MONITORING, EVALUATING & ASSESSING the IMPACT
of
GOVERNANCE PROGRAMMES

Summary of Literature and Practice Review

Cathy Shutt

2011

TABLE OF CONTENTS

INTRODUCTION

1. CARE International is a large, complex International NGO working at different levels in over 70 countries. Like other NGOs, it is looking to improve its effectiveness and accountability through enhancements to planning, monitoring and evaluation (M&E), and performance management systems.

2. Significant progress has already been made. Country programmes have adopted a programmatic approach. Detailed baseline studies support long-term (10-15 year) strategies and 'impact goals' defined for specific 'impact populations' that are defined as being of significant, often national, scale. Strategies are underpinned by 'theories of change' with associated chains of results.[1] Results are defined at three levels in each of the domains in which CARE works: inputs/activities/outputs; effects - defined as changes in behaviour and practices; and impacts - defined as equitable and durable improvements in human wellbeing and social justice.

3. The shift to a programmatic approach is enabling greater synergy between monitoring and evaluation (M&E) and learning. M&E is becoming a more analytical process to test theories of change and improve programmes.

4. This document summarises a brief review of literature and INGO practice related to governance M&E and impact assessment (IA). It is intended to inform the CARE UK's governance teams' future efforts to enhance monitoring, evaluation and impact assessment capabilities. More specifically, it summarises debates related to governance M&E and IA with the aim of stimulating discussion that will enable the selection and specification of appropriate governance indicators, as well as methodological approaches to measuring/tracking them. However it does not unpack in detail the issues arising from the challenge of trying to develop organisational systems that can respond to donor demands as well as organisational imperatives. These have been covered elsewhere.

5. CARE describes its governance theory of change in its Governance Programming Framework as: "*If citizens are empowered, if power-holders are effective, accountable and responsive, if spaces for negotiation are expanded, effective and inclusive, then sustainable and equitable development can be achieved. Change needs to take place and be sustained in all three domains to achieve this impact.*"

6. CARE outlines three domains of change for achieving this in its Governance Programming Framework:
   i. **Empowered citizens:** This domain is concerned with enabling the poor and marginalised, particularly women and girls, to be aware of their rights, to be capable of claiming their rights and to having a stronger voice to demand change. In short, the aim is to enable the poor to become active and empowered citizens.
   ii. **Accountable and effective public authorities[2] and other power holders** responsible for basic political, social and economic rights of their citizens.

---

[1] The term 'theories of change' is used here to relate to logics and assumptions that underpin projects and programmes.

[2] Public authorities are understood here as institutions (both formal and informal/modern and customary) that undertake core governance functions, including protection from external threats and management of external relations, peaceful resolution of internal conflicts; and providing and facilitating the provision of a range of collective goods and services - Evans, P (2010)

In this domain, CARE works directly with public authorities to improve their ability to fulfil their obligations and be more transparent, responsive and accountable, especially to its impact groups – women and other poor and marginalised populations.

iii. **Expanded and inclusive spaces[3] for negotiation** This domain is closely related to the previous two domains and results from their interaction. A focus on interactions between public authorities and citizens avoids the oversimplification of 'supply' and 'demand' sides of governance, and emphasises that the focus should be on the relationship of rights and responsibilities (i.e. a social contract) between public authorities and citizens in all its different forms.

7. The paper is organised as follows. It begins with a reflection on methods and goes on to explore issues emerging from current debates taking place within the development community about approaches to M&E and impact assessment within a results-based paradigm. Some of these are then unpacked with specific reference to literature on assessing and learning from governance programmes. This section discusses domains of change and possible indicator themes, as well as various approaches and methods that may be used for M&E and IA. It briefly covers debates about differences between objectives and subjective indicators. Considerations of distinctions between quantitative and qualitative indicators lead into a discussion of possibilities offered by participatory methods that can produce quantitative data. An analysis of INGO experience in monitoring and evaluating governance work follows. The concluding section focuses on possible implications for CARE arising from the report's findings.

REFLECTION ON METHODS

8. This review began with a search of published and grey literature covering several broad areas:
    - Academic and practitioner work related to monitoring and evaluation and impact assessment;
    - Academic articles and existing literature reviews on the monitoring, evaluation and impact assessment of governance projects and programmes;
    - Debates related to donors' use of governance indicators;
    - A variety of tools and guidelines related to governance M&E produced by the United Nations Development Programme (UNDP)'s Governance Centre and DFID;
    - Governance strategies of INGOs, e.g. Christian Aid, ActionAid and Oxfam GB (that are explored in more detail in another document).

9. Initial reading revealed considerable literature on M&E debates, governance assessment and indicators. Some of this (discussed below) is relevant to CARE's governance programming and M&E. However, the general picture emerging from the literature is that monitoring and evaluation of various aspects of governance work

[3] As a way of addressing the perceived "democratic deficit" of representative democracies, more direct forms of citizen engagement in governance have opened up new arenas for public participation, often referred to using the concept of "space". The literature often refers to two distinct types of space: invited spaces (provided by the government/state, where civil society has a voice and vote) and popular spaces (where CSOs come together at their own initiative to channel unrecognised demands, protest, provide services, solidarity etc.). The boundaries between invited and popular spaces are flexible. Cornwall, A (2004), "Introduction: New democratic spaces? The politics and dynamics of Institutionalised Participation – IDS Bulletin Nr 35.

has been somewhat weak.[4] A number of studies of the impact and effectiveness of citizen empowerment and transparency and accountability initiatives in various sectors that are instructive exist.[5] Yet there is a paucity of documented experience of governance M&E practice, particularly the use of indicators, by comparable organisations.[6] Thus the literature review was complemented by informal interviews and exchanges with practitioners working in comparable INGOs.[7] However, the information gathered during the course of these dialogues is not exhaustive. As the approaches used by NGOs are subject to change due to shifts in the funding environment, those reported here should be read as a means to raise issues that may need further discussion and exploration by CARE.

10. The initial terms of reference for the review focused on an indicator analysis as it assumed:
    - there was a one to one relationship between indicator themes and their attributive; quantitative/ qualitative; objective/ subjective properties
    - these relationships would be observable and could be analysed from examples of best practice.

11. Neither of these assumptions was found to hold in practice:
    - Indicators are defined as clues, symptoms and general aids to identifying (and often understanding) change. Indicators for a given change can be articulated using various units of analysis and therefore can be conceived and measured in a number of different ways.
    - There is no established 'best practice' relating to governance M&E.

    Thus, this report does not include a detailed indicator analysis per se. Instead it identifies a number of indicator themes and unpacks issues emerging in contemporary debates that may influence decisions about their specification, as well as IA and M&E approaches to measuring them.

ISSUES EMERGING FROM THE LITERATURE

MONITORING, EVALUATION AND IMPACT ASSESSMENT DEBATES

12. The age of austerity and pressure to demonstrate 'results' and 'value for money' (VFM) is contributing to lively debates about methodological approaches to evaluation and impact assessment.[8]

Debates concerning M&E, accountability and learning

13. Much of the literature about NGOs and monitoring and evaluation highlights dilemmas experienced by practitioners trying find approaches that on the one hand will meet donor demands to demonstrate results, and on the other be empowering

[4] O'Neill, T. Foresti, M. and Hudson A. (2007) Evaluation of Citizen's Voice and Accountability: Review of the Literature and Donor Approaches Report, Overseas Development Institute; Sarles, M. Evaluating the impact and effectiveness of USAID's democracy and governance programmes
[5] McGee, R. and Gaventa, J. (2010) 'Review of Impact and Effectiveness of Transparency and Accountability Initiatives ', Institute of Development Studies, Brighton provide a useful critical discussion of methods used to assess transparency and accountability initiatives
[6] GSDRC (2008) Helpdesk research report: monitoring and evaluation of participation in governance (07.05.10) *available at* *www.gsdrc.org*;
[7] See Annex 1 for further details
[8] e.g. see Eyben, R. and Guijt, I. (2011) Big Pushback Concept Note; Lucas, R. and Longhurst R. Lucas, H. and Longhurst, R. (2010) 'Evaluation: Why, for Whom and How?', *IDS Bulletin* 41.6, Brighton: Blackwell Publishing; Ravallion, Martin (2009) 'Should the *Randomistas* Rule?' *Economists' Voice*. Berkeley Electronic Press

for 'partners' and/or 'beneficiaries', enabling learning that will enable greater accountability to them.[9]

14. Although Results-based management (RBM) practices can encompass participatory methods, they are often criticised for reporting requirements that prioritise accountability to donors, encouraging INGOs to take credit for complex changes and impacts in ways that do injustice to their local partners and poor people.[10] The logical framework matrix (LFM), a management tool that epitomises and perpetuates the inequity that characterises relationships between international and local society actors, also reduces their effectiveness.[11] Its hierarchical logic, which is used inconsistently by donors and practitioners, can cause confusion that encourages regressive, disciplinary learning about rules and reporting, at the cost of progressive learning about social change.[12] Emphasis on quantitative output indicators can encourage practitioners to develop narrow activity-based understandings of the purpose of their work. This makes it difficult for them to perceive their roles in relation to the broader change strategies being pursued by their organisations.[13]

15. Aspects of this debate are about power: who decides what changes should count and how to assess or measure them? Who collects data and how? Who analyses and interprets it? Who learns during the process? How does this all influence answers to the questions: accountability for whom and for what?

16. Authors like Chambers argue that values and understandings about the nature of reality, knowledge and change that characterise scientific positivist, results-based paradigms contradict those underpinning a participatory worldview.[14] Results based management approaches are often criticised for suggesting change is linear and has predictable results that can be established by research conducted by outside experts.[15]Tools used within results based paradigms use reductionist approaches that tend to focus on isolated variables that inadequately reflect multiple contextual inter-related causes and power relations that influence people's wellbeing.[16]

17. There is growing acknowledgement that such evaluation methods are proving ill suited to governance[17] agriculture[18] and health[19] interventions involving complex

---

[9] Ebrahim (2005) NGOs learning and change; Wallace, T., Bornstein, L., & Chapman, J.(2006) *The Aid Chain: Coercion and Commitment in Development NGOs,* Intermediate Technology Development Group;
[10] e.g. Wallace, T., Bornstein, L., & Chapman, J. (2006)
[11] Hudock, A. (1999) *NGOs and Civil Society: Democracy by Proxy*, Cambridge, Polity Press; Fowler, A. (1998) 'Authentic NGDO partnerships in the new policy agenda for international aid: dead end or light ahead', *Development and Change* 29: 137-159;
[12] Shutt, C. (2006) 'Money matters in aid relationships', in (ed) R. Eyben *Relationships for Aid,* London: Earthscan, Eyben, R. Relationships for Aid
[13] Ebrahim (2005); Shutt, C. (2008) Power and Aid Relations: A Study of Cambodian NGOs and their International Funders', unpublished thesis, Brighton: IDS
[14] E.g. see Chambers, R. (2007) 'Who Counts? The Quiet Revolution of Participation and Numbers', *IDS Working Paper* 296, Brighton
Chambers, R. (2010) A Revolution Whose Time Has Come? The Win-Win of Quantitative Participatory Approaches and Methods, IDS Bulletin *Special Issue: People-centred M&E: Aligning Incentives So Agriculture Does More to Reduce Hunger*, Vol 41, Issue 6 p44-55
[15] See http://aidontheedge.info/2011/01/31/why-the-results-agenda-doesnt-need-results-and-what-to-do-about-it/ for further discussions about the problems with results based management practices
[16] Chambers (2010) 'Paradigms, Poverty and Adaptive Pluralism', *IDS Working Paper* 344, Brighton: IDS
[17] McGee and Gaventa (2010)
[18] Haddad, L. Lindstrom, J. and Pinto, Y. (2010) The sorry state of M&E in Agriculture: can people centred approaches help? IDS Bulletin *Special Issue: People-centred M&E: Aligning Incentives So Agriculture Does More to Reduce Hunger*, Vol 41, Issue 6
[19] Lucas and Longhurst (2010)

social factors that lead to unpredictable emergent change and effects. According to Andrew Natsios, former head of USAID, "*those development programs that are most precisely and easily measured are the least transformational, and those programs that are most transformational are the least measurable.*" [20]

Pros and cons of different approaches to evaluation and IA

18. Increased pressure to measure and demonstrate the impact and value for money of complex emergent programmes is stimulating more nuanced discussion about how different values and understandings of change influence approaches to evaluation and impact assessment.

