

UNCOVERING THE POLITICS OF 'EVIDENCE' AND 'RESULTS'. A FRAMING PAPER FOR DEVELOPMENT PRACTITIONERS

Rosalind Eyben

April 2013

www.bigpushforward.net

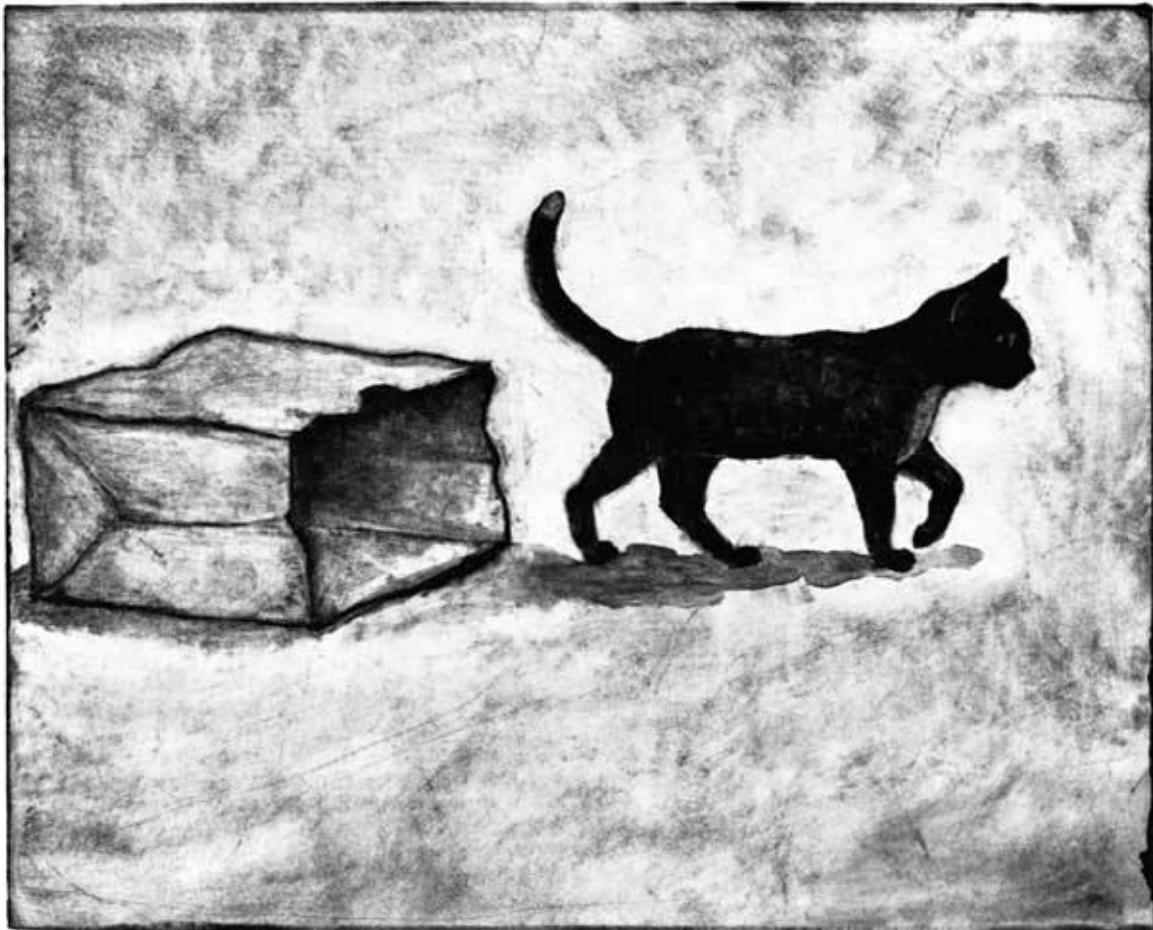


Image © Ben Killen Rosenberg / Getty Images

CONTENTS

CONTENTS.....	2
SUMMARY	3
1. The Readership	4
2. Aims and Approach	4
3. Results-and-Evidence Artefacts.....	6
3.1 The disciplining effect of artefacts.....	8
3.2. The rise and fall of artefacts.....	9
3.3. The effect on transformative agendas	10
4. Approaches Based on Results and Evidence	11
4.2 Results-based approaches	11
4.2 Evidence-Based Approaches.....	17
5. The Drivers of the Results-and-Evidence Agenda - and the Countervailing Trends.....	22
5.1 The need to be seen to be in control.....	22
5.2 Value for money and the politics of accountability	23
5.3 The sector’s internal dynamics	25
6. Conclusion	26
Endnotes.....	27

SUMMARY

Hard evidence, rigorous data, tangible results, value for money – all are tantalising terms promising clarity for the international development sector. Yet, behind these terms lie definitional tussles, vested interests and contested world views that this background paper to the Politics of Evidence Conference aims to make explicit and question. The aim is to encourage development practitioners to strategize in expanding the politico-bureaucratic space to make room for flexible and creative support of locally-generated and transformative change.

Disentangling the historical threads and origins of results-based management and evidence-based policy/programming has discovered a strong ‘family resemblance’. The discourses share a common epistemology or history of ideas and concepts. Both assume that evidence pertains only to verifiable and measurable facts and that other types of knowledge have no value: both a particular understanding of causality, efficiency and accountability. How and why have these discourses influenced the development sector and who is promoting them in which contexts? What has been the effect on the sector’s priorities and practices, and particularly its capacity to support transformative development?

Arguing the importance of being critically aware of how power sustains and reinforces the development sector’s results-and-evidence discourses, the paper explains how the resulting tools and methods, such as logical framework analyses and theories of change shape our working practices. Why and under what conditions do potentially useful tools - such as Theories of Change - mutate into coercive instruments that reduce the space for choice? These tools and methods can have perverse consequences because of their hidden and invisible power to determine what knowledge counts when hierarchical ways of working block communications and dialogue.

Just as tools and methods shape practice, so context specific practice shapes the tools. Their power is neither uniform nor constant. There *is* room for manoeuvre to expand approaches to evidence within the sector, enabled by an analysis of the politics of accountability and the sector’s internal dynamics.

The paper concludes with the contradictions arising from the political pressures to seem to be in control in a world of uncertainty and surprises and identifies these as potential opportunities for changing how the sector deals with results and evidence.

1. The Readership

This is a background paper for a Conference organised by the Big Push Forward, an initiative seeking political space for development practitioners to support transformative and locally owned development. Readers are development practitioners who want to support organised efforts to change the power relations and structures that create and reproduce inequality, injustice and the non-fulfilment of human rights - and who are concerned that today's narrow meanings of 'evidence' and 'results' risk undermining locally-owned and transformational development.*

Development practitioners are defined as people working in the international development sector - in a bilateral/ multilateral agency/ international NGO, as staff member or consultant; or employed by a government department or a non-governmental organisation in a low-income or middle-income country that receives financial aid. Or s/he might be located in a university/ think tank partially funded by the sector, a private sector consulting company or a philanthropic foundation. (I am located in the 'think tank' category, working in a policy research institute in the UK). We are employed by organisations receiving and/or disbursing funds in order to work with other intermediaries to achieve a contractually agreed purpose. These relationships are political. Power dynamics influence whose and what knowledge counts and which results matter in development policy and practice. The politics of 'results and evidence' is about how power determines the meanings that influence a course of action, for example what is considered 'robust evidence'. Such politics are not restricted to the development sector. Similar trends in policy and programming within donor countries – which have been extensively criticised¹ - share a genealogy that this Paper discusses. Readers outside the sector may thus also find elements of the Paper relevant and useful. Conversations across sectors about the politics of evidence can be mutually beneficial and help re-imagine and reconceptualise how to promote of transformational development.

“The conversations I hear in aid/development circles ... are happening as well with social change groups in the US ... but there has been little connection between these international and US-based groups.

Catherine Borgman-Arboleda <http://bigpushforward.net/archives/1786>

2. Aims and Approach

I seek to uncover how the diffuse power of New Public Management (p.13) shapes the design and evaluation of development projects and programmes through a discourse of 'results' and 'evidence'. The aim is to help practitioners strategize in expanding the politico-bureaucratic space for flexible and creative support of locally-generated and transformative change.

* I am grateful to Charlotte Cross for her background research for this paper and to Itil Asmon, Irene Guijt, Sulu Mathew, Marjoke Ostrom, Chris Roche, Cathy Shutt and Brendan Whitty for comments on earlier drafts. My time for writing was covered as part of a grant from the Swiss Development Cooperation Agency to the IDS Participation, Power and Social Change Team, and I gratefully acknowledge SDC's and the team's support. The opinions expressed are my own and do not necessarily reflect the views of SDC or of my colleagues.

