Chapter 2). Drawing on his experience of working for Oxfam Australia, he describes how Oxfam attempted to adopt an alternative approach that recognized multiple accountabilities and had a particular emphasis on 'bottom-up' or downward accountability. Describing the findings of an independent review into Oxfam's implementation of this alternative approach, Roche looks at how staff positionality has influenced diverse perceptions of how well the organization has been doing, thus making it difficult to secure organization-wide agreement for a single performance framework. To achieve this, we suggest, the organization would need to engage in a strategic conversation about accountability – a conversation that explicitly took account of internal power relations, as well as relations between Oxfam Australia and those with whom it interacts. And this is when the Australian government is ratcheting up the results agenda, cutting aid budgets drastically, and reducing development NGOs' room for manoeuvre. Roche suggests that a greater focus on understanding and sharing the practice of effective relational accountability and its benefits might be an important way forward.

Chapter 6 is also written from a practitioner perspective, and from inside another development NGO. Marjan van Es and Irene Guijt provide a historical account of how, rather than letting itself be pushed around by its principal funder (the Dutch government), Hivos has sought to stay faithful to its transformative development vision by taking the initiative. It

responded to the results agenda by elaborating its own approach to ToC as a means of strengthening its strategic accountability: 'Although compliance drained time and energy from staff and further reduced the M&E unit's space to shape processes for Hivos' own results information and learning needs, the M&E unit persisted with its investment in strategic, critical thinking.' Hivos refrained from using rigid ToC protocols, instead issuing a clear policy statement. Yet staff have now asked for guidelines to ensure a high-quality ToC that, for Hivos, includes political analysis. We wonder whether this raises an interesting contradiction for agencies such as Hivos. Should critical reflexive practice for social change be mandatory? What is the risk of a thought process becoming a mechanical artefact?

Appreciating that large international development agencies are internal political battlefields is a headline in our reading of Chapter 7. Ola Abu Alghaib describes what happened to a grass-roots disabled women's network after they succeeded in obtaining funding from a multilateral development organization, 'Women Advance' (a pseudonym). Like Rosario León, cited in Chapter 2, she and her colleagues found themselves obliged to revise and revise again results matrices that bore no relationship to what they were trying to do. 'In effect, staff felt that it had become a project designed by consultants based on a methodology prescribed by Women Advance's staff.' Abu Alghaib tracks her own learning journey from this encounter, including her research into the experience of other DPOs in the Middle East and North Africa. Echoing Whitty's point about effective organizations with low administrative capacity to manage the results agenda, she concludes that such organizations should collectively resist the donor pressure that makes them strangers to their own projects.

In the next chapter, another insider (turned researcher), this time in a government agency, Sida, uses a longer time frame to chronicle the organizational dynamics of push and response. In Chapter 8, Janet Vähämäki reveals a diversity of power relationships and perspectives over a period of 40 years with ministers, senior management, and other staff promoting and resisting the results agenda. For readers wanting to influence a bilateral agency, the chapter demonstrates the importance of undertaking an organizational power analysis, bearing in mind the fact that it is only too easy to assume that such agencies are homogeneous monoliths. In this amazing account, Vähämäki shows us something very different.

While Chapter 7 emphasizes the UN Convention of the Rights of Persons with Disabilities and how donors subscribing to the convention are in practice ignoring the DPOs' principle of 'nothing about us without us', Chapter 9 is about another convention – the Rights of the

Child – and looks at how that convention's focus on children's participation is often neglected in programme evaluations. Vicky Johnson analyses how managers and programme leaders ignored children's evidence she had compiled through participatory visual methods and describes how she negotiated the politics of evidence, her tactics, and the resistance she generated. She stresses the need to include people with decision-making power over strategies and budgets in participatory processes to ensure that children's voices are heard. She offers her experience-based framework for others using similar evaluation approaches to elevate marginalized voices.

The final case study concerns a key power practice of the results agenda: project reporting. In Chapter 10, Bernward Causemann and Eberhard Gohl reflect on their experience as 'squeezed middle' project managers employed in a multi-country initiative run by German NGOs and their Southern partners. The co-authors analyse why their Southern partners resisted submitting regular written reports to the German NGOs. Bearing in mind that reports rarely enhance the learning of either the grantee or the funder (the latter often not even reading them), they look at possible strategies for easier and improved reporting, as well as alternatives to reporting. Causemann and Gohl argue that, rather than using formal written reports, it is better to undertake critical analysis and reflection in a process of dialogue.

Lastly, in Chapter 11, Irene Guijt takes a forward-looking perspective, recognizing that the results and evidence agenda is here to stay for some time. Drawing on the book's case studies and the BPF conference report, she looks at how those working in international development can deal with the politics of evidence, in ways that do not compromise the space for supporting transformational development.