19. The popularity of Random Control Trials (RCTs), an experimental approach that appeared to emerge as the outright winner in 'causal wars' about assessing impact, is waning for both ethical and methodological reasons. Seeking to establish attributable impact using indicators to compare changes in a group that is randomly assigned to an intervention and 'a counterfactual' – a group that is not – is considered ethically unsound.[21] Moreover, according to Ravallion (2009), a lead researcher in the World Bank, who sees evaluation as a process to produce knowledge that can enhance development, RCTs provide little information about how interventions impact different groups within potential impact populations. He also notes they overlook spill over effects the project intervention might have on the 'counterfactual' group. For example a local government office aware of a donor funded RCT intervention in one district might reduce its budget there and allocate it to other districts instead. And approaches that are perceived to be of value to poor people and local actors spread spontaneously as has been the case for community led total sanitation approaches in Africa.[22]

20. Furthermore as RCTs focus on identifying and attributing change rather than explaining it, they are a-theoretical and have low external validity being unable to ascertain if effects of interventions in one location can be applied in other contexts. As a single method therefore, RCTs alone are unable to generate knowledge and learning that can be used for scaling up or replication.[23]

21. Given the methodological and ethical limitations (above) it is perhaps unsurprising that members of the donor community expect to see less emphasis on RCTs in the future.[24] The burgeoning literature on evaluation and IA in aid funded development appears to increasingly advocate for pluralist approaches and methods determined by particular evaluation purposes and contexts.[25]

22. Lucas and Longhurst's (2010) aptly entitled review of evaluation debates *'Evaluation: Why, for Whom and How?* draws attention to the contingent nature of evaluation approaches. They note approaches are shaped by the specific purposes of different users. Honing in on a typology of three major evaluation approaches: RCTs, theory of change and realistic evaluation, they explain why some approaches are better for accountability, while others learning. Noting the limitations of RCTs as a single approach, the authors argue some would likely favour its use in combination with results-based methods using outcome and impact indicators to test 'theories of

[20] Natsios, A. (2010) 'The Clash of the Counter-bureaucracy and Development', Center for Global Development
[21] Ravallion, M. (2009)
[22] Plan International has been involved in triggering CLTS in communities in Ethiopia and Kenya who have then gone to promote the approach in other communities.
[23] Ravallion, M. (2009)
[24] e.g. Jo Abbot the Deputy Head of DFID's Civil Society programme at DFID. She intimated she expected less emphasis on RCTs in the future at a BOND Meeting Feb 3rd, DFID TOR on impact assessment
[25] McGee and Gaventa (2010), Lucas and Longhurst (2010)

change' (TOC). This more pluralist methodological approach would explore not only *whether* an approach is working but also *how* and perhaps *why* with the aim of being able to establish attribution or contribution. The authors draw attention to differences between proponents of the 'theory of change' approach and those who favour 'realistic evaluation'.

23. Realistic evaluators tend to proceed from assumptions that change is complex and emergent and though their evaluations are theory based, they are arguably less committed to testing an *a priori theory* of change to establish if an approach has worked. Their evaluations may instead be guided by broad research questions that seek to first identify changes and then design methods to better understand mechanisms - how and why change happened. From this perspective it may be considered irrational and costly to seek to identify models unlikely to be replicable elsewhere. It is systematic learning they are interested in – exploring the contextual reasons why various groups are impacted by the same intervention in different ways with the aim of enhancing accountability by using learning to make interventions more effective. According to Lucas and Longhurst, their learning is driven by a desire to develop programme theories that may help explain the effects of interventions on specific groups in particular contexts.

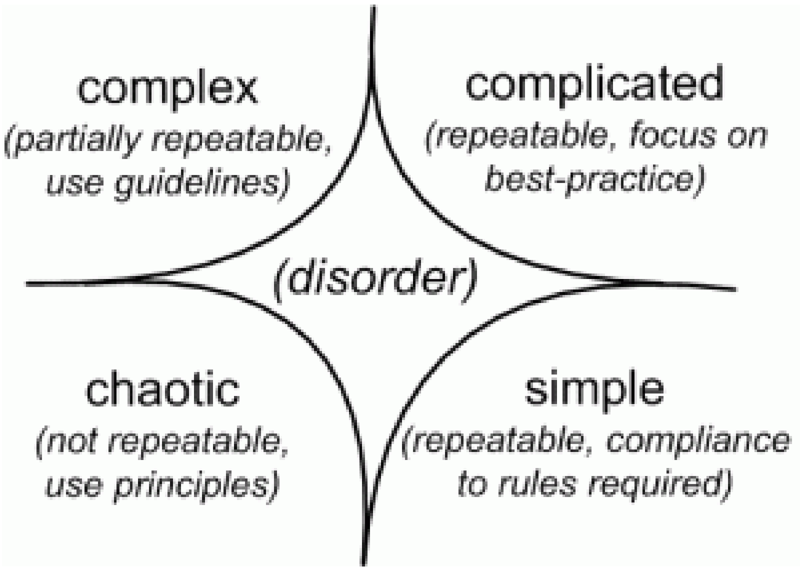
24. Another variant favoured by those who find complexity science of value to understanding aid interventions is 'developmental' or 'real-time' evaluation. Developmental evaluation draws together elements of TOC and realistic evaluation but goes further in a) adopting participatory evaluation principles that seek to empower those involved without seeking attribution and b) applying a complexity lens to distinguish between different types of change contexts using the categories: simple, complicated, complex and chaotic.[26]

Potential advantages of a 'developmental evaluation' approach?

25. Guijt (2010) shows how a complexity lens encourages questions about what can realistically be expected of evaluation and IA in contexts with different characteristics. In her paper she seeks to unpack links between developmental evaluation, learning and accountability:
    - *Simple cause effect relationships* such as can be found in a recipe enable easy accountability and the identification of 'best practice' that can be replicated through compliance.
    - Accountability can still be outcome oriented in *complicated systems* involving multiple actors and relationships, but only if M&E allows systematic learning as change is less predictable and therefore more can potentially go wrong. Although notions of 'best practice' can be applied to complicated systems, some might argue that it is only possible to identify and therefore try to replicate 'good practice'.
    - Change in *complex systems* is even more unpredictable and cause effect can only be understood retrospectively. Therefore while some elements of change might be understandable in ways that can be repeated with guidelines, others cannot. It is here where learning may be more about trying to make sense of what is happening in what essentially is 'emergent practice' in order to adjust interventions to make them more accountable.
    - In *chaotic systems* however, it is impossible to identify cause effect relationships. Efforts to assess patterns may enable some learning to improve

[26] e.g. see Guijt. I Accountability and Learning – exploding the myth of incompatibility between accountability and leaning in eds *dited by Jan Ubels, Naa-Aku Acquaye-Baddoo and Alan Fowler.*Capacity Development in Practice; Ramalingham, B. (2008) Evaluation and complexity, presentation from a Norad conference available at http://www.outcomemapping.ca/resource/resource.php?id=183

interventions, but this is best conceived as 'novel practice' unlikely to be repeated. At best learning can contribute to principle based accountability in terms of exploring whether actors did as well as could be reasonably be expected under rapidly changing circumstances.



D. Snowden Cynefin framework from Cognitive Edge[27]

26. Of course the categories above are not essentially distinct in practice, and the extent to which learning is thought to be able to contribute to replication in complex systems appears somewhat contested. Nevertheless Ramalingham (2008) includes a useful heuristic summary of key differences between complexity or developmental and more traditional results based approaches to evaluation.

Traditional	**Complexity-oriented**
Measure success against predetermined goals	Develop new measures and monitoring mechanisms as goals emerge & evolve
Render definitive judgments of success or failure	Provide feedback, generate learning, support direction or affirm changes in direction
Aim to produce generalisable findings across time & space	Aim to produce context-specific understandings that inform ongoing innovation
Creates fear of failure	Supports hunger for learning

Where is the IA and M&E debate now?

27. Methodological discussions about how to respond to donor demands for evidence of results for programmes taking place in complex systems are no longer merely quantitative/qualitative or 'economists versus the rest' debates.[28] They are providing

[27] Accessed from http://weblog.tomgraves.org/index.php/2010/02/21/chaos-and-cynefin/ March 19 2011
[28] This analysis comes from Henry Lucas who is working on evaluation methodologies at IDS and can be seen in a recent terms of reference for a DFID assessment on approaches to impact evaluation.

fresh opportunities to explore the complementarity and commensurability of elements of different evaluation and IA approaches described above, as well as how to conceptualise impact and accountability.

28. Drawing from recent reports that illustrate the ineffectiveness of traditional approaches to M&E and IA as means to enhance results, Jacobs (2010) argues it is better to conceive and measure NGO performance and accountability in terms of contributions to local institutions and citizen's efforts than ultimate developmental and democratic impacts.[29] For those that support this view the challenge becomes one of adopting tools and processes used in social accountability programmes to encourage feedback on their performance in relationships. Advocates of feedback loops position themselves between logical framework enthusiasts and those that favour participatory monitoring and evaluation.[30] The feedback approach includes some 'top down' metrics to evaluate NGOs performance according to organisational missions, together with more participatory evaluation tools that allow 'partners' to decide what counts, such as the widely cited citizen scorecard method developed by CARE in Malawi.

29. Such participatory approaches long criticised for being 'unscientific' or 'too qualitative' appear of growing interest to sceptics and donor organisations concerned with more traditional notions of accountability and impact. This is largely due to the increasing number of case studies that demonstrate participatory approaches can produce numbers that are at least as reliable as those created by more expensive quantitative methods.[31] Proponents of participatory approaches hope this will encourage wider adoption of evaluation and IA processes that empower 'beneficiaries' while encouraging learning and enhancements that can make interventions more accountable to them.

30. However, fundamental differences in opinion between people from various disciplinary backgrounds, with different values and varied beliefs about the existence of truth/value-free knowledge and thus the nature of 'evidence' mean not all differences can be easily resolved. An individual believing that knowledge is socially constructed and produced by processes shaped by power relations is unlikely to be comfortable with positivist approaches that claim to demonstrate objective truths. Similarly developmental evaluation approaches that start from a systems perspective that focuses on inter-relationships between variables are never going to be able to accommodate methodologies that aim to 'isolate variables'.[32] Differences in values will shape opinions about who should decide what counts, choose indicators, collect data, perform analysis and communicate results.

31. It is commonly accepted that there is a need to move beyond M&E and IA debates about data collection methods and focus more on what values, understandings of the nature of knowledge and the politics of interpretation mean for evaluation practices. Groups within the Institute of Development Studies are currently proposing a project that will support such discussions between academics from different disciplinary backgrounds.

[29] http://ngoperformance.org/performance/ and Jacobs, A. (2010) Creating the missing feedback loop, IDS bulletin Vol 41 No 6, IDS Brighton
[30] see Jacobs, A., Barnett, C, & Ponsford, R. (2010) Three Approaches to Monitoring: Feedback Systems, Participatory Monitoring and Evaluation and Logical Frameworks, IDS Bulletin Vol 41, Issue 6

[31] Examples of these can be found in Chambers (2007) and Chambers (2010)
[32] Danny Burns personal communication Feb 2011

EVALUATING AND ASSESING THE IMPACT OF GOVERNANCE PROGRAMMES

32. The broad debates on monitoring and evaluation and impact assessment discussed above are mirrored in those concerned with enhancing M&E and learning from democratic governance programmes.

33. Critiques of instrumental motives for upward accountability (mentioned above) are reflected in discussions about donor motives for assessing governance. Although some reports suggest that donors' primary aims are to improve programming and policy dialogue,[33] others contend that donor desires to 'measure governance' and demonstrate effectiveness are fundamentally misguided and contrary to the spirit of harmonisation enshrined in the Paris Declaration.[34] How relevant such debates will be in the new era of 'development effectiveness' and value for money that seems to be driving agencies towards individualistic efforts to demonstrate and attribute results to their interventions remains to be seen.[35]

What domains of change are of interest and what indicators should be used?

34. Whatever their motives, there now exist a plethora of international governance indicators that have been developed for a wide range of purposes, exploring various dimensions of governance change.

35. The emergence of such indicators has implications for the evaluation and impact assessment of long-term governance programmes in terms of their availability and suitability as secondary sources of data. They can also be viewed as a useful aid to inform programming theories of change and the design of systems for the evaluation and impact assessment of more discrete projects involving smaller populations.