But this is *not* a paper about the methodologies and underlying principles in providing such support. Many resources about participatory and democratic methods are already available to development practitioners.² For example, the Swiss Development Cooperation Agency (SDC) recently commissioned a ‘Good Practice Note’³ on the monitoring and evaluation of empowerment, which highlights the methodological principles for enabling development programmes to support changes in power relations. Rather, I am concerned with the discourses and administrative practices that are blocking the many well-intentioned development practitioners from putting these principles in practice. The SDC paper just cited notes the tension between a flexible and enabling role in supporting an empowerment process of social change and the obligation to demonstrate results and fulfil internal rules and regulations. That tension is no longer creative. The obligation to demonstrate results is increasingly dominant, reducing the opportunities for flexible and responsive approaches that support to empowering processes requires. The Paper looks at why and how this came about. It includes quotes from comments posted on the Big Push Forward website and from contributions to the internet debate that Chris Roche and I had with DFID’s Chief Scientist and Chief Economist.⁴

Box 1. A set of tools quite distinct from what we were actually doing

‘I tried and failed many times to explain to our donor organizations why processes had an importance beyond the results they achieved. The results-based framework within which we operated existed in the context of complex power relationships... Sometimes we found ourselves talking openly and finding support from among the donors, while at other times we had to conceal our true objectives and ensure that the results-based, logical framework outputs were achieved... We found ourselves adopting a language and a set of tools – technical activity reports, expenditure reports and products – quite distinct from the work we were actually doing...

When our consortium members protest or refuse to abide by parameters that do not reflect their political reality, how am I to respond? Has the time come to liberate ourselves from the project tools that force us to frame our activities within parameters established by others?’

Source: Rosario León (Bolivia) www.thebrokeronline.eu/Series/Stories-from-Aidland/The-ghost-in-the-aid-machine

I do not accept the radical criticism that the international development sector serves only to maintain an inequitable global status quo and that progressively-minded practitioners are naive to imagine they can work from within the sector to help make a better world.⁵ Rather, I take the view that the sector’s complexity offers many opportunities for supporting transformation, that such opportunities have been seized at times in the past to good effect, and such opportunities may still be seized today – provided we are critically aware of how power reinforces the development sector’s results-and-evidence discourse, and that we challenge this discourse with good alternatives.

The Paper seeks to answer **frequently asked questions by development practitioners** concerned about the current trend. These include:

- Why and under what conditions do potentially useful approaches - such as theories of change - mutate into coercive instruments?
- Where and how did the results-based management and evidence-based policy and programming originate?
- Why are ‘results’ and ‘evidence’ discourses increasingly influential, particularly in the international development sector?
- Which organisational actors and interactions are promoting these discourses?
- Who is coming under pressure and how are some people resisting?
- What are the effects of the power of these discourses on transformative agendas?
- What are the possibilities to create spaces for different approaches?

I start by looking at the ‘institutional artefacts’ – rules and procedures - that translate the discourse of results and evidence to reality of practice. I then unpack ‘results’ and ‘evidence’ and find strong family resemblance between the two discourses. These share a certain way of knowing the world (epistemology), including assumptions that evidence pertains to verifiable and measurable facts as categories of things⁶ and a particular understanding of causality, efficiency and accountability.

“It is no secret that often the use of log frames, SMART tables, and even many efforts at theories of change are developed... primarily for establishing a framework for funder mandated evaluations that are essentially symbolic in nature”.

Juliette Majot <http://bigpushforward.net/archives/532/comment-page-1#comment-26>

I look at how and why these discourses have played out in the development sector, including what and who is driving the agenda. I make a first stab at identifying some of the contradictions and trends that might help identify the loci and strategies of resistance. In that respect, the Conference aims to surface resistance and make the results-and-evidence agenda a legitimate subject of debate. Chris Roche has blogged⁷ that people fear for their individual careers or for their organisational standing because should they object they risk aggravating donors and putting their funding at risk. Thus it is hoped this paper and the Conference itself will help such people gain the *courage* and *confidence* to navigate political space and put pressure on the system so that front-line practitioners like Rosario (Box1) can use donor funds to respond effectively to their local realities.

3. Results-and-Evidence Artefacts

Development projects and programmes are increasingly being planned, appraised, implemented and evaluated within a ‘results’ and ‘evidence’ framework. This framework uses a particular meaning of ‘evidence’ as what works to solve a problem as the basis for delivering ‘results’, reported upon and sometimes evaluated. Results-and-evidence discourses shape our working practices through ‘artefacts’, such as logical framework analyses. Various results-and-evidence artefacts are used at various stages in the development sector’s funding cycles.

Results artefacts are used very widely within the sector for planning, implementation, monitoring and evaluation. They are reporting, tracking and disbursement mechanisms that include:

- Base-line data
- Results reports
- Progress reviews
- Performance measurement indicators
- Logical framework analysis
- Risk register
- Theories of Change
- Payment by Results

Evidence artefacts are popular with DFID, USAID, the World Bank and the Gates Foundation and are concerned with finding out what works best and therefore delivering value for money. These artefacts are used to *paper choice* of intervention, for appraising proposals and for evaluating effectiveness and impact with respect to value for money. Evidence artefacts include:

- Randomized control trials
- Systematic reviews
- Cost-effectiveness analysis
- Option appraisal
- Social return on investment
- Business cases
- Impact evaluation.

Results and evidence artefacts can be very useful. But they can also be misused to make a travesty of reality when power relations insist on their use in inappropriate circumstances (Box 2).

Box 2. Abusing Cost/Benefit Analysis

At a meeting called by the Director General to determine whether the agency should start a programme to get girls into school in a certain fragile state, the education programme officer for the agency's fragile states unit was asked by the D.G. to provide the unit cost of educating a girl through primary school. He was irritated by her reply that there was insufficient evidence to answer that question, and accused her of sloppiness. She looked round the table for colleagues' support but they were all looking down at their laps. When the meeting was over, she asked them why they did not come to her rescue. They answered, 'You should not have resisted. Just invent a number and give it to him'.

Source: Adapted from an account given to the author

From the case in Box 2, it would appear that the Director General's belief that cost benefit analysis is not only desirable but always feasible is confirmed each time subordinates' play the game by providing spurious numbers. His assumption about the universal applicability of cost benefit analysis will be reinforced leading to even more vigorous demands that such a tool be used in all circumstances.

3.1. The disciplining effect of artefacts

Whether we find an artefact useful in our endeavours will influence how we feel about it. Our personality, experience and kind of job - for example whether we are independent consultants or in-house staff - may also influence our response. But the emotional and power effects such artefacts have on us also depends on the kind of organisation we work for - not only with respect to its position in the aid chain but also to its institutional culture and leadership.

Results and evidence artefacts are intended to improve clarity and thought but can have **perverse consequences** when used in an organisational environment in which hidden and invisible power⁸ determines what knowledge counts and hierarchical ways of working block communications and dialogue.⁹

The artefacts, more than the discourse that has produced them, shape practitioners' lives. Language can be ignored, but artefacts influence every day of work in the development sector. They are '**technologies of power**', implemented and enforced by authority, but often also internalised so that no obvious external control is required (Box 3).¹⁰ With internalisation, artefacts take on a life of their own, independent of the authority that had initially required their use. An organisation may voluntarily adopt such artefacts in the absence of any mandatory donor requirement – and hire the staff with the skills needed to use them.

Box 3. Internalising the Results Agenda

The bureaucratic imperative has reached down to the grassroots, and local organisations have internalised the results agenda. An INGO participant at a Big Push Forward meeting in New York recounted how a conversation they managed to have with their donors allowed the INGO to push back on the results agenda and to support its local partners without requiring log frames or other quantifiable output-oriented mandatory procedures. Yet, when she visited partners to invite stories of change, they provided her with lists of quantified results - what they had learnt that international organisations expected from them.

Source: <http://bigpushforward.net/archives/1786>

Auto-surveillance creates a disciplined practitioner. When in a position of authority, such practitioners demand that subordinates follow rules and procedures according to how these have been internalised and interpreted, deaf to arguments that what they are asked to do may be contrary to what the originating authority might be requiring or expecting. Or even when their use *is* mandatory, a grant-receiving organisation may be more exigent and controlling in how they are used than had ever been envisaged by the artefacts' originators. It is common to find intermediary organisations 'more Catholic than the Pope' in applying the results-and-evidence doctrine.

On the other hand, the disciplinary power of artefacts is **far from totalizing**. Effects are contingent on personality, seniority, experience, capacity for reflexivity, etc., as well as the institutional culture of the person's organisation. Two grant recipients, equally dependent financially on a funding agency, may respond differently in the extent to which they internalise an artefact's power and impose it on others.