[A]References

Ackerman, J.M. (2005) *Human Rights and Social Accountability*, Social Development Paper 86, Washington: World Bank <http://portals.wi.wur.nl/files/docs/gouvernance/HumanRightsandSocial0AccountabilityFINAL.pdf>.

Clegg, S. and Haugaard, M. (2009) *Handbook of Power*, London: Sage.

Clegg, S., Courpasson, D. and Phillips, N. (2006) *Power and Organizations*, London: Sage.

Denzin, N.K. and Giardina, M.D. (eds) (2008) *Qualitative Inquiry and the Politics of Evidence*, Walnut Creek: Left Coast Press.

Ebrahim, A. (2005) 'Accountability myopia: losing sight of organizational learning', *Nonprofit and Voluntary Sector Quarterly* 34(1): 56–87.

Fox, J. (2014) 'Social accountability: what does the evidence really say?', keynote speech at the Global Partners Forum of the Global Partnership for Social Accountability, 14–15 May 2014.

Fry, R.E. (1995) 'Accountability in organizational life: problem or opportunity for non-profits?', *Non-profit Management and Leadership* 6: 181–95.

Goetz, A.M. and Jenkins, R. (2005) *Reinventing Accountability: Making Democracy Work for Human Development*, Basingstoke: Palgrave Macmillan.

Green, D. (2013) 'So what do I take away from the great evidence debate? Final thoughts (for now)', in From Poverty to Power [blog] <http://oxfamblogs.org/fp2p/so-what-do-i-take-away-from-the-great-evidence-debate-final-thoughts-for-now/>.

Guijt, I. (2010) 'Exploding the myth of incompatibility between accountability and learning', in J. Ubels, N.-A. Acquaye-Baddoo and A. Fowler (eds), *Capacity Development in Practice*, London and Washington DC: Earthscan.

Guijt, I. (2014) 'Feedback loops – new buzzword, old practice?' in BetterEvaluation [blog] <http://betterevaluation.org/blog/feedback_loops_new_buzzword_old_practice>.

Guijt, I., Brouwers, J., Kusters, C., Prins, E. and Zeynalova, B. (2011) *Evaluation Revisited: Improving the Quality of Evaluative Practice by Embracing Complexity, Conference Report*, Wageningen: Centre for Development Innovation, Wageningen University and Research Centre <http://edepot.wur.nl/169284>.

Hulme, D. (2010) 'Lessons from the making of the MDGS: human development meets results-based management in an unfair world', *IDS Bulletin* 41(1): 15–25.

Melamed, C. (2011) 'If not results, then what? The risks of not having a results agenda' in From Poverty to Power [blog] <http://oxfamblogs.org/fp2p/if-not-results-then-what-the-risks-of-not-having-a-results-agenda/>.

Morse, J.M. (2006) 'The politics of evidence', *Qualitative Health Research* 16(3): 395–404.

Mosse, D. (2007) *Power and the Durability of Poverty: A Critical Exploration of the Links Between Culture, Marginality and Chronic Poverty*. CPRC Working Paper 107, Manchester: Chronic Poverty Research Centre.

Newell, P. and Bellour, S. (2002) *Mapping Accountability: Origins, Contexts and Implications for Development*, IDS Working Paper 168, Brighton: Institute of Development Studies.

Piron, L.H. (2005) 'Rights-based approaches and bilateral aid agencies: more than a metaphor?', *IDS Bulletin* 36(1): 19–30.

Reeler, D. (2007) 'A three-fold theory of social change and implications for practice, planning, monitoring and evaluation', Cape Town: Community Development Resource Association <http://www.cdra.org.za/threefold-theory-of-social-change.html>.

Roche, C. and Kelly, L. (2012) *The Evaluation of Politics and the Politics of Evaluation*, DLP Background Paper 11, Birmingham: Developmental Leadership Program, University of Birmingham <http://www.gsdrc.org/go/display&type=Document&id=4340>.

Wallace, T. and Porter, F. with Ralph Bowman (eds) (2013) *Aid, NGOs and the Realities of Women's Lives: A Perfect Storm*, Rugby: Practical Action.

---

[1] GPE, 'Report of the country grants and performance committee part 1: operational framework for requirements and incentives in the funding model of the global partnership for education and results- based financing pilot' (BOD/2014/05—DOC 03, May 2014), p. 4.

[2] <http://bigpushforward.net/archives/502>.

[3] Chatham House rules require participants to protect the anonymity of those contributing to discussions by not disclosing to a wider public which particular person said what.

[4] The difficult process of raising finance was an intriguing illustration of the politics of the results and evidence agenda. Was it because most funders viewed the debate that we had triggered as regressive and unhelpful? Or was it because they feared being associated with controversy and how this might affect their standing with DFID/AusAid or others?

[5] We deliberately avoid an over-academic discussion of concepts of power. Readers interested in going deeper into the complexity of the concept are recommended Clegg and Haugaard (2009).