36. Holland and Thirkell's (2009) *Measuring Change in Voice and Accountability Work* (pages 36 to 49)[36] and UNDP's *Sources for Democratic Governance Indicators*[37] are two particularly useful publications that summarise the focus of and methods used to measure various indicators. The meta level indicators discussed in these publications are produced in very different ways – some rely on key informants, others large scale surveys that use random sampling techniques to explore citizens' subjective perceptions of different aspects of governance e.g. the Afro Latin and Asia Barometers. Some are composite and include what are commonly thought of as 'objective' indicators e.g. counts of criminal prosecutions etc. Others, such as the IDEAS State of Democracy (SOD) index, use slightly more participatory processes to determine exactly which indicators should be used to review and reform aspects of local and national government in different context. SOD areas of interest include: the rule of law; access to justice; civil and political rights; economic and social rights; citizenship; elections; etc.[38]

---

[33] OECD DAC, 2008, 'Survey of Donor Approaches to Governance Assessment', Organisation for Economic Cooperation and Development - Development Assistance Committee, Paris

[34] Irish Aid Advisory Board (Booth, David & Verena Fritz) (2008) Final Synthesis Report: Good governance, aid modalities and poverty reduction: from better theory to better practice *Research Project (RP-05-GG) of the Advisory Board for Irish Aid* available from www.odi.org.uk,

[35] Draws from Rosalind Eyben's observations of the shift in discourse from aid to development effectiveness.

[36] Holland, J. and Thirkell, A. (2009), *Measuring Change and Results in Voice and Accountability Work*, DFID Working Paper 34

[37] (UNDP undated) http://www.undp.org/oslocentre/docs04/Indicator%20Sources.pdf

[38] Institute of Democracy and Electoral Assistance Assessment framework http://www.idea.int/sod/framework/index.cfm

37. While some of these datasets appear to have potential utility as secondary data for long- term programmes, it is impossible to assess their suitability without reference to specific empirical examples. Although some are only available in limited geographic areas, the situation is changing rapidly as they spread. However even if they are available, they will only be useful if they are: updated in line with programme timelines; disaggregated for specific demographic groups of interest; and specified in units of analysis that fit with the broader monitoring and evaluation approaches used by the systems of which they are a part.

How should indicators be specified, defined and used?

The term indicator is sometimes associated with linear notions of change and targeted outcomes. While some indicators are used in this way, a broader understanding of indicators being a signal of change makes them intrinsic to any evaluation or impact assessment. Although they can be targeted, they can also be developed retrospectively through emergent participatory practice or inductive desk based methods such as those used by Gaventa and Barrett (2010). The literature on governance M&E and IA, e.g. Dahl-Østergaard, T., Schulz, K, & Svedberg B. (2008) and McGee and Gaventa (201O) gives the impression that indicator 'selection' and specification has tended to be retrospective for the purpose of particular studies.

Overall the literature suggests that indicator choices take place in complex or chaotic unpredictable systems. Final decisions are inherently contingent and require an iterative process that includes engagement with discussions about evaluation methods discussed earlier. They can be aided by reference to the criteria for good indicators adapted from UNAIDs (2008) below.

Criteria for a good indicator:

- Is needed and meaningful
- Has substantial merit – tracks significant change
- Has been tested and works
- Is feasible – there are resources and capabilities required to collect data and analyse it
- Is consistent with and adds meaning to overall indicator sets – the combination of indicators found in programmes are mutually supportive
- If quantitative - is defined in metrics that are consistent with the overall system of which it is a part
- Is fully specified:
  - Has a clear purpose and rationale
  - Qualitative aspects as well as quantitative aspects are well defined
  - Is identified with clear methods to collect data that allow it to be measured
  - Specifies how frequently data should be collected
  - Is disaggregated so that changes for particular groups of interest can be tracked
  - Includes guidelines on how to interpret what changes mean to local people and communicate contextual notions of change in terms or units that can be understood by 'outsiders'

Adapted from UNAIDs 2008 cited in Holland and Thirkell (2009)

38. Some broad, common themes manifested in international indicators that can guide choices of domains of change for governance programmes and play a role in indicator choices are summarised in the table below. It should be noted that they are not necessarily mutually exclusive. Process indicators related to deepening the culture of participatory governance found in CARE's expanded spaces for negotiation domain can also indicate citizen empowerment and enhanced state effectiveness outcomes.

CARE's Domains of Change	**Broad domains of change that shape commonly used governance indicators**
Empowered Citizens	• Nature and scope of civil and political knowledge • Forms/nature of participation in civil society and governance • Voter turn out

	• Preferences for/attitudes to democracy, power, reform, • Capacities for collective action • Number/range of civil society organisations, movements, networks and alliances • The degree of linkages between various CSOs • Sense of freedom, empowerment, security • Social cohesion
State Effectiveness (Representative, Democratic, Responsive and Accountable)	• Free and fair elections • Democratic role of political parties • Citizen confidence and trust in institutions at various levels • System responsiveness at various levels especially (as a result of actions of democratically elected representatives) • Availability of budget and other information at various levels • Legal frameworks, institutions and access to justice • Civilian control of military and police • Anti corruption legislation mechanisms and rule of law • Existence and implementation of right to information law • Respect for free speech and association • Dispute resolution • Decentralisation: decisions taken at appropriate levels of government • Delivery of quality services and functions • Budget allocations for issues of importance to citizens • (Access to) free media • Right to citizenship • Citizens able to fully participate in public life • Equal access to services and resources • Equal access to civil and political rights • Equal access to economic social and cultural rights • Equality of access to political office (women's quotas)
Expanded spaces for negotiation	• Greater state -citizen engagement in all phases of governance cycles: advocacy policy and budget formulation, budget policy monitoring, • New spaces for citizens to engage and participate in public affairs • New issues discussed in spaces • Proposals put forward by citizens enacted by government

39. Whether a particular change and indicator will be considered evidence of an output, outcome or impact level change is likely to be contingent on the particular organisations and individuals involved in designing programmes and M&E systems as is made evident by discussions below.

What do complex governance change pathways mean for attribution?

40. Debates about the validity of some of the international governance indicators raise questions about how the outputs, outcomes and impacts of governance programmes should be conceived as well as measured. These are most easily discussed with reference to an example of a voice and accountability results based chain, such as that described in Holland and Thirkell (2009), which assumes impact is expressed in terms of developmental changes in people's lives.

Input	Process	Output (Capacity)	V&A Output (Behavioural change)	Outcome	Impact
Resources	Capacity-building Network building	Capacity to engage	Direct collective/ individual civic engagement with the state	Access to resources or service provision	Improved well being

41. Aspects of arguments about how the impact of democratic governance programmes should be conceived are ideological or political as illustrated in a paper by John Gaventa and Greg Barrett (2010). Based on a rigorous, inductive meta-evaluation of case studies, they show that citizen engagement can have both positive and negative impacts on people's lives. Within their analysis they argue it is a mistake to perceive changes such as increased voice and empowerment merely as instrumental means to MDG type outcomes as described in the chain above. In other words, the paper makes a case for viewing the *impacts* of active citizenship approaches in terms of people exercising the right to participate, becoming politically empowered and achieving dignity.[39]

42. Other debates about how to conceive impact are more methodological, but it is impossible to untangle them from notions of accountability and attribution. Most simply put they force questions about who should be accountable for what types of changes in complex systems involving multifarious power relationships between different actors.

43. Several strands of literature emphasise that there is little evidence that many of assumptions underlying theories of change underpinning governance programmes hold in practice. Kaufmann and Kraay, the principle architects of the World Bank Institute's 6 KKZ composite indicators use examples of corruption legislation to illustrate the complex links between policy decisions and policy outcomes.[40] They argue that 'objective' *du jure* policy decisions related to corruption bear little relation to the *de facto* experiences and perceptions of governance outcomes on the ground. A number of other authors also challenge commonly held assumptions about links between citizen voice and government responsiveness and accountability,[41] transparency and accountability[42] and between democratic and development outcomes.[43]

44. Such debates suggest the above results chain, which does not unpack assumptions about state behaviour, would prove an inadequate theory of change for impact

[39] Gaventa, J. and Barrett G. (2010) So what Difference Does it Make?: Mapping the outcomes of citizen engagement, Working Paper DRC, IDS
[40] Kaufmann, D. and Kraay, A.(2007a) Governance Indicators, Where are We and Where Should we Be Going?, World Bank Working Policy Research Paper
[41] Rocha Menocal, A & Sharma, B. (2008) *Joint Evaluation of Citizens' Voice and Accountability*: synthesis report – DFID, London
[42] McGee, R. and Gaventa, J. (2010)
[43] e.g. Goetz, Anne Marie (2005) Social justice, gender justice, and local government *Background and concept note for: Assessing the social justice outcomes of participatory local governance, a project of the Ford Foundation Local Governance and Learning group (LOGO)*
Houtzager, P., Acharya, A. & Gurza Lavalle A.(2007) Associations and the exercise of citizenship in new democracies: evidence from Sao Paulo and Mexico City *IDS Working Paper 285* IDS, Brighton; Logolink (2008) Draft research reading guide *unpublished*; Manor, J.(2004) Democratisation with inclusion: political reforms and people's empowerment at the grassroots *Journal of Human Development* 5/1, pp 5-29; Mosse, D. (2004) Power relations and poverty reduction *in Power, Rights, and Poverty: Concepts and Connections, a working meeting sponsored by DFID and the World Bank March 23-24 2004,* edited by Ruth Alsop,

assessment design. Holland and Thirkell (2009) actually do go much further and usefully unpack other sample results based chains and map them onto various dimensions of DFID's governance framework. These illustrate how complex and governance change pathways can get and the contingent nature of outcomes and impacts. For example one suggests that voice and accountability can be thought of as an impact of increased state institutional capability that creates an environment in which people can exercise civil rights.

45. Given governance changes are so complex and ill understood Kaufmann and Kraay advise against putting too much emphasis on easy to measure *de jure* quantitative indicators such as policy changes as they are unreliable proxies for developmental impacts defined in terms of changes in people's lives.[44] And when it comes to assessing the developmental impacts of programmes that aim to influence state capacity or responsiveness the World Bank's term 'most likely association' may be the best recourse.[45] Several donors recommend using multiple indicators and methods to assess governance change. As Munck argues, undue attention to tracking quantitative data may divert attention from qualitative information that better informs understanding about processes of change that lead to developmental or democratic outcomes.[46]

46. Given the complex systems in which governance interventions take place, it is not surprising that conceptualisations of democracy and governance impact are also contingent in evaluation practice. In a meta-evaluation of Sida programmes, for example, evaluators describe the impact of a programme to enhance cooperation between the South African Police Service and the Swedish National Police Board in terms of "good awareness of human rights amongst the police both related to victims and detainees." [47] Other indicators of impact for this programme included "reduced response time", "improved suitability of new recruits," and improved equity and representation of different ethnic groups overall in the police force. In Vietnam the impact of a programme designed to enhance the capacity of a government institution to perform its function and produce statistics was defined as just that without reference to how statistics may influence people's lives. And for a democracy and human rights programme in Nicaragua, the evaluators describe impact in terms of "developing platforms for dialogue between the Nicaraguan government and civil society".

47. Related, in their evaluation of the impact of five donor-led voice and accountability initiatives Rocha Menocal and Sharma (2008) conclude that donor expectations of voice and accountability initiatives contributing to poverty alleviation and MD goals are unreasonably high. They thus question the 'absolutist' interpretations of impact implied in Holland and Thirkell's chain, suggesting impacts might better be conceived in terms of what Holland et al categorise as outputs - changes in behaviour and practice of public officials and policy actors. This appears an argument in favour of limiting conceptualizations of impact within short scale programmes to changes that take place in less complex parts of systems where there is greater possibility of proving an intervention has contributed to change.

[44] Kaufmann, D. and Kraay, A.(2007a)
[45] O'Neill, T. Foresti, M. and Hudson A. (2007)
[46] Munck, Gerardo L (2008) *Measuring democracy: a bridge between scholarship and politics* – Johns Hopkins University Press, Baltimore Md.,
[47] Dahl-Østergaard, T., Schulz, K, & Svedberg B. (2008) Experiences and Lessons Learnt from Sida's Work with Human Rights and Democratic Governance, Department for Democracy and Social Development SIDA Evaluation 2008:29

48. A recent terms of reference on Impact Assessment issued by DFID suggests the complexity of governance M&E and IA challenges is well understood, but the search for pluralistic approaches to impact attribution continues. The TOR acknowledges that RCTs or quasi-experimental approaches are not suitable or possible for many programmes that aim to strengthen democracy and accountability; accountable and responsive government; security and conflict prevention; combating gender based violence; citizen empowerment and community action. Thus its aim is to seek a repertoire of approaches that that may be more suitable for establishing attribution.[48]

What approaches and methods can be used to for governance IA and evaluation?