I suspect there may be '**a squeezed middle**' of programme officers (those responsible for managing grants including in-country programmes) that most strongly experience these artefacts as tools of a diffuse tyranny. The squeezed middle is pressurized to behave against their better judgement, while they try to protect front-line practitioners and partner organisations from the deleterious effects of such artefacts. The 'squeezed middle' either mock or vent their anger and then cynically comply. They try to do the least possible to secure the funding, possibly with a nod and a wink from a sympathetic bureaucrat in the donor organisation who is equally despairing of mandatory requirements. Cynical compliance may sometimes be accompanied by secret resistance: people carry on working according to their own professional judgement, while reporting up the system what they perceive to be ridiculous numbers. Although websites like 'Stuff Expatriate Aid Workers Like' provide space to let off steam,¹¹ compliance and resistance consume energy and enthusiasm.

3.2. The rise and fall of artefacts

How artefacts are used - and with what intentions - seems to change over time. I was one of the early advocates of the Logical Framework in DFID in the late 1980s. I had learnt to use it while working as a consultant for USAID, and found it useful in encouraging colleagues to ask themselves why they wanted to undertake certain activities. A log frame, I argued, requires you to state a purpose – for example, increased incomes of landless families – and then work out what you could do to achieve that purpose. British projects at the time did the opposite: an available technology or resource was identified and then a half-hearted effort was made to justify its use by trying to show how it would improve the world. I also appreciated the Log Frame's assumptions column, which in the early days has been used to identify the dynamic social and political environment of aid projects, thus introducing uncertainty and requiring iterative planning. Did other routes to improving landless people's incomes become apparent during project implementation? Has an unforeseen event occurred which requires a quite different kind of response?

However, nowadays, the log frame has become a rigid tool, demanding ever more precise and pre-determined 'results' with SMART indicators. As a consequence a new artefact, the Theory of Change, has been introduced to open things up by again asking questions about assumptions and conditions. A struggle has been taking place within DFID about whether ToC should be a mandatory requirement, with objectors worried that any such requirement risked making it another micro-technology of power that requires obedience rather than independent and above all, critical thought.

As practitioners modify artefacts to fit with everyday organisational practice, the artefacts' power in enforcing the results-and-evidence discourse may wane. Then, once again, new artefacts may be introduced to shore things up. Writing about 'performance measurement systems', Michael Power notes they operate in a near-continual cycle of reform. **'Dreams of measurement for control purposes'** are articulated; these are shown to be defective and/or leading to adverse unintended consequences; new measures and refinements are proposed'.¹² These changes may make the artefact more controlling, as happened with Log Frames, but sometimes refinement may mean a dilution of excessive control, as with Systematic Reviews, whose designers are showing signs of greater flexibility in the definition of evidence,¹³ possibly because of academic involvement in their design and application with consequent greater openness to other points of view.

3.3. The effect on transformative agendas

The increasing dominance of these artefacts coincides with and may be a consequence/cause of the aid funding landscape changing, notably with donors' interest in rights and social transformation declining. The general proposition articulated by the former Director of USAID is that work that emphasizes measurable outcomes tends to drive out work that produces immeasurable ones, 'That crowding out of less measurable activities has in turn led to a greater emphasis on service delivery instead of institution building and policy reform as the predominant programmatic approach to development.'¹⁴

Little research has been done so far on the influence of results and evidence artefacts on how and why aid is spent.¹⁵ However, a recent study of historical changes to donor support to women's rights organisations¹⁶ found that donor emphasis on management systems and reporting allowed organisational strengthening and enhanced effectiveness, particularly when there was a sense of partnership and trust within a long-term relationship. On the other hand, short-term and fluctuating project-related rather than core funding, combined with the donor pooling of funds decrease the potential for good quality, direct relationships. Because it discouraged experimentation and risk taking, today's emphasis on inter-organisational competition for increasingly scarce funds were seen as detrimental to mutual efforts to secure social transformational outcomes.



"Dear Mr. Gandhi, We regret we cannot fund your proposal because the link between spinning cloth and the fall of the British Empire was not clear to us."

Written by *M. M. Rogers* and illustrated by *Ari: R. Fátzal, Wahyu S., Ary W.S.*
Creative team for Search for Common Ground in Indonesia

4. Approaches Based on Results and Evidence

*‘The study of politics is largely a matter of recovering the meanings that inform actions’.*¹⁷

This section uncovers the origins and affinities of the ‘results and ‘evidence’ discourses in the politics of public sector management. Recognizing their historical genesis within a *particular* context and time is an important step in challenging their ubiquity and relevance for *all* times and purposes.

‘Results’ and ‘evidence’ have a common an understanding of causality, efficiency and accountability that originated in and remains more prevalent in countries with an Anglo-Saxon empiricist tradition. However, through the dominance of English-language-based global institutions such as the World Bank, they are discourses spreading widely within the international development sector and into developing countries. David Mosse has called them ‘travelling rationalities’.¹⁸

This, the Paper’s longest section, starts with results-based approaches. Descended from an earlier discourse, management by objectives, they entered the development sector about ten years ago, leading to an emphasis on quantifiable, relatively easy to measure targets. Below I enquire into their public sector origins and investigate them as an element of New Public Management, informed by the theory that desired outcomes are achieved through incentives (carrots and sticks) that control people’s behaviour. I review critiques, including a brief discussion of the implications for the framing of ‘accountability’ before examining ‘results’ in the development sector.

“At the heart of the debate of what counts as evidence is a contest of different development philosophies”.

Sam <http://bigpushforward.net/archives/1786#comments>

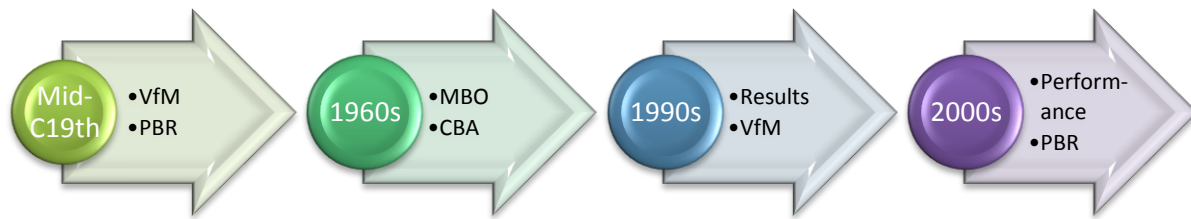
I then move to evidence-based approaches in public sector management that share a common intellectual heritage with results, although in this case they derive from recent changes to medical practice. These approaches fit easily into the ethos of a development sector that historically tended to treat people as subjects requiring treatment rather than as citizens with voice. Their recent use by the sector means that so far few agencies are promoting them. I consider the likelihood of them spreading more widely.4.1

4.1. Results-based approaches

Results-based approaches rely solely on quantification. Statistics have been instruments of progressive change in the hands of reformers and campaigners using numbers to reveal social injustice and to inform sensible, fair policy making. However, statistics have also been put to use as instruments of repression, for categorizing and controlling deviant domestic populations - suicides, prostitutes, criminals etc – people different from us. This may explain why bureaucratic quantification was most fully developed and applied in the colonies where everyone was different from the colonizing power.¹⁹ Aggregate numbers have the effect of turning people into de-contextualised, homogenized ‘others’.

4.1.1 History

Figure 2 Historical trajectory of the ‘results’ discourse ²⁰



The late 18th century saw enormous and rapid social, economic and political change as Europe and North America industrialised and proceeded to colonize the rest of the world. The positivist power of numbers appeared to tame uncertainty in an era of such rapid change.²¹ By the mid-19th Century, the politico-managerial public sector predilection for measurable facts led in England and Wales to introducing ‘payment by results’ (PBR) into the elementary schools at a time when the discourse was of efficiency, value of money (VfM), competition and a balanced budget.²² PBR was discredited and abolished at the end of the century due to the increased bureaucracy and administration costs of verifying the results.²³

The precursor of results, ‘Management by Objectives’ (MBO), was adopted in the 1960s by the Democratic administration in the United States as ‘planning programming and budgeting’, with associated artefacts such as cost-benefit analysis (CBA) and risk assessment.²⁴

Box 4. Sweden - an Early Adopter

Back in the 17th Century, long before modern economic theory was established, the idea of principal and agent establishing a contractual agreement about the objectives to be achieved was an early attempt by central government to exert more control over its semi-autonomous agencies. Sweden became an early European adopter. In 1970, the Swedish National Audit Office was emphasising that each agency had to be results-oriented and results-conscious. That a Social Democratic government introduced MBO illustrates that this discourse fitted as well into a big-state as into the minimal-state ideology associated with the Right.

Source: G.Sundstrom (2006) Management by Results: Its Origin and Development in the Case of the Swedish State, *International Public Management Journal*, 9:4,

Adopted by both left- and right-wing governments (particularly in the Anglo-Saxon world)²⁵ ‘objectives’ had morphed into ‘results’ by the mid-1990s. Again, the impetus came from a United States Democratic administration borrowing from private-sector accounting the rationalisation of resource allocation decisions.²⁶

All this was part of a broader shift in public-sector management approaches, commonly known as ‘**New Public Management**’ (NPM) that emerged in the late 1970s. Today prevailing in various forms in most OECD countries it is exported through their development agencies to aid-receiving countries. In addition to ‘results’, NPM includes linking resource allocation to performance, competition between providers of services, greater discipline and parsimony in resource use, and adoption of what is represented as private-sector management practices.²⁷ Thus accountancy expanded from checking whether the money had been spent in accordance with the rules, to checking on efficiency and effectiveness. This required not only determining *how* to achieve efficiency and effectiveness through ex-ante cause-effect logic, but also determining through auditing ‘rituals of verification *whether* the pre-determined results had been achieved’.²⁸ **Accountability for results** or performance against pre-established objectives is a response to the ‘principal-agent problem’, a theory positing that because individuals are assumed to be always in pursuit of their own selfish interests, policy intentions are likely to be subverted by those designated to implement them. The application of this theory can be observed when organisations introduce positive and negative incentives to align actors’ interests with those of the agenda-setters. Changing behaviour through ‘incentive structures’ that may include bonuses as carrots and threat of dismissal as sticks, has become so ‘naturalised’ in modern management that most of us do not appreciate it is a practice derived from just one among many contested theories that seek to explain social reality and human behaviour.

The application of results-based management artefacts and the growth of ‘the audit culture’²⁹ resulted in strong reactions and criticism from professional bodies, including teachers and

academics. They found that their discretionary space was shrinking through the demand to deliver results determined by central government. An increasing amount of their time was being devoted to performance measurement and reporting against targets, to the detriment of time spent actually doing their job. Critics argued that performance indicators, when used for control, ‘are unreliable: they do not measure performance itself, distort what is measured, influence practice towards what is being measured and cause unmeasured parts to get neglected’.³⁰

As a countervailing trend, Complexity Theory became popular from the 1990s onwards to inform ‘a more realistic and democratic approach to achieving policy goals than the audit culture of performance management’.³¹ Chapman, illustrating his argument with reference to the UK National Health Service, contrasts bounded with unbounded problems. With the first there is broad agreement on the nature of the problem; there is some mutual understanding of what a solution would look like; and there are limits to what is required in terms of the time and resources required for their resolution.³² With unbounded problems, however, there is no agreement about the diagnosis and therefore on the actions required; no possibility of an eventual permanent solution because solutions generate new problems; and therefore no



way of determining the quantity and type of resources needed. Because NPM treats all problems as bounded ones, perverse consequences arise.

“The reality of many interventions is that programmes are based on the 3-5 year ‘projectible change’ model: here’s a problem, here’s an amount of money, here’s what we want to see, go to work.What this then leads to in VFM terms is the need to very quickly find data that supports both the initial assertion and the need of the implementing agency that it is succeeding. The result is poor use of evidence, leading to poorly informed decision making. So what is a possible alternative? Simplicity within complexity. So: here’s an issue, what is happening, how can we support this, how’s that going, how can we do that differently or better?”

Jake Allen <http://www.oxfamblogs.org/fp2p/?p=13327>

People have also challenged NPM’s interpretation of accountability as ‘the delivery of results’. The methods demanded of us to be *more accountable* have the effect of our becoming ever *less responsible* for seriously enquiring of ourselves how we can most usefully contribute to transformative social change and be held accountable for our commitment in that respect (Box 5). Thus, given that any actor has only partial understanding of a complex system and given how power constructs knowledge, an alternative approach to accountability is ‘responsible action’, as demonstrated through adaptive learning and multi-stakeholder dialogue. Such an approach requires balancing measuring outcomes with tracking responsiveness in the change process. It focuses on the ‘how’ as much as the ‘what’.³³

Box 5. ‘Strategic accountability’

“Critically important – and one of the places where accountability and learning converge – is that accountability can also be taken to mean taking responsibility for oneself. Understanding what you’ve done, being able to respond to questions about the basis of strategic decisions, the underlying theory of change and, of course, how money was spent. Such *strategic accountability* seeks to answer the question ‘Did I/others/organizations/institutions act as effectively as possible?’ In this sense, accountability is intrinsically about identity – feeling committed to one’s ideas and strategies (Fry, 1995). Ebrahim (2005) echoes this by saying that: ‘Organizational learning is more likely if internal accountability to mission, rather than upward accountability to donors, papers NGO reporting’. Being held accountable thus means having ‘respond-ability’.”

Source: Guijt, I. (2010). Exploding the Myth of Incompatibility between Accountability and Learning’. J. Ubels, N.-A. Acquaye-Baddoo and A. Fowler, *Capacity Development in Practice.*: Earthscan p. 283

These critiques of the results agenda – particularly of its perverse consequences – may influence UK domestic policy arenas such as health and policing.³⁴ Nevertheless even here there continues a trend to tackle complex issues where the voice of those affected is muted, such as youth unemployment, as a problem of under-performance of front-line workers. Refusing to abandon their diagnosis, agenda-setters are increasing the dosage of the prescription by the introduction of **Payment By Results**.

Box 6. Payment by Results (PBR)

Payment by Results is where ‘commissioners of services (e.g. a government) pay the service providers only after a pre-determined result has been achieved and independently verified. The logic of PBR is that there is a manageable level of risk in achieving the result and that service providers must be incentivised to play a more active role (co-production). In a [recent speech](#), the British Prime Minister stressed that PBR helps tackle intractable social problems. PBR is being increasingly used in the UK and USA to find jobs for the long-term unemployed; break the cycle of re-offending; and placing children more promptly in stable foster care. Because the social distance of unemployed, offenders and children makes them as unknowable to those in power as erstwhile colonial subjects, these are people most likely to be subjected to this kind of treatment.

Otherwise known as ‘performance-based incentives’ (USAID) or ‘payment for success’ (KPMG), Payment by Results is premised on a predictable world. PBR, has become increasingly popular in those OECD countries where NPM was first adopted and where less rigorous systems have been tried and judged to have failed. In the United States, PBR has returned again to schools. Reporting on one such scheme, the Economist commented: ‘You are transferring from a system where the agents are (to a degree) public-spirited individuals to one that motivates agents to be self-interested’.³⁵

4.1.2 Development results

The development sector, while influenced by the wider public-sector management discourse, has also innovated with artefacts and practices. The Log Frame, for example, was first used by USAID in the late 1960s to promote management by objectives, and was taken up domestically in the USA in the 1990s. The ‘results’ (as distinct from ‘objectives’) language became general in the sector after the 2002 Monterrey Conference on Financing for Development. There it was agreed that improved development effectiveness required a focus on results and a systematic approach to the monitoring of the actions of all parties. In 2003 the Joint Venture on Managing for Development Results, incorporating bilateral and multilateral donors, was set up within the DAC Working Party on Aid Effectiveness. The Paris Declaration on Effective Aid (2005) included ‘results’ among its five pillars. In Sweden the results language had been in use since the 1970s. When in 2007, Sida launched its ‘managing for results’, this was justified as a response to demands of the development sector rather than to its government’s forty-year-old discourse.

Box 7. Delivering Results

‘The development machine has many levers but only one engine – *aid*. Aid drives the gears that launch the projects that churn out money and deliver *results*. I have been one of the machine’s lowly operators, a local servant in Aidland. I applied the machine’s instruments and tools to manage and control the people to ensure a sustained production of measurable results.

The donor agencies operate remotely, using fine antennas to connect to the head office, making just enough contact to allow me to hear their instructions. These are delivered in a neutral voice. The donors must keep their hearts hardened and never establish friendly relations with those on the factory floor – otherwise they might learn about another reality. A reality very different from the one for which their machine was designed’.

Source: Rosario León (Bolivia) www.thebrokeronline.eu/Series/Stories-from-Aidland/The-ghost-in-the-aid-machine

Likewise, Switzerland cited the Paris Declaration when publishing a primer on results-based management (RBM). The primer introduced a revised log frame with respect to ‘the results chain [that] clearly shows the plausible, causal relationships among the elements’ and stresses the importance of base-lines and ‘key performance indicators’ (KPIs).³⁶ The primer’s example is tertiary-level training for bankers.

Two years earlier DFID had also introduced a revised log frame that included mandatory base-line numbers. The Dutch Government revised its co-financing agreement with its big INGOs, obliging them to make a single global log-frame with respect to the programmes of multiple partners in many different countries and requiring aggregate data on outputs and outcomes from different countries. In 2010, with a change in government and an even greater emphasis on results alongside the introduction of a new artefact – the ‘business case’ - DFID revised yet again its log frame and, like the Swiss, emphasised the results chain that ‘must be based on evidence about what has worked in the past, so this is a real opportunity to take account of all the lessons learned, evaluation and research evidence available that underpins the design of the project.’³⁷ The worked example concerned training health professionals – like the Swiss example, providing easily quantifiable data. Health and education interventions appear most amenable to counting (Box 8).