49. Methodological challenges and political pressures are unsurprisingly creating both implicit and explicit demands for pluralist approaches to the evaluation and IA of governance programmes and the use of indicators within them. Debates suggest the possible benefit of using a complexity lens to consider which elements of programme systems can be considered simple enough to facilitate the identification of causal links necessary to establish attribution. Such analysis could also reveal where and when it might be beneficial to use theory of change approaches versus realistic approaches etc.

50. A common theme emerging from discussions about exploring and/or measuring change in governance programmes is the important role that context variables play in change processes, particularly political cultures and informal embedded power relations.[49] And Gaventa and Barrett (2010) emphasise context may not always influence changes in expected ways. This cautions against adopting M&E or IA approaches and indicators that encourage country comparisons and/or simplistic assumptions regarding possibilities for replicating promising programmes in very different locations.[50]

51. There is a growing literature that notes the challenges of applying results based logical framework approaches for governance interventions and proposes a number of alternatives:
    - In their research on voice and accountability evaluations O'Neill et al (2008) argue that theory of change, or realistic approaches to evaluation discussed earlier are more suitable for governance programmes than traditional results based management methods.
    - Gaventa and Barrett (2010) use a novel 'realistic' inductive method to produce quantitative data retrospectively from qualitative case studies.
    - In a specific paper on evaluating policy influence around a theory of change approach, Jones (2011) outlines a number of outcomes/domains of change that can be used to guide various approaches to evaluating the effects of influencing programmes. He also suggests a variety of methods and tools that can be applied to the task of 'measurement'/assessment.[51]

---

48 DFID (2011) Impact evaluation terms of reference
49 e.g. see McGee and Gaventa, J. (2010)
50 Goetz, A.(2005)
Robinson, M (2004) Resources, citizen engagement and democratic local governance, a background paper *unpublished*
51 Jones, H. (2011) A guide to monitoring and evaluating policy influence, Overseas Development Institute Background Notes, ODI

Influencing approach	Outcomes; what to measure	How; tools
Evidence and advice	Outputs	Evaluating research reports, policy briefs and websites
	Uptake and use	Logs; new areas for citation analysis; user surveys
	Influence	RAPID outcome assessment; Episode studies; Most Significant Change
Public campaigns and advocacy	Target audience attitudes, behaviour, etc	Surveys, focus groups, direct responses
	Media attention	Media tracking logs, media assessment
	Media framing and influence	Framing analysis; coverage
Lobbying approaches	Actors; relationships; policy process and institutions	Recording meetings; tracking people; interviewing key informants; probing influence

- A considerable number of domains of change, indicators and tools that can be used to evaluate and assess local governance effectiveness using participatory approaches are summarised and analysed in the UNDP (2009) Measuring Local Governance Guide.[52]
- McGee and Gaventa's (2010) summary of approaches and methods that could be used to assess transparency and accountability initiatives has broad relevance to other aspects of democratic governance programmes.

Approach and *example of its use with TAIs*	**Good for…**	**Less good for….**
Experimental approaches e.g. RCTs *Service delivery: Random testing of demand-led Vs top-down interventions in education in Madagascar, Nguyen and Lassibille (2008)*	-Isolating impact of a particular intervention -'Before'/'after'- or 'with'/'without'-type comparisons, via use of control group -High internal validity (i.e. permits conclusions about causal relationships) -Measuring, counting -Eliminates selection bias, through random approach -Most immediately applicable to service delivery TAIs where 'with'/'without' contrasts are most tangible	-Capturing the unexpected or unforeseen -Recording what actually happened, as opposed to ascertaining the achievement of the specific change expected -Explaining nuances of causality or change processes; can miss complex interactions of multiple variables - Capturing spill over effects - External validity (i.e. RCTs do not produce findings that allow valid generalisations to be made from this case to other research contexts) -Assessing contribution in light of other contributors, as opposed to affirming attribution -Ethics: possibly unethical to

[52]UNDP (2009) Measuring Local Governance Guide, UNDP Democratic Governance Group, UNDP Oslo Governance Centre, Norway

		involve 'untreated' control group -Less useful when TAIs' expected outcomes are intangible or non-material - Exploring different ways trials impact different populations - Empowerment and learning to enable more accountability to poor people
Quantitative survey *Service delivery:* *Assessment of impact of citizen report cards on Bangalore public sector performance (Ravindra 2004)*	-Drawing generalisable conclusions on basis of representative sample (i.e. high external validity) -Perceived as objective -Can generate numbers, which can be more persuasive than prose for some audiences -Measuring how far the TAI has been implemented (its effectiveness), or measuring impact as proxied by indicator(s) - e.g., % of people benefitting	-Explaining the degree of implementation attained, or understanding what made the impact happen -Capturing what is not easily quantifiable or countable -Sampling can be difficult in some environments - Empowerment and enabling learning that will enhance accountability to poor people
Qualitative case studies *e.g. Aid transparency: Assessment of workings of World Bank Inspection Panel (Clark et al 2003)*	-Purposive sampling (e.g. to focus on successful cases or failures) - Can be used longitudinally within a real-time evaluation framework and provide before /after or – Alternatively can be retrospective to trace back from a given outcome (positive or negative) the various factors and dynamics that gave rise to it, including complex contexts (i.e. high internal validity) -Unpacking underlying theory of change -Can be used inductively, adapting analytical focus as case unfolds -Detailed explanation of one case – but can be conducted within comparative or multiple framework, to enhance explanatory power and external validity - Looking at effects on different groups within the population - Can incorporate participatory methods and or quantitative surveys etc	-Comparability over time or with other cases; regrettably few multiple-case comparative studies are available -Drawing general or representative conclusions -Knowing significance of findings, beyond specific cases
Qualitative Stakeholder interviews *e.g. Natural resource governance: Evaluations of EITI (Rainbow Insight 2009)*	-Capturing positioned viewpoints of differently-placed stakeholders in multi-stakeholder initiatives -Easily combining with direct verification or observation methods	-Time-intensive and generates copious qualitative data
Official Indices and rankings	-At-a-glance comparative information -Mobilising the low scorers into	-Explaining reasons or contexts behind scores or rankings - Being understood in ways that

e.g. Budgets: Open Budget Survey of International Budget Partnership (www.openbudgetindex.org/)	action through peer- or reputational pressure or shame	mean they are appropriately used
Participatory approaches *e.g. Freedom of Information: 'People's Assessment' of progress of India's Right to Information law (Raag 2009) Care Malawi's citizen scorecards*	-Encompassing different 'indicators' and perceptions, especially those of users/participants/intended beneficiaries -Building stakeholders' ownership and participation in the initiative as a whole - Empowerment and learning that can immediately contribute to outcomes - Enhancing 'downward 'accountability - Facilitating power analysis	-Replicating across many or diverse contexts -Deriving quick, yes/no answers - Representativeness if not used with random sampling techniques and approaches that mitigate power dynamics affecting data produced in groups
Outcome mapping *A general accountability and transparency programme: AcT (Accountability in Tanzania) http://www.accountability.or.tz/home/*	-Detecting, and understanding changes in behaviours, relationships and/or activities of people and organizations -Tracing emergent change, including unforeseen aspects, in complex contexts -Actively engaging stakeholders in learning - Embracing participatory axiology and worldview - Can incorporate participatory methods and thus be empowering and enhance downward accountability -Facilitating power analysis	-Demonstrating initiative's direct contribution to development impacts -Demonstrating fulfilment of prescribed goals -Producing generalisable findings
'Most Significant Change' approach *Anti-corruption DFID GTF programme by Transparency International (Burge 2010)*	- Participatory activity involving stakeholders in deciding what change or 'impact' is worth analysing, and what value to assign to it -Recognising and capturing complexity and unexpected dimensions -Generating learning for participants - Embracing participatory axiology and worldview - Can be empowering and enhance downward accountability - Could include power analysis	-Producing generalisable findings -Generating upward accountability to funders over pre-specified outcomes -Time-consuming, resource-intensive, single-initiative focus - Representativeness if not used with random sampling techniques and approaches that mitigate power dynamics affecting data produced in groups
Other story- or narrative-based methods that can incorporate participatory approaches such as most significant change	-Exploring, describing and making sense of processes involving many actors and multiple relationships - Exploring how and why change does or does not happen - Encouraging critical thinking by being provocative - Exploring power - Communicating results in creative	-Producing generalisable findings -Time-consuming, resource-intensive, single-initiative focus - Finding simple solutions - Fitting normative requirements

e.g. Action Aid's Critical Stories of Change	and engaging ways - Enabling learning to enhance practice and changing behaviour -Embracing complexity including unexpected results - Exploring effects of interventions on different groups - Including participatory methods	

Adapted from McGee and Gaventa (2010)

52. It should be noted that elements of the approaches are not mutually exclusive, and current efforts to find better ways to measure impact are sometimes producing somewhat unexpected efforts to couple elements of approaches. For example the Governance Teams at IDS is currently exploring whether participatory methods can be used within random control trials. Whether the approach will be judged to be commensurable with proponents of a participatory paradigm remains to be seen.

53. One finding emerging from the IETA study is that there is little evidence to suggest that participatory approaches, such as citizen scorecards, which are often inherent parts of social accountability projects, are being used for the evaluation of their outcomes. Emerging donor interest in participatory evaluation and impact assessment suggests there may growing opportunities to incorporate such methods in different evaluation approaches.[53] Decisions about adopting participatory methodologies require engaging with debates about differences between objective and subjective measures of change, as well as the pros and cons of quantitative versus qualitative approaches to data collection.

The relevance of debates about objective & subjective indicators

54. Particularly useful are ideas from Kaufman and Kraay's defence of the World Bank Governance indicators that begin to unpack unhelpful dichotomies that influence attitudes toward to indicator choices and data collection methods.[54]

55. Kaufman and Kraay argue that complex links between *de jure* and *de facto* changes mean that evaluation of governance is all about subjective perception. They also implicitly reject the value free nature of knowledge and argue that few indicators are truly objective. D*e jure* indicators, often assumed to be 'objective', can involve some level of subjective selection by decision makers. Perhaps surprising is their claim that the margins of error in the measurement of perception aspects of the World Bank indicators are less than those for 'objective' measures.[55,56]

56. The Afro, Asia and Latin America Barometers that are based on citizens' perceptions of government performance have added enormous credibility to the use of

[53] DFID and Sida are both exploring the potential of participatory evaluation and impact assessment.
[54] Kaufmann, D. and Kraay, A. (2007) On measuring governance framing issues for debate
[55] Kaufmann, D. and Kraay, A.,(2007) *Governance Indicators, Where are We and Where Should we Be Going?,* World Bank Working Policy Research Paper
[56] It should be noted that data on facts such as income level etc are often collected through methods that rely on subjective recall. Elements of subjectivity can also be introduced in categorisation, ranking etc and analysis.

perception-based data to assess changes in different aspects of governance programmes.[57]

The relevance of debates about quantitative versus qualitative indicators?

57. New debates are providing space to explore the relative strengths and weaknesses of using quantitative and qualitative data when specifying indicators and deciding data collection methods.

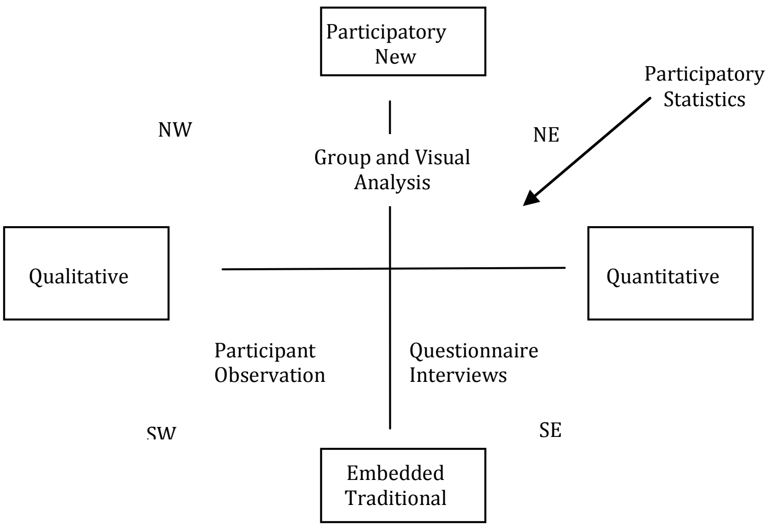
Figure 3.2 Qualitative and quantitative dimensions of poverty appraisal

More 'qualitative' research	More 'quantitative' research
1. Non-numerical information	Numerical information
2. Specific (contextual) population coverage	General (non-contextual) population coverage
3. Active population involvement	Passive population involvement
4. Inductive inference methodology	Deductive inference methodology
5. Broad social sciences disciplinary framework	Neo-classical economics (and natural sciences) disciplinary framework

Source: Gabarino, S & Holland, J, (2009) Adapted from Kanbur (2003, 1)

58. The table above, usefully suggests that while there may be differences between quantitative and qualitative approaches it is better to think of them in terms of different points along a continuum, rather than being essentially distinct. Although quantitative survey methods are often associated with disciplines that favour deductive methods, evidence that qualitative data can be successfully coded, ranked and quantified to produce statistics appears to be opening up interest in more participatory and inductive approaches to producing numbers.

59. Robert Chambers (2010) argues that as participatory approaches produce reliable statistics they provide win-win solutions, enabling learning and empowerment that can facilitate accountability to donors and poor people alike. The diagram below illustrates where participatory statistics fit conceptually with other data collection methods.



[57] Holland forthcoming

Adapted from Chambers, R (2008). Revolutions in Development Inquiry, by Jeremy Holland

Possibilities for participatory evaluation and impact assessment of governance work

60. Traditionally participatory approaches have struggled to gain widespread acceptance by donors and policy makers. Some of the reasons are reflected in Chambers' (2007) analysis of the opportunities and trade offs when comparing traditional approaches with participatory evaluation methods:
    - Standardised, closed and commensurable versus participatory processes that open, diverse and empowering: The more standardised the process, the more extractive and less empowering and accommodating of local priorities and realities it is likely to be. The less standardised it is, the harder the outcomes are to analyse.
    - Scale, quality, time, resources, and ethics: Smaller scale participatory processes take more time, and more resources, but can allow for higher quality analysis linked to deeper local ownership and impact but loss of "precision of inference"; and vice versa.
    - Quality of facilitation versus speed, scale and cost of implementation: To achieve good facilitation of participatory processes requires time and resources devoted to careful selection of facilitators, their training and then their supervision in the field. This may add to costs and slow implementation and limit its scale, even if the outcomes are still highly cost-effective compared with alternatives.
    - Sampling for representativeness versus ease, spontaneity and participatory self-selection of convening groups: Care in selection, in judging size of group, and observation and facilitation of participatory processes can offset these dangers but takes time and effort and can entail a loss of spontaneity.

61. The trade offs require important consideration when choosing evaluation approaches and data collection methods, but a growing number of examples illustrate that with adequate planning some of perceived disadvantages of participatory processes can be overcome:
    - *Standardisation and commensurability*: In Malawi and Uganda participatory approaches have been used to evaluate outcomes and assess impact at scale using rigorous representative samples that enabled inference of population estimates. This involved working in larger number of areas than is usual in participatory studies and the incorporation of statistical principles to ensure tools could meet demands of standardisation and comparability of data produced across sites. Although statisticians argued such approaches could be used to inform national policy and become part of national statistical systems they faced a number of institutional constraints, which are arguably more related to mindsets than possibilities.[58]

Using participatory methods to assess the impact of health policy on the rural poor.

Uganda Participatory Poverty Assessment Process (UPPAP) revealed the significance of health shocks on the vulnerability of the poor that had become a major election issue. In March 2001, the President of Uganda scrapped user fees for Government health units. This decision was also informed by finding that it did not raise much revenue.[59]

Time series data on outpatient attendances later showed that the combined effects of abolishing fees and increasing the supply of health services increased outpatient attendance by 84% between 2000/01 and 2002/03. Immunisation rates among children also increased

[58] Holland forthcoming

[59] Yeates et al 2008 cited Holland and Thirkell (2009)

dramatically (e.g. the proportion of children who received their third DPT immunisation increased from 48% to 84% between 2000/2001 and 2002/03).

To confirm that the pick up in outpatient attendance was progressively weighted towards poorer households, the WHO/MoH conducted **participatory research** with carefully-sampled communities. A **wealth ranking** exercise conducted by villagers revealed that since the abolition of user fees, the poorest quartile had consistently used Government health centres more than any other group: in 2002 the poorest quartile used these facilities at a rate of 0.99 visits per person per year whereas for the wealthiest group the rate was 0.77. There was only limited data available for the period preceding the abolition of fees, but the average utilisation rate for the months of January and February 2001 was 0.52 and 0.42 respectively. Subsequent analysis of the 1999/2000 and the 2002/3 **household surveys** confirmed that poorer income groups had increased their utilisation of services more than richer groups.

From Holland forthcoming

- *Scale:* another example of the rigorous use of participatory approaches at scale is that of a land rights social movement in Bangladesh that generated its own set of quantitative indicators for monitoring and evaluating empowerment. They mapped these indicators onto a matrix, as a baseline for annual participatory M&E. Participants assigned a happy face to those indicators that had been achieved and an unhappy face to those that had not been achieved, in the process generating learning that prompted an action plan for making better progress on those indicators in future. Data from individual groups was amenable to aggregation and uses by outsiders described as "secondary stakeholders".[60]
- *Aggregation:* Several examples of successful aggregation from focus groups are discussed in Chambers (2008:p 114), which include studies in urban violence by Caroline Moser and others in Jamaica and Latin America.
- *Representativeness:* Both of the above participatory monitoring and evaluation examples have used sampling techniques that have been judged as being as rigorous and at least, if not more, reliable than traditional alternatives by statisticians at Reading university.[61]

62. As Chambers (2010b) argues, participatory methods can be used to quantify almost any qualitative dimension of change open to human judgement and have often been found to deliver data that is more accurate than surveys due to the triangulation opportunities provided by group activities. Concerns about some data being too sensitive to discuss in groups has largely proved unfounded. In instances where this is a concern anonymising techniques, such as asking people to use identical pens to write or draw responses to sensitive questions on screwed up pieces of paper that are passed to a facilitator, can be used.[62]

Participatory Methods for Generating and Measuring Quantitative Indicators:

Counting: examples are social and census maps that tend to be useful for households and headcounts or anything which is common knowledge.

Calculating amounts: examples come from triggering processes for Community-Led Total Sanitation where local people calculate the quantities (e.g. cartloads for the whole community) of shit produced by their households in a day, multiply out for longer periods, and add up for

[60] Jupp D et al, 2010. "Measuring Empowerment? Ask Them: Quantifying qualitative outcomes from people's own analysis. Insights for results-based management from the experience of a social movement in Bangladesh", Sida Studies in Evaluation 2010:1
[61] Holland, J. forthcoming
[62] Robert Chambers (March 2011) personal communication

the whole community, concluding sometimes with community cartloads per annum. **Participatory measuring** can be found with timber stocks, water flows, crop yields, arm circumferences, and land use areas from participatory GIS mapping and modelling. **Estimating and comparing** –can be used to illustrate changes or trends over time and/or to give an idea of proportional values attributed to various material assets; social relationships; elements of wellbeing; different government services; government institutions etc. The ten seed technique (Jayakaran 2002) is a common method in which ten seeds are allocated between two or more options to illustrate how relative changes or preferences.[63] Alternatively it can involve the allocation of 100 seeds, stones, or other counters to give percentages. Although usually better if performed in groups where there is more opportunity for triangulation, the methods could be adapted to accommodate individuals. The methods are versatile and can be used to explore and generate quantitative data on attitudes, emotions, intentions, and client satisfaction. Other methods often used to establish relative values include scoring or preference ranking, matrix ranking and matrix scoring.

Adapted from Chambers (2007) which includes more references to specific examples.

63. Essential for participatory evaluation and IA initiatives that aim to simultaneously contribute to citizen empowerment as well as learning to enhance programme accountability is allowing citizens to choose their own indicators. A particularly useful resource in this regard is the UNDP publication 'Measuring Democratic Governance: A framework for selecting gender sensitive and pro-poor indicators. This publication emphasises the importance of enabling women and particularly excluded groups to select indicators that represent changes in areas of government accountability and responsiveness that are of particular concern to them.

64. Striking in literature reviewed is frequent mention of community scorecards, often associated with CARE Malawi that can be used to evaluate changes in perceptions of the quality of public service delivery. These ideas have been taken forward in work with PACT that has sought to apply scorecard principles to the development of local governance barometers.[64] Such approaches encourage communities to get together with providers and discuss the quality of services and demand changes, thus proving a mechanism that can be both integral to achieving programme outcomes while also measuring them. Community scorecards could be adapted and fit well within theory of change, realistic and development approaches to evaluation. As they can integrate evaluative processes as a means to achieve outcomes with feedback loops that may provide means to enhance CSO accountability, it can be argued that they may provide particularly cost effective solutions.[65]

65. Citizen's scorecard are sometimes confused with a more top down approach to feedback: citizen's report cards. Differences are compared in the table below.

[63] Jajayaran 2002 (2002) The Ten Seed Technique, World Vision, China. Download at www.fao.org/participation/Ten-Seed%20Technique-Revised.pdf (accessed 7 November 2007)

[64] Bloom, E. Sunseri, A., & Leonard, A. (2007) Measuring and Strengthening Local Governance Capacity the Local Governance Barometer, PACT and USAID

[65] Roche, C. (2009)

Table 1: Key Differences between Citizen Report Cards & Community Scorecards

The Citizen Report Card	The Community Scorecard
• Unit of analysis is the household/individual • Information collected via a survey questionnaire • Relies on formal stratified random sampling to ensure that the data is representative of the underlying population • The major output is the actual perceptions assessment of services in the form of the report card • The media plays the major role in generating awareness and disseminating information • Conducted at a more macro level (city, state or even national) • More useful in urban settings • Time horizon for implementation is long (about 3-6 months) • Intermediary plays a large role in conducting the survey and data analysis • Feedback to providers and the government is at a later stage after media advocacy	• Unit of analysis is the community • Information collected via focus group interactions • Involves no explicit sampling. Instead the aim is to ensure maximum participation of the local community in the gathering of information. • Emphasis here is less on the actual scorecard and more on achieving immediate response and joint decision-making • Relies more heavily on grass-roots mobilization to create awareness and invoke participation • Conducted at a micro/local level (village cluster, and set of facilities) • More useful in rural settings • Time horizon for implementation is short (about 3-6 weeks) • Role of intermediary is mostly as facilitator of the exercise • Feedback to providers is almost immediate and changes are arrived at through mutual dialogue during the interface meeting

Source Ackerman (2005)

Risks of abuse of quantitative data

66. It is not only methods for producing numbers that need to be considered, so do risks of them being abused. In an piece that explores the psychology of numbers, Pollitt (2008) notes that figures produced through complex governance assessment processes, particularly those used to establish composite indicators, are often poorly understood by people other than those who produce them.[66]

67. Poorly understood figures pose certain risks. An OECD report on the "use and abuse of governance indicators" draws on Kaufmann's own observations and provides examples of misinterpretations of indicators such as Transparency International's index, which is sometimes understood as an absolute measure of a country's corruption instead of a relative ranking.[67] This suggests that tools used for data collection need careful consideration and the potential pros of different approaches to quantifying - even in participatory research- carefully considered.