Mainstream commentators like Owen Barder have pointed out the risks of focusing on results. First it may add to bureaucratic overload. Second, it may make aid less strategic and short-termist. Third, it may impose the wrong priorities. Fourth, it may ignore equity. Fifth, it may create perverse incentives. Sixth, it may inhibit partnership. Seventh, the results information is all bogus anyway as claims about results must rely on assumptions about the counterfactual which are usually flawed or incomplete.³⁸

Box 8a. DFID's Results

Secure schooling for 11 million children – more than we educate in the UK but at 2.5% of the cost.

Help vaccinate more children against preventable diseases than there are people in the whole of England.

Provide access to safe drinking water and improved sanitation to more people than there are in Scotland, Wales and Northern Ireland.

Save the lives of 50,000 women in pregnancy and childbirth and stop 250,000 newborn babies dying needlessly.

Source: www.dfid.gov.uk

Box 8b. World Bank's Results

What counts as a development result? The World Bank Group and its country partners are documenting development outputs and outcomes, to better understand the wider impact on results. The focus, for example, is on counting both the number of hospitals or rural health clinics built and the impact that their services had on the lives of people who used them. Did access to health care improve? How many children were inoculated against infectious diseases? What was the impact on under-five child mortality? Were more children able to go to school because of better access?

Source: <http://web.worldbank.org>

Nevertheless, instead of addressing these concerns, agencies appear to be strengthening the focus on results and related managerial instruments. Ignoring its use 150 years ago, DFID sees PBR as a 'promising new instrument'³⁹. In 2008 the World Bank's Health Results Innovations Trust Fund (HRITF) was established, supported by the governments of Norway and the United Kingdom.⁴⁰ USAID is rolling PBR out across the world in health and family planning programming; the Global Partnership for Output-Based Aid, established in 2008 by the World Bank and DFID, is financing PBR projects in a variety of sectors and countries and is also supporting AusAid and the Asian Development Bank in developing such projects.⁴¹ 'Cash on delivery' is a version of PBR promoted by the Centre for Global Development.⁴²

4.2. Evidence-Based Approaches

Entering the development sector from medical practice, 'Evidence' has acquired a particular meaning relating to 'what works' – a narrow discourse in which the 'how' of context and process is ignored. It also deflects attention from the ideologies and values that shape development initiatives.

The technical question of 'what works' exacerbates the sector's tendency to see people as subjects requiring treatment rather than as citizens with political voice. Power may silence any challenge to the technical framing of 'the problem', foreclosing discussion of the structural causes and consequences of inequity and how to tackle these. To act 'technically' in a politically complex context can make development practitioners pawns of powerful vested interests and therefore by default, albeit unintentionally, political actors.

4.2.1 History

Whereas ‘results’ can be traced back to utilitarian concerns for efficiency, the origins of ‘evidence’ lie in the mid-19th Century French innovation of experimental medicine.⁴³ However, ‘evidence-based medicine’ was developed less than thirty years ago in Canada, whence it spread to the UK and the USA. Randomized trials were cited as the ‘gold standard’ for judging whether a treatment does more good than harm.⁴⁴ As this particular understanding of evidence – as distinct from its broader meaning, rapidly moved into other policy arenas, it became associated with the same issues of efficiency and effectiveness that had generated the results discourse.⁴⁵

The rise of objectivity typified in the evidence discourse may be a consequence of the Cold War’s end and the disappearance of ideological battles between left and right in Europe, where it became fashionable for policy not to be driven by ideology. In the UK, the 1997 New Labour government was keen to demonstrate its decisions would not be driven by socialist ideology but by technical, un-biased evidence.⁴⁶ Yet, even within medicine, the leap from evidence-based clinical practice into evidence-based health policy was subjected to challenge. For example, an article in the British Medical Journal argued that policy was shaped by institutional arrangements, values and beliefs and by a variety of different sources of information.⁴⁷

Meanwhile, opponents of evidence-based education questioned the homology between education and medicine, including the approach’s linear cause-effect thinking and its dismissal of values.⁴⁸ Evidence-based approaches with their experimental or quasi-experimental design were criticised as inapplicable to complex social policy contexts. Tony Harrison, bucked the editorial trend in a book promoting evidence-based approaches, argued these did not work at the wider community level where multiple perspectives come into play and there is no shared agreement about the problem.⁴⁹ In other words, policies are an outcome of a political process in which various kinds of evidence may be used to inform opinion but these are not usually the deciding element. Hence, the reference to ‘evidence-informed’ policies.⁵⁰

“In the UK, health policy is decided by a great many number of factors or appeals (evidence, sure, but also values, tradition, biases, political calculations, etc.).... But health policy in heavily Aid dependent countries) is decided mainly by evidence (or what often passes as evidence at the time) and usually by foreign experts, albeit with the acknowledgement of the importance of politics (which is not the same as participating in politics)”.

Enrique Mendizibal <http://www.oxfamblogs.org/fp2p/?p=13327>

4.2.2 Evidence-based development

Embarking on a social-transformation initiative through development aid is not the same as administering a pill.⁵¹

Evidence-based approaches share with results a common intellectual heritage of rational choice theory concerned with individual behaviour. This is ‘methodological individualism’ that economics shares with medicine. It is distinct from other, holistic social sciences that look at the relations between people and the culture and history that shape these. **Methodological individualism** is reflected in DFID’s new draft evaluation policy ‘Learning What Works to Improve Lives’. The policy is about quick wins based on targets as articulated in numbers of people. This contrasts with a holistic approach that supports societal change for reducing poverty.

DFID’s current approach is homologous to colonial bureaucracies whose categorizing, counting and objectifying developed a way of working that saw people as objectified subjects requiring intervention and treatment.

Similar methods were then applied back home to control the socially deviant and the poor. In both instances, people were objectified when they lacked political voice and therefore silenced from challenging the technocrats’ framing of ‘the problem’ and from offering alternative ways of thinking about the issue.

“When the attempt is to test ‘if something or what works’, it is not a simple case of observing X leads to Y. These tests will have to explain: a) why what is being tested by ‘us’ is relevant for the context (otherwise we are using people to test our ideas for our purposes)”.

Hakan Seckinelgin <http://www.oxfamblogs.org/fp2p/?p=13327>

Evidence-based approaches aim to **build an anti-politics firewall**. Development assistance becomes a ‘technical’ best-practice intervention based on rigorous objective evidence, delivering best value for money to domestic taxpayers and recipient country citizens, without interfering in that country’s politics. They are the latest manifestation of a certain long-standing approach to development that speaks to the sector’s ‘need to overlook its internal involvement in the places and problems it analyses and present itself instead as an external intelligence that stands outside the objects it describes’.⁵²

Evidence-based approaches in development have been driven by a relatively small group of institutional actors and individual academics that share a commitment to improving impact evaluation. According to



Robert Picciotto (former Director of the World Bank's Evaluation Department), a stalemate in the battle between the macro-economists Sachs and Easterly about aid effectiveness provided the opportunity for a group of young micro-economists at the Massachusetts Institute of Technology (MIT) to shift this debate to a clinical examination of specific development interventions. 'The battle cry of the new development economists was sounded by the MIT Poverty Action Lab's charismatic co-founder (Esther Duflo) when she famously declared during a World Bank Conference on the evaluation of development effectiveness held in 2003: "Just as randomized evaluations revolutionized medicine in the 20th century, they have the potential to revolutionize social policy during the 21st century".⁵³

Box 9. The African Laboratory

In the 1930s Africa was seen as 'a living laboratory' to achieve improvements in the welfare of the populations. Evidence-based approaches are reviving development as laboratory idea. In 2012 the World Bank established a [Gender Innovation Lab](#) to design 'innovative interventions to address gender inequality and to develop rigorous research projects in order to produce evidence on what works and what does not'. Jeffrey Sachs' Millennium Villages have been framed as [laboratories to lift people out of poverty](#)'. The best-known is the J-Pal Poverty Action Lab, whose mission is to reduce poverty 'by ensuring that [policy is based on scientific evidence](#)'.

The Poverty Action Lab (now named JPal) has received much favourable publicity for its impact evaluations. Dr. Duflo particularly has had much media coverage, including since publishing with Abhijit Banerjee (co-founder and director of JPAL) *Poor Economics*, named one of the top books of 2011 by The Economist, the Financial Times and Goldman Sachs. Finding out what works looks like eminent common sense to business-minded people, which explains why evidence-based practice has become popular for example with private philanthropy.

Enthusiastic support and very substantial funding from large private foundations, particularly Gates and Hewlett, generated enormous momentum behind this evidence-based movement, with many professional ties between these two foundations and those who share their objective within development agencies such as the World Bank, DFID and USAID.

According to 3IE - an organisation created by donor agencies in 2008 to enhance development effectiveness through the promotion of evidence-based policy making - high-quality **impact evaluations** that measure the net change in outcomes amongst a particular group, or groups, of people that can be attributed to a specific programme. They have narrowed the debate from the older OECD Definition of 'impact' which is 'the positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended'. Hence, increasingly impact evaluation (IE) tends to be associated with experimental and quasi-experimental methods because of their 'objectivity'. USAID evaluation policy today states: 'Impact evaluations in which comparisons are made between beneficiaries that are randomly assigned to either a treatment or a control group provide the strongest evidence of a relationship between the intervention under study and the outcome measured.'⁵⁴ At the same time, donors in support of this approach are promoting the *demand* for IEs in aid-recipient governments.

Evidence-based approaches in development have not as yet spread as widely within the donor government and NGO parts of the development sector as have results-based approaches. Nevertheless, even in countries like France where there is a strong intellectual tradition of the philosophy of ideas, compared to the a-theoretical empiricism of English-speaking countries, evidence-based approaches are being discussed in the development literature.⁵⁵ Their penetration into developing countries is triggering the same debate as in the countries where these originated. In South Africa, Andries DuToit argues that “designing ‘good’ pro-poor policy is not something that can be guaranteed by focusing on generic, technical ‘best practice’ skills: rather, it requires a clear focus on the substantive issues at hand, and the concrete institutional, fragile and contested political nature of the South African state”.⁵⁶

Meanwhile, there has been a strong negative reaction among mainstream schools of development evaluation is typified by Robert Picciotto, cited earlier, who concludes that the bubble is bound to burst because evidence-based approaches are appropriate only for ‘relatively simple interventions, the effects of which are realized in a short period of time and are large relative to other potential influences’.⁵⁷

Box 10. The Gates Foundation at the Heart of Evidence-Based Approaches in Development

The Gates Foundation is credited with having been extremely influential in promoting the evidence-based agenda. For example, in a January 2012 [interview](#) with Channel 4 news, JPAL’s Executive Director commented on Gates’ influence ‘in shaping the way governments give their funding, the priorities they choose and how they focus on innovation, crediting Gates with having been one of the major advocates of prioritising cost-effective, proven solutions. [Similar sentiments](#) were expressed by the president of the Center for Global Development, who described the Gates Foundation’s emphasis on results as one reason for its success. During a recent visit to the UK, the Gates stressed the importance of ‘return on investment’ and the need to increase impact assessment. The Gates Foundation was instrumental in setting up 3IE and is its biggest annual donor. 3IE funds impact evaluations and systematic reviews.

To conclude, evidence-based approaches are a recent phenomenon closely linked and helping reinforce a deeper and more influential results discourse. I now turn to the deeper drivers that make the results and evidence agendas so influential in today’s development sector.

5. The Drivers of the Results-and-Evidence Agenda - and the Countervailing Trends

This section starts a search for countervailing trends generated by the contradictions inherent in the drivers of the agenda. Expanding the political space to support locally-owned transformative development practice may depend on identifying and exploiting such contradictions. Three inter-connected trends are identified (I hope readers will discern others). These are the political pressures to be seen to be in control in a world of uncertainty and surprises; the politics of accountability and the sector's internal dynamics in response to the changing environment of international aid.

5.1. The need to be seen to be in control

The desire for control, symptomatic of a refusal to engage with complexity in a dynamic and uncertain world, has created both elaborate performance measurement systems and an emphasis on quick deliverables – typified by bed-nets to tackle malaria. Such interventions demonstrate the aid agency is in charge and what is achieved can be attributed to the donor – important for the politics of accountability (5.2) – explaining the resurgence of popularity in branding.⁵⁸

Box 11. 'They challenged the standardization of the tools'

Once again, the tension between our bureaucratic strait-jacket and the needs of those we were helping became a challenge. Some members of the consortium objected to what was expected of them in the annual operating plan. They challenged the standardization of the tools and methodologies suggested by the facilitator team, which were meant to produce knowledge about the issue of citizenry in each community. They developed their own strategies to gather information and understand the processes they were following.

Source: Rosario León (Bolivia) www.thebrokeronline.eu/Series/Stories-from-Aidland/The-ghost-in-the-aid-machine

The rights-based approaches of the 1990s – and all that these meant with respect to the messiness of political voice and shifts in power relations - have been replaced by managerial control of people's lives. Interestingly the resistance of recipient governments and local NGOs to such control – documented in academic literature⁵⁹ – is disregarded in the promotion of the results-and-evidence agenda that, to paraphrase Mark Duffield, is a **performance of the will to govern**⁶⁰ to a domestic audience. Nevertheless, many mainstream development managers – and evaluators – are well aware this *is* a performance and thus in practice are more flexible than the public face of their agency might indicate (see Box 4).⁶¹ Sensible conversations with donors are sometimes possible and always worth trying. Furthermore, evaluators are becoming more open and aware of alternatives, enabling organisations to learn about how to support social change in complex and dynamic contexts.⁶² The problem is that much bureaucratic practice has not yet shifted, is oftentimes contradictory and - without challenge - may never

shift sufficiently to create the space for appropriate choice.

The **urge for control** may become pathological in the face of two uncomfortable realities worth tracking in terms of the opportunities they provide for supporting transformative action. The first is the irony that aid is increasingly likely to be funnelled to countries where chronic poverty is sustained by unstable political environments.⁶³ Rich countries' may desire to control these countries for security reasons but it is exactly here that results and evidence artefacts are least likely to be credible. Aid agencies will have to enhance their performance to domestic audiences or change their script. Secondly, the growing economic and political importance of middle-income countries - former aid recipients - adds to the uncertainty of official donor agencies and may temporarily contribute to exaggerated claims of controlling the circumstances in which they face competition from other donors not interested in results and evidence. The alternative take of the rising powers on development cooperation may encourage recipient governments to reject the evidence-and-results agenda demanded by conventional donors.

5.2. Value for money and the politics of accountability

“My top priority will be to secure maximum value for money in aid through greater transparency, rigorous independent evaluation and an unremitting focus on results.” DFID Minister Andrew Mitchell⁶⁴

Value for Money (VfM) is a normative idea conventionally concerned with achieving maximum economy, efficiency and effectiveness of resources. Publicly checking on the delivery of VfM is a performance⁶⁵ typically, for example in Parliamentary committee processes. The content of such performances along with commentaries in the media are important drivers of ideas about what is value for money and how to achieve it. These processes in turn give rise to the demand for certain types of information that privileges certain tools and methods – and the kinds of development programmes donors are prepared to fund.

“Why is all the focus at the planning stage when we all know they create perverse incentives (recent example from our work is women owning land kilometres away that they could get to only once a week, rather than having access to or power over communal land nearby due to target of ownership)?”

Laurie Adams http://bigpushforward.net/archiv_1#comment-21es/516/comment

With increasing need to demonstrate that aid budgets are delivering VfM during a period of austerity, ‘results’ and ‘evidence’ become part of the politics of accountability. Pressure is exerted, not so much by taxpayers as by Parliaments claiming to respond not only to taxpayers’ concerns about corruption and misuse of funds but more broadly to whether aid is delivering the best possible value for money. The Institute of Fiscal Studies, a British think tank, commented

‘Because the spending occurs elsewhere in the world, there is a relative lack of public [meaning the UK public] scrutiny of the budget’s effectiveness – voters can’t experience the effectiveness of aid spending in the way they can experience their local school, hospital or police force. This argues for an even greater degree of transparency and clarity about spending decisions and effectiveness than is seen in the rest of public spending’.⁶⁶

Box 12. Power and Politics Issues Relating to Value for Money

What aid is and who owns it. Is it an investment by donors or an entitlement?

What is it trying to achieve – transactional development or transformational change in power relations?

Whose values or standards are used to assess whether aid funded initiatives are effective and to whom should intermediaries primarily be responsible – donors or poor people?

How change happens – is it linear and predictable on the basis of past evidence or complex and emergent shaped by local contexts?

What can be known and understood about the causes of change – and the relative (un)importance of international development agencies in change processes.

Source: Cathy Shutt <http://bigpushforward.net/archives/1747#more-1747>

The results-and-evidence artefacts seem attractive ways to prove effectiveness to an audience that has little inside knowledge of the multi-stakeholder dynamics of even the simplest project. The drive for depoliticisation manifested in the evidence discourse connects to the privileging of accounting performance, which is viewed as technical and objective.

In discussing the **rise of accountancy** in the UK, McSweeney⁶⁷ notes the fantasy that accounting calculations can identify ‘the absolute truth’ and that accounting is thus able to transcend party-political distinctions. Self-evidently, once a particular discipline gains a discursive foothold in public-sector policy and practice it seeks to strengthen its position therein, generating a path dependency so that year by year its influence becomes more dominant. Large accountancy firms such as KPMG and PWC are involved in helping official development agencies develop management tools and approaches. They are also responsible for the selection and the monitoring and evaluation performance for some programmes, including NGO challenge funds such as the Dutch Government MDG3 fund and the DFID Girls Education Challenge programme (for which PWC has overall management responsibility), the latter with a payment-by-results pilot. In USAID, an increasing number of senior staff are no longer development professionals but management and compliance experts ‘such as accountants, lawyers, auditors, and procurement officers’.⁶⁸ The current Ministers for International Development in Sweden and the UK are both professional accountants.