68. Moreover, Pollitt points out that numbers can be given more influence than they merit because of the peculiar power they command.[68] They are often assumed to be objective, even though most processes used to aggregate figures involve margins of error as well as an element of subjective classification bias. Despite Kaufmann and Kraay's efforts to stress the methodological limitations of the KKZ's many authors still question their validity, appropriateness and the transparency of methods used for their construction.[69]

69. Dialogues between Kaufmann and Kraay and their detractors make evident that any governance indicator is likely to contain errors and must be fully specified and explained with assumptions related to the theories of change underpinning its use,

[66] Pollitt, C. (2008) 'Moderations in all things': Governance quality and performance measurement, paper for the structure and organization of government ASOG meeting Gottenburg
[67] Arndt, C. & Oman, C. (2006) *The Uses and Abuses of Governance Indicators,* Organisation for Economic Coperation and Development
[68] Pollitt, C. (2008)
[69] Thomas, M. (2009) *What do the Worldwide Governance Indicators Measure,* European Journal of Development Research

and estimates of margins of error being made transparent.[70] Their points are well made and contain useful lessons for participatory practitioners who despite emphasis on reflexivity, are sometimes accused of failing to adequately expound on methodological limitations such as those outlined in the tradeoffs described earlier.[71]

[70] e.g. see Thomas's critique of the World Bank's Governance Indicators that questions whether they are measuring the right things, and argues that a lack of transparency around their construction makes that difficult to know.
[71] E.g. by Cooke, B. and Kothari, U (eds) (2001) The case for participation as tyranny, in *Participation the new tyranny,* Zed Books London.

BINGOs' APPROACHES TO GOVERNANCE M&E

EXAMPLES OF COMMON GOVERNANCE INDICATORS USED by NGOs(Some are under specified)

Type of change	**Indicator theme (the categorisation against specific themes is contestable)**
Input	• Spend £
Citizen Capacity	• Knowledge/understanding of rights • Ability to organize organisations and events
Government Capacity	• Knowledge about citizens' rights • Ability to facilitate meetings/engage with citizens
Citizenship/ Empowered citizens	• Number of CBO organizations formed • Number of federations formed • Number of CBOs strengthened • Number of CSOs functioning/working together with media research institutes on advocacy issues • Number of awareness raising activities conducted by marginalised groups • Number of service delivery initiatives implemented by young people • Active engagement of stakeholders, confidence and willingness to get involved in the future • Number of CSOs using accountability tools or processes in relations with citizens • Number of lobbying activities conducted by marginalised people
Citizenship changes – in terms of Northern advocacy	• Number (and quality) of lobbying activities conducted by groups in the North • Number of media pieces on related issues instigated by supporters in the North • Number (and quality) of links between people/organizations here and in the South
Negotiated Development (State – Citizen engagement in processes to enhance accountability and responsiveness(	• Number of people who have had contact with elected representatives • Levels of participation of different kinds of stakeholders in invited spaces • Number (and quality) of meetings between marginalized groups and relevant authorities • Numbers (and quality) of participation of specific groups, e.g. women, ethnic minorities, in meetings or accountability processes (sometimes articulated as numbers of people other times as in numbers of organizations or locations) • Number of people or groups monitoring government plans • Number of proposals presented to government authorities • Level and quality of citizen monitoring and oversight • Level and quality of monitoring evidence generated e.g. number of exchanges with media, no on radio slots, number of meetings with service providers on access quality and relevance of services, • Number (and quality) of local governance mechanisms in which citizens are engaged • Government and MPs' perceptions of quality of interactions with CSO coalitions
State Effectiveness: Responsiveness &	• Dissemination of information by authorities • Communities report responsiveness in terms of budget allocations/resources responses to specific claims

Type of change	Indicator theme (the categorisation against specific themes is contestable)
Accountability ‘Outcomes’ (State behaviour)	• Number of policies or plans revised that support the rights of citizens • Number of sources and/or amounts of devolved funding available to citizens • Number of proposals or plans approved or incorporated into policy • Transparency and access to information • Equality and access to decision making • Level of citizen trust in state institutions • Government adopt participatory budgeting M&E systems (put forward by partner) • Mechanisms/spaces to facilitate citizen engagement institutionalised by government • Government adopt policy as a result of CSO advocacy • Citizens’ perceptions of quality of state interactions • Increase in number of partners/coalitions who say they benefit from exchanges with government • Number of service delivery mechanisms established and/or strengthened to respond to the needs and rights of young children
Access to services and rights	• Budget commitments (especially for) marginal groups • Access to services or material benefits for (especially for) marginal groups • Access to/realization of political, civil, economic and social rights
Well being change in terms of Midis	• See MDG indicators and other theme
Well being change in terms of citizen empowerment?	• Changes in their life as a result of participation in social accountability initiatives • Transformation of power through new relationships, access to new networks

CONCLUSIONS AND IMPLICATIONS for CARE

70. Emerging debates related to the challenges of evaluating outcomes and assessing the impact of aid programmes, particularly those related to governance are extremely relevant to challenges facing CARE. Decisions about how to evaluate and assess the impact of CARE's governance work are taking place in a complex if not chaotic unpredictable aid system! Challenges are far broader than merely choosing indicators and are shared by other comparable INGOs.

71. The literature indicates the relevance of CARE's theory based approaches to conceiving evaluation as a means to test and enhance understanding of complex governance change pathways. However, pieces that emphasise the contingent nature of governance changes, often taking place in complex systems, raise questions about the extent to which it may be possible to identify models that can be replicated or scaled up elsewhere. Drawing on developmental evaluation literature may help to establish which aspects of change can be thought of as existing within simple/complicated systems where there may be scope for 'attribution' and replication versus those likely to prove too complex and chaotic.

72. In practice CARE will need to use a variety of approaches to evaluate and assess the impact of its work. The literature suggests the challenge is too complex to enable easy identification of best practice M&E. As recommended by SIDA and the World Bank, assessing democratic governance requires the use of quantitative, qualitative, objective and subjective indicators.[72] The suitability of indicators and approaches can only be established with reference to particular project and programme contexts in which some of the criteria adapted from UNAIDS good indicator criteria can be applied.

73. It might be interesting to explore the extent to which the Governance Action Research Initiative, which is proving successful in terms of promoting learning that can enhance effectiveness and value for money might be able to accommodate quantitative participatory methods that would meet donor requirements for quantified results in ways that make a persuasive case for provable contribution.

74. Given the pressure to produce numbers and CARE's commitment to participatory processes, seeking ways to incorporate community scorecards, a recognised method already associated with CARE, might prove a productive way to quantify perceptions related to state effectiveness and expanded spaces for negotiation. There may exist opportunities to couple this with elements of the 'Measuring Empowerment Index' framework from Bangladesh (currently being tested by CARE US) to explore changes in the Empowered Citizens domain.

75. Where possible M&E and IA methods should endeavour to explore positive & negative outcomes as well as intended and unintended consequences of programmes, using power analysis to help to identify factors that facilitate or hinder 'good change'.

76. Any efforts to pilot participatory numbers at scale should be pursued in collaboration with individuals qualified to develop research/sampling designs that will be considered sufficiently rigorous by donors and academics. The advice of those with experience producing participatory numbers using methods and units of analysis that allow standardisation, aggregation and comparison should also be sought. Careful consideration of how various power relations might affect data produced in processes

---

[72] Kaufmann and Krayy (2007) SADEV (2006)

such as those described in discussions of tradeoffs of participatory numbers (above) is vital.[73]

77. Although challenging, it may be practically possible to develop methods to produce numbers that can be compared across locations and countries. But given the risks of abuse attempting such comparison is probably best avoided unless part of meaningful qualitative country case studies that explore how historical political, economic, social, cultural factors influence change processes in various locations.

78. Using broad governance meta-level indicators as a basis for more contextually defined country level indicators appears relatively common. However, the 'problematic' of how to approach it in CARE may be particularly challenging because of the sometimes competing demands of ambitious organisational programmatic M&E imperatives reflected in the shift to a programme approach and donor reporting requirements. Some specific areas of concern are likely to be:

- .
- *Finding ways to use donor project or programme level M&E to 'prove' its contribution to longer term programme change.* It is generally accepted that links between governance programmes and developmental impacts are hard to prove, but there is increasing pressure to try and attribute or prove contribution. Thus in order to make a convincing argument for its contribution to national level impact, CARE must ensure project level M&E approaches are rigorous and commensurable with secondary data used to establish programme level impact.
- *Availability of secondary data that will meet the above criteria.* It still remains to be seen whether secondary data that fits the above criteria will be available for impact populations defined by CARE and reported at times that coincide with CARE's programme cycles.
- *Defining and measuring baselines:* Plan and Christian Aid's experiences suggest that if there is intention to measure change against baselines, indicators need to be clearly defined and carefully articulated to enable comparison. One of the challenges that may face CARE is whether baselines conducted during the process of developing programme strategies will be considered acceptable by donors deciding to provide financial support several years later.

79. CARE's Governance Programming Framework (GPF) stresses the importance of CS accountability. This includes a focus on the organisation's accountability. Exploring ways to adapt and expand the use of citizen scorecards appears a promising means to further this aim.[74] Such considerations should include possible value for money advantages resulting from making participatory processes, which are integral to programme theories of change, evaluation tools. They could simultaneously enhance accountability to partners and donors. Given CARE's brand association with community scorecards and the local governance barometer, such approaches could be marketed to exploit this apparent niche.

80. Recent briefing notes from ODI and INTRAC as well as methods used by Oxfam to evaluate its Climate Change campaign may provide useful ideas about how best to approach the evaluation and IA of aid effectiveness advocacy work.[75]

[73] UNDP, 2007, 'Governance Indicators: A Users Guide', 2nd Edition, United Nations Development Programme includes some useful tips about this as do several chapters in Cooke, B. and Kothari, U (eds) (2001) The case for participation as tyranny, in *Participation the new tyranny,* Zed Books London.
[74] Roche, C. (2009) Promoting Voice for Choice: Exploring Innovations for Australian NGO Accountability for Development Effectiveness
[75] Jones, H. (2011); O'Flynn, M. (2009), Tracking Progress in Advocacy: Why and how to monitor and evaluate

REFERENCES

Ackerman, J. (2005*) Human Rights and Social Accountability, World Bank, Social Development Papers, Participation and Civic Engagement, Paper No. 86, Washington DC: World Bank*

Arndt, C. & Oman, C. (2006) *The Uses and Abuses of Governance Indicators,* Organisation for Economic Cooperation and Development *OECD, http://www.oecd.org/document/25/0,3746,en_2649_33935_37081881_1_1_1_1,00.html*

Barahona, C. & Levy, S. (2003) *'How to generate statistics and influence policy using participatory methods in research'*. Statistical Services Centre Working Paper, University of Reading http://www.reading.ac.uk/ssc/workareas/participation/How_to_generate_stats_and_influence_policy.pdf

Barahona, C. (2010) Randomised control trials for the impact evaluation of development initiatives: A statistician's Point of view, Institutional Learning and Change Working Paper 13 http://www.cgiar-ilac.org/files/publications/working_papers/ILAC_WorkingPaper_No13_Randomised%20Control%20Trials.pdf

Bloom, E. Sunseri, A., & Leonard, A. (2007) Measuring and Strengthening local governance capacity the Local Governance Barometer, PACT and USAID http://pdf.usaid.gov/pdf_docs/PNADK250.pdf

Booth, D, & Fritz, V,(2008) Final Synthesis Report: Good governance, aid modalities and poverty reduction: from better theory to better practice *Research Project (RP-05-GG) of the Advisory Board for Irish Aid* available from www.odi.org.uk

Burge, R. (2010) 'Final Report: Learning from DFID's Governance and Transparency Fund (GTF): Tools, Methods and Approaches.' Unpublished Report. TripleLine/KPMG. June.

Burnell, P. (2007) (ed) Evaluating Democracy Support: Methods and Experiences, International Institute for Democracy and Electoral Assistance and SIDA available from: http://www.idea.int/publications/evaluating_democracy_support/index.cfm

Catley, A., Burn, J., Abebe, J. and Suji,O.(2008) Participatory impact assessment: a guide for practitioners, Feinstein International Center downloadable from: http://www.preventionweb.net/english/professional/publications/v.php?id=9679

Chambers, R. (2007) 'Who Counts? The Quiet Revolution of Participation and Numbers', *IDS Working Paper* 296, Brighton: IDS available from: http://www.ntd.co.uk/idsbookshop/details.asp?id=1006

Chambers, R. (2010a) 'Paradigms, Poverty and Adaptive Pluralism', *IDS Working Paper* 344, Brighton: IDS, available from: http://www.ntd.co.uk/idsbookshop/details.asp?id=1187

Chambers, R. (2010b) A Revolution Whose Time Has Come? The Win-Win of Quantitative Participatory Approaches and Methods, IDS Bulletin *Special Issue: People-centred M&E: Aligning Incentives So Agriculture Does More to Reduce Hunger*, Vol 41, Issue 6 p44-55

Clark, D., Fox, J. and Treakle, K. (2003) *Demanding Accountability: Civil Society Claims and the World Bank Inspection Panel.* Oxford: Rowman & Littlefield.