The increasing influence of accountancy companies, is also related to **an ideology of the private sector** - that greater efficiency and effectiveness could be achieved by the public (and voluntary) sectors through the elements of New Public Management that imitate the private sector, such as explicit standards and measures of performance, control of outputs, competition, etc. Private-sector foundations and companies established to encourage charitable giving – particularly in the United States - have been instrumental in promoting the results agenda and its spurious homology with bottom-line profits.

On the other hand, at least outside the United States, there is more opportunity for pushing back as development practitioners learn the negative effects of using SMART performance indicators as targets – known in Economics as Goodhart’s Law, which states that “*when a measure becomes a target, it ceases*

to be a good measure". Their distortive effect - for example the police focusing on easy-to-solve crimes and hospitals treating easy-to-cure illnesses - is well known to the public in donor countries, and more could be done to draw the public's attention to how this is playing out in the funding of aid programmes, as well as forming coalitions for change with domestic policy actors.

"Should NGOs be openly challenging some of the assumptions and demands being made on them in return for funding and challenging the demands for targets, results based approaches, etc., as teachers, health staff and even the army have done in recent years?"

Tina Wallace <http://bigpushforward.net/archives/1759>

5.3. The sector's internal dynamics

*'If they write a cheque today, they want a result tomorrow.'*⁶⁹

The results-and-evidence agenda is technical and self-declaredly apolitical - although arguably highly political in terms of addressing an increasingly political problem of public support for aid. De-politicisation has enabled official agencies and NGOs to preserve their access to funds despite changes in the domestic political climate. In the UK (readers may know of instances from other countries), in the latter years of the Labour Government, officials were already elaborating an approach to results and evidence that distanced DFID from its early years under the radical Minister Clare Short. Her commitment to cross-government policy coherence for global poverty reduction was quietly dropped (as it has been in Sweden) and her stress on not distinguishing the British contribution to development from the collective effort was replaced with 'UK Aid', picked up and further emphasised by the incoming centre-right government. However, DFID's success in protecting the aid budget in a time of austerity has subjected it to jealous scrutiny from other Departments, along with intense parliamentary and media interest, obliging it to further exaggerate its claims to deliver VfM, as has been also the case in Australia.

Traditional aid agencies, both official and non-governmental, are also predicted to face increased competition in delivering social welfare from private foundations and individual charitable giving. Kharas and Rogerson argue that if social welfare programmes were financed and delivered more efficiently and effectively by private philanthropists and social impact investors (they do not say how this will be known, but presumably herein is the power of RCTs), then official aid agencies delivering such programmes will lose credibility with their citizens. These citizens, in turn, will learn they get a greater-poverty-reducing 'bang for the buck' through individual giving – and they will no longer be as supportive of paying taxes for official aid agency programmes.⁷⁰ This would then encourage official aid agencies to demonstrate they can equally deliver VfM, disregarding issues of local ownership and sustainability. However such a driver for reinforcing results and evidence discourses may be undermined in donor countries such as Denmark,⁷¹ where there is a strong emphasis on security and development. Here is an opportunity to push back against the results agenda by pointing to the contradiction between exercising control over bounded problems and the challenges of working in highly emergent and fragile contexts (where 30% of the current aid budget is spent),⁷² which requires a more flexible approach.

As noted above, a major driver of the results agenda is the competition for resources between official aid agencies and other government spending departments. The same driver is at work between multilateral agencies and INGOs competing for donor resources, as well as among think tanks like the one where I

work. In all instances, the competition makes organisations willing not only to comply with funders' management and monitoring requirements but sometimes through internalisation of power to support their funders' agendas wholeheartedly, for example in the case of development studies institutes by establishing impact evaluation centres or running postgraduate training courses on the topic. Likewise, the intense competition among NGOs is likely to increase their enthusiasm to demonstrate they are delivering VfM by presenting proposals for projects that can easily be measured.⁷³ Although within these NGOs others may be resisting or searching for alternatives.

More optimistically, the current rapid proliferation of artefacts could be an indicator of a dominant discourse under challenge in times of uncertainty and rapid change. Moreover, the greater the efforts of bureaucracy to exert control through an increase in mandatory procedures and over-lapping protocols, and the confusion over what is Value for Money or a Theory of Change, may create more room for manoeuvre and interpretation, posing opportunities, Cathy Shutt⁷⁴ argues, for 'radical reformers' interested in more participatory and democratic forms of aid management.

6. Conclusion

*A first and necessary step in counteracting the force of any discourse is to recognise ... its capacity to become hegemonic, 'to saturate our very consciousness, so that it becomes the ... world we see and interact with, and the commonsense interpretations we put on it..'*⁷⁵

I have sought to disentangle the historical threads and origins of results-based management and evidence-based policy/programming and have identified how, in which contexts and by whom they are used today, including their integration into the development sector and the change this had on the sector's priorities and practices. Armed with these insights, practitioners should be better placed to avoid power working invisibly to make us concur with what we know are inappropriate methods for designing and assessing programmes with multiple pathways of change.

Box 13. Conclusions from an experience of pushing back against the imposition of an artefact

Within a context of the desperate desire to get the grant that was generating blind compliance by the grantee's leadership and in a situation where power relations were blocking communications between donor and grantee, after a year of struggle, resistance within the grantee organisation succeeded due to:

- Solidarity and collective action
- Gaining support of allies within the donor organisation
- Appropriating the donor discourse and giving it an alternative meaning
- Seizing the opportunity offered by a surprising event.

Source: <http://bigpushforward.net/archives/1969#more-1969>

Organisations and individuals are not fully free but restricted by consequences of the actions of others. Nevertheless, the opportunities for individual and collective agency are manifold: the doctrines of evidence and results are not only challenged by the academy (as evidenced by the substantial body of literature on which this Framework Paper draws) but also many development practitioners who are finding room for manoeuvre to push back and create the space for alternative framings. These successes need to be reflected upon and shared.

The Big Push Forward believes that poverty reduction is not a technical problem but requires significant social change, and that this social change is, and must be, both political and locally led. Therefore, this requires alternatives to rigidly linear, project-based aid modalities that demand omniscience before they have even begun. However, this would suggest significant implications for external actors such as the development practitioners for whom this Paper has been written. We need to be self-aware to avoid disempowering others. It requires undertaking power analyses with ourselves factored in - as organisations and individuals who can make a positive or negative contribution, often inadvertently. It means engaging with a wider and more diverse group of policy actors in the state, civil society and the private sector; whenever possible, supporting debate, locally-driven problem solving, and independent research.⁷⁶

Endnotes

¹ See Lapsley, Irvine. 'New Public Management: The Cruellest Invention of the Human Spirit.' *ABACUS* 45, no. 1 (2009): 1-21; Hood, Christopher, and Guy Peters. 'The Middle Aging of New Public Management: Into the Age of Paradox?' *Journal of Public Administration Research and Theory* 14, no. 3 (2004): 267-82 ; Power, M. (1997) *The Audit Society: Rituals of Verification*. Oxford: Oxford University Press,

² www.participatorymethods.org

³ <http://cdi.mecon.gov.ar/docelec/az2249.pdf>

⁴ <http://www.oxfamblogs.org/fp2p/?p=13396>

⁵ Those interested in such a debate are recommended E.O. Wright's *Envisioning Real Utopias* (2010)

⁶ For a longer discussion of this point, see my paper 'Relationships matter. The best kept secret of international aid?' in *Investing in the Immaterial* at http://www.capacity.org/capacity/export/sites/capacity/documents/topic-readings/fa_cdra_digest.pdf

⁷ <http://bigpushforward.net/archives/1759>

⁸ www.powercube.net

⁹ See 'Murky Fields of Opposition' <http://bigpushforward.net/archives/1969>

¹⁰ Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972-1977*. Vintage. p. 155

¹¹ See for example <http://stuffexpataidworkerslike.com/2011/05/30/rcts/>

¹² Power, M. (2004). Counting, control and calculation: Reflections on measuring and management. *Human Relations*, 57(6), 765-783.

¹³ See the discussion on the BPF website

¹⁴ A, Natsios 'The Clash of the Counter-Bureaucracy' www.cgdev.org/content/publications/detail/1424271 p. 61

¹⁵ Although there are an increasing number of doctoral studies underway. For details contact Big Push Forward convenor, Brendan Whitty.

¹⁶ Mukhopadhyay, M. and Eyben, R. (2011) '*Rights and Resources. The Effects of External Financing on Organising for Women's Rights*' www.pathwaysofempowerment.org/Rights_and_Resources.pdf

¹⁷ Bevir, M., & Rhodes, R. A. (2008). Authors' Response: Politics as Cultural Practice. *Political Studies Review*, 6(2), 170-177.p. 170

¹⁸ http://ceas.iscte.pt/ethnografeast/papers/david_mosse.pdf

¹⁹ Anderson, B. (2006). *Imagined communities: Reflections on the origin and spread of nationalism*. Verso.

²⁰ Author's figure

²¹ Lambert, K. A. (1993). The taming of chance. *History of European Ideas*, 17(4)

²² Madaus, G.F., Ryan, J.P., Kellaghan, T., & Airasian, P.W. (1987). Payment by Results: An Analysis of a Nineteenth-Century Performance-Contracting Programme. *The Irish Journal of Education/Iris Eireannach an Oideachais*, 80-91.

²³ *ibid.*

²⁴ Espeland, W. N. (1997). Authority by the numbers: Porter on quantification, discretion, and the legitimation of expertise. *Law & Social Inquiry*, 22(4)

²⁵ C. Hood (1991) A Public Management for all seasons, *Public Administration* Vol. 69 (3-19).

²⁶ In 1994 Vice President Gore presented a report to President Clinton entitled *From Red Tape to Results: Creating a Government that Works Better and Costs Less*. Cited in Gray, A., and Jenkins, B. (1995) From public administration to public management: reassessing a revolution? *Public Administration* 73 Spring (75-99)

²⁷ Hood 1991, *op. cit.*; McSweeney, B. (1994). 'Management by accounting'. in (eds) A. Hopwood and P. Miller, *Accounting as social and institutional practice*, Cambridge, Cambridge University Press, p. 237.

²⁸ Power, M. (1997) *The Audit Society. Rituals of verification* Oxford, Oxford University Press

²⁹ Strathern, M. (2000) *The Audit Culture*

³⁰ Elton, L. (2004) Goodhart's Law and Performance Indicators in Higher Education, *Evaluation & Research in Education* 18:1-2 (120-128).

³¹ Blackman, T. (2001) 'Complexity theory and the new public management', *Journal of Social Issues* vol. 1, issue 2 p.1, www.whb.co.uk/socialissues/tb.htm

³² <http://www.demos.co.uk/publications/systemfailure2>

³³ For an excellent discussion of this point see <http://findwhatworks.wordpress.com/2013/01/16/pritchett-feedback-loops-and-the-accountability-conundrum/>

³⁴ Two recent official reports from the first quarter of 2013 are (1) *The Mid Staffordshire NHS Foundation Trust Inquiry, Independent Inquiry into care provided by Mid Staffordshire NHS Foundation Trust, January 2005 – March (2009)* Volume I Chaired by Robert Francis QC, HC375-I http://www.midstaffsinquiry.com/assets/docs/Inquiry_Report-Vol1.pdf; and

(2) *Southwark Sapphire Unit's local practices for the reporting and investigation of sexual offences, July 2008 – September 2009* <http://www.ipcc.gov.uk/news/Pages/IPCC-finds-failings-in-the-working-practices-of-Southwark-Sapphire-Unit-between-July-2008-and-September->

³⁵ <http://www.economist.com/blogs/freeexchange/2012/08/teacher-incentives>

-
- ³⁶ R. Batliner et al (2011) ‘A Primer on Results-Based Management’ SECO Economic Cooperation and Development’. www.seco-cooperation.admin.ch/themen/01100/index.html?
- ³⁷ www.dfid.gov.uk/Documents/publications1/how-to-guid-rev-log-fmwk.pdf
- ³⁸ <http://www.owen.org/blog/5483>
- ³⁹ www.dfid.gov.uk/Documents/pilots-payments-results.pdf
- ⁴⁰ www.norad.no/en/tools-and-publications/publications/evaluations/publication?key=393126
- ⁴¹ www.gpoba.org/gpoba/sites/gpoba.org/files/Docs/GPOBA%20Annual%20Report%202012.pdf
- ⁴² http://www.cgdev.org/section/initiatives/_active/codaid and associated critiques
<http://www.cgdev.org/doc/Cash%20on%20Delivery%20AID/Derenzio%20Woods.pdf>
- ⁴³ Claude Bernard, *An Introduction to the Experimental Study of Medicine* (1865).
- ⁴⁴ D. Sackeit et al. *BMJ* VOL 312, 13 JANuARY1996
- ⁴⁵ Carson 2002 http://www.ssps.unimelb.edu.au/__data/assets/pdf_file/0007/93553/evidence-based-policy-and-practice.pdf
- ⁴⁶ Solesbury *The Ascendancy of Evidence* cited in Carson op.cit.
- ⁴⁷ Black, *BMJ* Vol, 323 4 AUGUST 2001
- ⁴⁸ Biesta, G. (2007). Why “what works” won’t work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory*, 57(1), 1-22.
- ⁴⁹ T. Harrison 2000 ‘Urban policy. Addressing wicked problems’ in (eds) Nutley, S. M., Davies, H. T. O., & Smith, P. C. (2000). *What works?: Evidence-based policy and practice in public services*. The Policy Press., p. 209 <http://motahari.edu.umsu.ac.ir/ebmfile/What%20works%20Evidence%20based%20policy%20and%20practice%20n%20public%20services.pdf#page=224>
- ⁵⁰ <http://illinois.edu/lb/article/72/68987>
- ⁵¹ Picciotto, R. (2012). Experimentalism and development evaluation: Will the bubble burst?. *Evaluation*, 18(2), p, 221
- ⁵² T. Mitchell 1995 in J. Crush (ed) *The Power of Development* p. 130
- ⁵³ Picciotto op,cit p. 214
- ⁵⁴ http://transition.usaid.gov/policy/evalweb/USAID_Evaluation_Policy_FAQ.pdf
- ⁵⁵ Labrousse, A. (2010). Nouvelle économie du développement et essais cliniques randomisés: une mise en perspective d’un outil de preuve et de gouvernement. *Revue de la régulation. Capitalisme, institutions, pouvoirs*, (7).
- ⁵⁶ A. Du Toit http://www.plaas.org.za/sites/default/files/publications-pdf/PLAAS_WorkingPaper21dutoit_0.pdf

⁵⁷ Picciotto 2012 p.227

⁵⁸ <http://bigpushforward.net/archives/1700>

⁵⁹ E.g. Ebrahim, A. (2005). *NGOs and organizational change: Discourse, reporting, and learning*. Cambridge University Press.; Townsend, J. G., Porter, G., & Mawdsley, E. (2004). Creating spaces of resistance: development NGOs and their clients in Ghana, India and Mexico. *Antipode*, 36(5), 871-889.

⁶⁰ Duffield, M. (2002). Social reconstruction and the radicalization of development: aid as a relation of global liberal governance. *Development and change*, 33(5), 1049-1071.

⁶¹ See also <http://devpolicy.org/pushing-forward-for-more-accountability-and-less-theatre/>

⁶² See Ausaid's recent impact evaluation paper <http://www.ode.ausaid.gov.au/publications/documents/impact-evaluation-discussion-paper.pdf> and DFID's commissioned report by Nick Stern et al. <http://www.dfid.gov.uk/Documents/publications1/design-method-impact-eval.pdf> . See also www.betterevaluation.org

⁶³ H. Kharas and A. Rogerson 2012 'Horizon 2025: creative destruction in the aid industry' <http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7723.pdf>

⁶⁴ <http://www.dfid.gov.uk/news/latest-news/2010/andrew-mitchell-appointed-secretary-of-state/>

⁶⁵ Cathy Shutt 2011 'Confusion, Competition and Collaboration: Defining Value for Money' <http://aidconference.files.wordpress.com/2011/05/shutt-paper.pdf>

⁶⁶ www.ifs.org.uk/budgets/gb2012/12chap7.pdf, p. 161

⁶⁷ McSweeney, B. (1994). 10 Management by accounting. *Accounting as social and institutional practice*, 24, 237.

⁶⁸ Natsios *op.cit* p. 50

⁶⁹ Comment made at Big Push Forward meeting in New York, October 2012

⁷⁰ Kharas and Rogerson, *op.cit*.

⁷¹ Stepputat, F. (2012). Knowledge production in the security–development nexus: An ethnographic reflection. *Security Dialogue*, 43(5), 439-455.

⁷² Heppell, T., and Lightfoot, S. (2012). 'We will not balance the books on the backs of the poorest people in the world': Understanding Conservative Party Strategy on International Aid. *The Political Quarterly*, 83(1), 130-138.

⁷³ Shutt *op.cit*

⁷⁴ Shutt *op.cit*

⁷⁵ Apple, 1979, cited in B. Davies (2003) 'Death to Critique and Dissent? The Policies and Practices of New Managerialism and of 'Evidence-based Practice'. *Gender and Education* 15,1 p. 102.

⁷⁶ C. Roche and L. Kelly (2012) 'The Evaluation of Politics and the Politics of Evaluation' DLP
Background Paper no. 11

<http://www.dlprog.org/ftp/info/Public%20Folder/The%20evaluation%20of%20politics%20and%20the%20politics%20of%20evaluation.pdf.html>