---

advocacy projects and programmes Starling, S. (2010)

Cooke, B. and Kothari, U (eds) (2001) The case for participation as tyranny, in *Participation the new tyranny,* Zed Books London.

Cornwall, A (2004), "Introduction: New democratic spaces? The politics and dynamics of Institutionalised Participation – IDS Bulletin Nr 35

Cugelman, D and Oterao (2010) Basic Efficiency Resource:
A framework for measuring the relative performance of multi-unit programs, Alterspark
http://www.alterspark.com/uploads/cugelman_otero_ber-whitepaper_v22.pdf

Dahl-Østergaard, T., Schulz, K, & Svedberg B. (2008) Experiences and Lessons Learnt from Sida's Work with Human Rights and Democratic Governance, Department for Democracy and Social Development SIDA Evaluation 2008:29

Davies, R. and Dart, J (2004) The Most Significant Change Technique: A guide to its use http://www.mande.co.uk/docs/MSCGuide.pdf

Deaton, A. (2009) Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development, *Research Program in Development Studies*, Center for Health and Wellbeing, Princeton University

DFID, (2010) Impact Evaluation Terms of Reference, Unpublished DFID UK

Ebrahim (2005) NGOs and Organizational Change. Discourse, Reporting and Learning, Cambridge University Press.

Eyben, R. (2009) Stories of Empowerment: Framework Document for DAC POVNET Task Team. Unpublished Note.
http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DCD/DAC/POVNET(2009)5/ADD1&docLanguage=En

Eyben, R. and Guijt I. (2010) The Big Pushback concept note unpublished

Fowler, A. (1998) 'Authentic NGDO partnerships in the new policy agenda for international aid: dead end or light ahead', *Development and Change* 29: 137-159;

Gabarino, S & Holland, J, (2009) Quantitative and Qualitative Methods in Impact Evaluation and Measuring Results Issues Paper, GDRC UK http://www.gsdrc.org/docs/open/EIRS4.pdf

Gaventa, J. and Barrett, G. (2010) 'So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement', *IDS Working Paper* 347, Brighton: IDS, available from: http://www.ntd.co.uk/idsbookshop/details.asp?id=1197

Goetz, Anne Marie (2005) Social justice, gender justice, and local government *Background and concept note for: Assessing the social justice outcomes of participatory local governance, a project of the Ford Foundation Local Governance and Learning group (LOGO)*

GSDRC (2008) Helpdesk research report: monitoring and evaluation of participation in governance (07.05.10) *available at www.gsdrc.org*

Guijt, I. (2010) Accountability and Learning – exploding the myth of incompatibility between accountability and leaning in eds e*dited by Jan Ubels, Naa-Aku Acquaye-Baddoo and Alan Fowler.* Capacity Development in Practice
http://www.snvworld.org/en/Documents/21%20Accountability%20and%20Learning%20-%20Exploding%20the%20Myth%20of%20Incompatibility%20between%20Accountability%20and%20Learning%20-%20Irene%20Guijt%20-.pdf

Haddad, L. Lindstrom, J. and Pinto, Y. (2010) The sorry state of M&E in Agriculture: can people centred approaches help? IDS Bulletin *Special Issue: People-centred M&E: Aligning Incentives So Agriculture Does More to Reduce Hunger*, Vol 41, Issue 6

Holland, J. and Thirkell, A. (2009), *Measuring Change and Results in Voice and Accountability Work*, DFID Working Paper http://www.dfid.gov.uk/Documents/publications1/evaluation/meas-chge-res-voice-acc-wk.pdf

Holland, J. (forthcoming) Concept note on participatory evaluation (unpublished)

Hudock, A. (1999) *NGOs and Civil Society: Democracy by Proxy*, Cambridge, Polity Press

Houtzager, P., Acharya A. & Gurza Lavalle, A. (2007) Associations and the exercise of citizenship in new democracies: evidence from Sao Paulo and Mexico City *IDS Working Paper 285* IDS, Brighton

Hyden, G. et al. (2008) 'Governance Assessments for Local Stakeholders: What the World Governance Assessment Offers', Overseas Development Institute, London

Jacobs, A., Barnett, C, & Ponsford, R. (2010) Three Approaches to Monitoring: Feedback Systems, Participatory Monitoring and Evaluation and Logical Frameworks, IDS Bulletin *Special Issue: People-centred M&E: Aligning Incentives So Agriculture Does More to Reduce Hunger*, Vol 41, Issue 6

Jajayaran, R. (2002) The Ten Seed Technique, World Vision, China. www.fao.org/participation/Ten-Seed%20Technique-Revised.pdf

Jones, H. (2011) A guide to monitoring and evaluating policy influence, Overseas Development Institute Background Notes, ODI http://www.dochas.ie/Shared/Files/4/A_guide_to_monitoring_and_evaluating_policy_influence.pdf

Jones, H. and Hearn, S. (2009) Outcome Mapping a realistic alternative for monitoring and evaluation, ODI Background Note, Overseas Development Institute http://www.odi.org.uk/resources/download/4118.pdf

Jupp, D. and Ibn Ali, S., with C. Barahona, 2010 (2010) Measuring Empowerment? Ask Them - Quantifying Qualitative Outcomes from People's Own Analysis: Sida Evaluation Series 2010: 1, Sida, Stockholm http://www.aline.org.uk/pool/measuring-empowerment-ask-them.pdf

Kapoor, I., 2004, 'Donor Participatory Governance Evaluation: Initial Trends, Implications, Opportunities, Constraints', Journal of International Development, 16 (2), 157-170

Kaufmann D., Kraay A., Mastruzzi M. (2003) "Governance Matters III: governance indicators for 1996- 2002", The World Bank Institute, Draft for comment http://www.worldbank.org/wbi/governance/pubs/govmatters3.html

Kaufman, D. and Kraay, A. (2007) On Measuring Governance: Framing Issues for Debate Issues paper prepared for the January 11 2007 Roundtable on Measuring Governance, World Bank, Washington, D.C. http://siteresources.worldbank.org/INTWBIGOVANTCOR/Resources/1740479-1149112210081/2604389-1167941884942/On_Measuring_Governance.pdf

Kaufman, D. and Kraay, A.,(2007) *Governance Indicators, Where are We and Where Should we Be Going?,* World Bank Working Policy Research Paper http://ideas.repec.org/p/wbk/wbrwps/4370.html

Laberge (2008) Claiming the MDGs: an Empowerment Framework, UNDP http://www.undp.org/oslocentre/docs08/Framework_1_-_Claiming_the_MDGs_-_Final.pdf

Lucas, R. and Longhurst R. Lucas, H. and Longhurst, R. (2010) 'Evaluation: Why, for Whom and How?', *IDS Bulletin* 41.6, Brighton: Blackwell Publishing

Logolink (2008) Draft research reading guide *unpublished*

Malamed, C. (2011) Does development give poor people what the want? Overseas Development Institute Briefing Note http://www.odi.org.uk/resources/download/5360.pdf

Manor, James (2004) Democratisation with inclusion: political reforms and people's empowerment at the grassroots *Journal of Human Development* 5/1, pp 5-29,

McGee, R. and Gaventa, J. (2010) 'Review of Impact and Effectiveness of Transparency and Accountability Initiatives ', Institute of Development Studies, Brighton, available from: http://www.ids.ac.uk/index.cfm?objectid=7E5D1074-969C-58FC-7B586DE3994C885C

Mosse, D. (2004) Power relations and poverty reduction *in Power, Rights, and Poverty: Concepts and Connections, a working meeting sponsored by DFID and the World Bank March 23-24 2004,* edited by Ruth Alsop

Munck, G. (2008) *Measuring democracy: a bridge between scholarship and politics* – Johns Hopkins University Press, Baltimore Md.

Natsios, A. (2010) 'The Clash of the Counter-bureaucracy and Development', Center for Global Development www.cgdev.org/content/publications/detail/1424271

Neubert, S. (2010) Description and examples of MAPP: Method for Impact Assessment of Programmes and Projects, Bonn: German Development Institute www.ngo-ideas.net/mediaCache/MAPP/

Nguyen, T. and Lassibille, G. (2008) 'Improving Management in Education: Evidence from a Randomized Experiment in Madagascar.' Draft Paper. Available at http://www.povertyactionlab.org/sites/default/files/publications/93_Nguyen_Improving_Management.pdf

O'Neill, T. Foresti, M. and Hudson A. (2007) Evaluation of Citizen's Voice and Accountability: Review of the Literature and Donor Approaches Report, ODI http://www.dfid.gov.uk/Documents/publications1/evaluation/citizen-voice-accountability.pdf

OECD DAC, 2008, 'Survey of Donor Approaches to Governance Assessment', Organisation for Economic Cooperation and Development - Development Assistance Committee, Paris www.oecd.org/dataoecd/58/32/42258487.pdf

Oxfam's PPA Baseline Report

Pollitt, C. (2008) 'Moderations in all things': Governance quality and performance measurement, paper for the structure and organization of government ASOG meeting Gottenburg http://www.qog.pol.gu.se/working_papers/SOG%20papers/Pollitt%20-%20SOG%20Conference%20Nov08.pdf

Quinn- Patton, M. (2009) Developmental Evaluation, Presentation CES http://www.evaluationcanada.ca/distribution/20090601_quinn_patton_michael_a.pdf

Ramalingham, B. (2008) Evaluation and complexity, presentation from a Norad conference available at http://www.outcomemapping.ca/resource/resource.php?id=183

Ravindra, A. (2004) 'An Assessment of the Impact of Bangalore Citizen Report Cards on the Performance of Public Agencies.' Evaluation Capacity Development Working Paper no.12, Operations Evaluation Department. Washington, DC: World Bank

Reimenschneider, N. (2010) "Comparing quantitative results from three different stories: do they tell the same story?", *OPM Briefing Notes* 2010 (01), Oxford, Oxford Policy Management (cited in Holland forthcoming)

Ravallion, M. (2009) 'Should the *Randomistas* Rule?' *Economists' Voice*. Berkeley Electronic Press, www.bepress.com/ev

Robinson, M. (2004) Resources, citizen engagement and democratic local governance, a background paper *unpublished*? www.logolink.org/index.php/**resource**/download/512

Rocha Menocal, A & Sharma, B. (2008) *Joint Evaluation of Citizens' Voice and Accountability*: synthesis report – DFID, London
http://www.dfid.gov.uk/Documents/publications1/evaluation/voice-accountability-synthesis.pdf

Roche, C. (2009) Promoting Voice for Choice: Exploring Innovations for Australian NGO Accountability for Development Effectiveness

SADEV (2006) Swedish Agency for Development Evaluation How to trace results of democracy support Policy Brief.
http://www.sadev.se/Documents/Publikationer%202006/SADEV%20Report%202006_1.pdf

Sarles, M. (2007) Evaluating the impact and effectiveness of USAID's democracy and governance programmes, in Burnell, P. (ed) Evaluating Democracy Support: Methods and Experiences, International Institute for Democracy and Electoral Assistance and SIDA
http://www.idea.int/publications/evaluating_democracy_support/index.cfm

Shutt, C. (2006a) 'Money matters in aid relationships', in (ed) R. Eyben *Relationships for Aid,* London: Earthscan, Eyben, R. Relationships for Aid

Shutt, C. (2008) Power and Aid Relations: A Study of Cambodian NGOs and their International Funders', unpublished thesis, Brighton: IDS

Starling, S. (2010) Monitoring and evaluating advocacy: lessons from Oxfam GB's Climate Change campaign Development in Practice, Volume 20, Number 2, April 2010

Thomas, M. (2009) *What do the Worldwide Governance Indicators Measure,* European Journal of Development Research

UNDP Governance Assessment Portal http://www.gaportal.org/areas-of-governance

UNDP, (2004) 'Governance Indicators: A Users Guide', 2nd Edition, United Nations Development Programme, UNDP Oslo Governance Centre, Norway,
http://www.undp.org/oslocentre/docs04/UserGuide.pdf

UNDP (2006) Indicators for Human Rights Based Approaches to Development in UNDP Programming: A Users' Guide http://gaportal.org/sites/default/files/HRBA%20indicators%20guide.pdf

UNDP, (2006) Measuring Democratic Governance: A framework for Selecting Pro Poor and Gender Sensitive Indicators, UNDP Democratic Governance Group, UNDP Oslo Governance Centre, Norway
http://www.undp.org/oslocentre/docs06/Framework%20paper%20-%20entire%20paper.pdf

UNDP, (2006) A Guide to Measuring the Impact of Right to Information Programmes, UNDP Democratic Governance Group, UNDP Oslo Governance Centre, Norway
http://www.undp.org/oslocentre/docs06/A%20Guide%20to%20Measuring%20the%20Impact%20of%20Right%20to%20Information%20Programmes%20-%20final%20(11%2004%2006).pdf

UNDP (2007) Sources for Democratic Governance Indicators, UNDP Oslo Governance Centre, Norway www.undp.org/oslocentre/docs04/Indicator%20Sources.pdf

UNDP (2008) Review of results-based management at the United Nations, Report of the Office of Internal Oversight Services http://www.un.org/ga/search/view_doc.asp?symbol=A/63/268

UNDP (2009) Measuring Local Governance Guide, UNDP Democratic Governance Group, UNDP Oslo Governance Centre, Norway http://www.undp.org/oslocentre/docs09/LGGuide2July.pdf

Indicators of parliamentary performance budget processes
http://www.parlcent.ca/indicators/budget_process_e.php

Wilson, I. (undated) Some practical sampling procedures for development research, http://www.reading.ac.uk/ssc/workareas/participation/Some_practical_sampling_procedures.pdf

Wallace, T., Bornstein, L., & Chapman, J.(2006) *The Aid Chain: Coercion and Commitment in Development NGOs,* Intermediate Technology Development Group.

World Bank Empowerment in Practice 2005 http://siteresources.worldbank.org/INTEMPOWERMENT/Resources/Empowerment_in_Practice.pdf

Annex 1 INGO practitioners consulted during the process of research

Name	**Organisation**
Doug Orr	Plan Former Learning and Impact Manager Plan UK
Jake Phelan	Current Learning and Impact Manager Plan UK
Olivia MacDonald	GTF Programme Manager Christian Aid
Amy Pollard	Aid Effectiveness Policy Officer CAFOD
Jo Rowlands	Governance and Institutional Accountability Advisor Oxfam GB
Claire Hutchings	Monitor, Evaluation and Learning Team Oxfam GB
Laurie Adams	Head of Impact Assessment and Shared Learning ActionAid
John Gaventa	Chair of Oxfam GB, author of recent article on citizen engagement

Annex 2 Summary of Selected Resources

Some of the documents and resources noted in the review that may be used for developing governance M&E systems and choosing indicators are briefly summarised in the text box below.

Purpose/Utility	**Resource**
Exploring different evaluation approaches	Lucas, R. and Longhurst R. (2010) 'Evaluation: Why, for Whom and How?', *IDS Bulletin* 41.6, Brighton: Blackwell Publishing O'Neill, T. Foresti, M. and Hudson A. (2007) Evaluation of Citizen's Voice and Accountability: Review of the Literature and Donor Approaches Report, ODI http://www.dfid.gov.uk/Documents/publications1/evaluation/citizen-voice-accountability.pdf Quinn Patton, M. (2009) Developmental Evaluation, Presentation CES http://www.evaluationcanada.ca/distribution/20090601_quinn_patton_michael_a.pdf Ramalingham, B. (2008) Evaluation and complexity, presentation from a NORAD conference available at http://www.outcomemapping.ca/resource/resource.php?id=183
Exploring links between accountability and learning	Guijt. I (2010) Accountability and Learning – exploding the myth of incompatibility between accountability and leaning INTRAC, http://www.snvworld.org/en/Documents/21%20Accountability%20and%20Learning%20-%20Exploding%20the%20Myth%20of%20Incompatibility%20between%20Accountability%20and%20Learning%20-%20Irene%20Guijt%20-.pdf Roche, C. (2009) Promoting Voice for Choice: Exploring Innovations for Australian NGO Accountability for Development Effectiveness
Analysing the utility of global governance indicators	Holland, J. and Thirkell A. (2009), *Measuring Change and Results in Voice and Accountability Work*, DFID Working Paper http://www.dfid.gov.uk/Documents/publications1/evaluation/mea

	s-chge-res-voice-acc-wk.pdf UNDP (2007) Sources for Democratic Governance Indicators, UNDP Oslo Governance Centre, Norway www.undp.org/oslocentre/docs04/**Indicator**%20**Sources**.pdf
Guidelines for approaches that encourage the identification of pro-poor and gender sensitive indicators by those groups.	UNDP, (2006) Measuring Democratic Governance: A framework for Selecting Pro Poor and Gender Sensitive Indicators, UNDP Democratic Governance Group, UNDP Oslo Governance Centre, Norway http://www.undp.org/oslocentre/docs06/Framework%20paper%20-%20entire%20paper.pdfUNDP guide on governance and empowerment.
Guide to choosing indicators within rights based approaches	UNDP (2006) Indicators for Human Rights Based Approaches to Development in UNDP Programming: A Users' Guide http://gaportal.org/sites/default/files/HRBA%20indicators%20guide.pdf
Frameworks for conceiving social justice outcomes of participatory local governance	Goetz, Anne Marie (2005) Social justice, gender justice, and local government *Background and concept note for: Assessing the social justice outcomes of participatory local governance, a project of the Ford Foundation Local Governance and Learning group (LOGO)* Related http://www.idrc.org/tehip/ev-111764-201-1-DO_TOPIC.html
Defining citizenship outcomes and domains of change	Gaventa, J. and Barrett, G. (2010) 'So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement', *IDS Working Paper* 347, Brighton: IDS, available from: http://www.ntd.co.uk/idsbookshop/details.asp?id=1197
Choosing indicators and thinking about approaches for voice and accountability democracy support evaluation and IA	Holland J. and Thirkell A. (2009), *Measuring Change and Results in Voice and Accountability Work*, DFID Working Paper http://www.dfid.gov.uk/Documents/publications1/evaluation/meas-chge-res-voice-acc-wk.pdf O'Neill, T. Foresti, M. and Hudson A. (2007) Evaluation of Citizen's Voice and Accountability: Review of the Literature and Donor Approaches Report, ODI http://www.dfid.gov.uk/Documents/publications1/evaluation/citizen-voice-accountability.pdf SADEV (2006) Swedish Agency for Development Evaluation How to trace results of democracy support Policy Brief. http://www.sadev.se/Documents/Publikationer%202006/SADEV%20Report%202006_1.pdf
Thinking about relationships between quantitative and qualitative methods	Gabarino, S & Holland, J, (2009) Quantitative and Qualitative Methods in Impact Evaluation and Measuring Results Issues Paper, GDRC UK http://www.gsdrc.org/docs/open/EIRS4.pdf Gaventa, J. and Barrett, G. (2010) 'So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement', *IDS Working Paper* 347, Brighton: IDS, available from: http://www.ntd.co.uk/idsbookshop/details.asp?id=1197
Introduction to Outcome Mapping	Jones, H. and Hearn, S. (2009) Outcome Mapping a realistic alternative for monitoring and evaluation, ODI Background Note, Overseas Development Institute http://www.odi.org.uk/resources/download/4118.pdf
Thinking about how to measure empowerment	World Bank Empowerment in Practice 2005 http://siteresources.worldbank.org/INTEMPOWERMENT/Resources/Empowerment_in_Practice.pdf

	Gaventa, J. and Barrett, G. (2010) 'So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement', *IDS Working Paper* 347, Brighton: IDS, available from: http://www.ntd.co.uk/idsbookshop/details.asp?id=1197 Goetz, Anne Marie (2005) Social justice, gender justice, and local government *Background and concept note for: Assessing the social justice outcomes of participatory local governance, a project of the Ford Foundation Local Governance and Learning group (LOGO)* Jupp, D. and Ibn Ali, S., with C. Barahona, 2010 (2010) Measuring Empowerment? Ask Them - Quantifying Qualitative Outcomes from People's Own Analysis: Sida Evaluation Series 2010: 1, Sida, Stockholm http://www.aline.org.uk/pool/measuring-empowerment-ask-them.pdf
Designing approaches to establish relative contributions of different units to complex programmes	Cugelman, D and Oterao (2010) Basic Efficiency Resource: A framework for measuring the relative performance of multi-unit programs, Alterspark http://www.alterspark.com/uploads/cugelman_otero_ber-whitepaper_v22.pdf
Selecting indicators of parliamentary performance	Indicators of parliamentary performance I budget processes http://www.parlcent.ca/indicators/budget_process_e.php
Thinking about how to measure local governance	Bloom, E. Sunseri, A., & Leonard, A. (2007) Measuring and Strengthening local governance capacity the Local Governance Barometer, PACT and USAID http://pdf.usaid.gov/pdf_docs/PNADK250.pdf Goetz, Anne Marie (2005) Social justice, gender justice, and local government UNDP (2009) Measuring Local Governance Guide, UNDP Democratic Governance Group, UNDP Oslo Governance Centre, Norway http://www.undp.org/oslocentre/docs09/LGGuide2July.pdf
Approaches for the participatory identification of democratic governance indicators	Institute of Democracy and Electoral Assistance Assessment framework http://www.idea.int/sod/framework/index.cfm http://www.idea.int/sod/
Guide to using the most significant change technique	Davies, R. and Dart, J (2004) The Most Significant Change Technique: A guide to its use http://www.mande.co.uk/docs/MSCGuide.pdf
Introduction to sampling procedures	Wilson, I. (undated) Some practical sampling procedures for development research, http://www.reading.ac.uk/ssc/workareas/participation/Some_practical_sampling_procedures.pdf

Sourcing information on participatory approaches to M&E and IA and participatory statistics	Chambers, R. (2007) 'Who Counts? The Quiet Revolution of Participation and Numbers', *IDS Working Paper* 296, Brighton: IDS http://www.ntd.co.uk/idsbookshop/details.asp?id=1006 Chambers, R. (2010) A Revolution Whose Time Has Come? The Win-Win of Quantitative Participatory Approaches and Methods, IDS Bulletin *Special Issue: People-centred M&E: Aligning Incentives So Agriculture Does More to Reduce Hunger*, Vol 41, Issue 6 p44-55 Jupp, D. and Ibn Ali, S., with C. Barahona, (2010) Measuring Empowerment? Ask Them - Quantifying Qualitative Outcomes from People's Own Analysis: Sida Evaluation Series 2010: 1, Sida, Stockholm http://www.aline.org.uk/pool/measuring-empowerment-ask-them.pdf Cateley, A., Burn, J., Abebe, J. and Suji,O.(2008) Participatory impact assessment: a guide for practitioners, Feinstein International Center downloadable from: http://www.preventionweb.net/english/professional/publications/v.php?id=9679 Neubert, S. (2010) Description and examples of MAPP: Method for Impact Assessment of Programmes and Projects, Bonn: German Development Institute www.ngo-ideas.net/mediaCache/MAPP/ Parks, W., Gray Felder, D., Hunt, and Byrne, A. Who Measures Change: An introduction to participatory monitoring and evaluation for communication for social change http://www.intrac.org/data/files/resources/177/who_measures_change.pdf
Documenting stories of empowerment	Eyben, R. (2009) 'Stories of Empowerment: Framework Document for DAC POVNET Task Team. Unpublished Note. http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=DCD/DAC/POVNET(2009)5/ADD1&docLanguage=En
Thinking about CSO approaches to accountability	Roche, C. (2009) Promoting Voice for Choice: Exploring Innovations for Australian NGO Accountability for Development Effectiveness http://ngoperformance.org/related-initiatives/feedback-examples/ http://ngoperformance.org/2011/01/19/keystone-partner-survey-powerful-new-data-on-ngos'-performance/
Designing evaluation approaches for advocacy	Jones, H. (2011) A guide to monitoring and evaluating policy influence, Overseas Development Institute Background Notes, ODI http://www.dochas.ie/Shared/Files/4/A_guide_to_monitoring_and_evaluating_policy_influence.pdf O'Flynn, M. (2009), Tracking Progress in Advocacy: Why and how to monitor and evaluate advocacy projects and programmes http://www.intrac.org/data/files/resources/672/Tracking-Progress-in-Advocacy-Why-and-How-to-Monitor-and-Evaluate-Advocacy-Projects-and-Programmes.pdf Starling, S. (2010) Monitoring and evaluating advocacy: lessons from Oxfam GB's Climate Change campaign Development in Practice, Volume 20, Number 2, April 2010