

## The Sorry State of M&E in Agriculture: Can People-Centred Approaches Help?

Lawrence Haddad, Johanna Lindstrom and Yvonne Pinto<sup>i</sup>

*“The next Green Revolution ....must be guided by small-holder farmers, adapted to local circumstances, and sustainable for the economy and the environment”*

Bill Gates, World Food Prize speech, October 2009<sup>ii</sup>

### Abstract

This paper argues that if the multiple purposes of M&E were recognised and pursued it would help align the incentives of funders, implementers, M&E service providers, and intended beneficiaries to increase the impacts of agriculture on poverty. In reality, these multiple purposes are rarely pursued, leading to the weak provision of meaningful M&E. We make the case that M&E in agriculture is not immune to this, and it may even be more susceptible due to agriculture’s unique properties. We argue that one possible way forward is people-centred M&E, which looks for ways to balance multiple accountabilities. People-centred M&E embraces farmer feedback, focuses on incentivising learning within organisations and finds ways to share M&E information more openly.

### Introduction

Done well, monitoring and evaluation (M&E) has the potential to make enormous contributions to development practice and theory.<sup>iii</sup> Good M&E can make projects work better, assess impact, steer strategy, increase stakeholder ownership, build the capacity of stakeholders to hold programme financiers and implementers to account and share learning more widely.<sup>iv</sup>

This paper argues that if these multiple purposes of M&E were pursued it would help align the incentives of funders, implementers, M&E service providers, and intended beneficiaries to increase the impacts of agriculture on poverty. In reality, these multiple purposes are rarely pursued, leading to the weak provision of meaningful M&E. We make the case that M&E in agriculture is not immune to this, and it may even be more susceptible due to agriculture’s unique characteristics.

Part 1 of this paper sets the scene and explores the role of M&E within the overall aid system, looking particularly at how the aid effectiveness agenda and the economic and financial crisis has promoted a greater emphasis on aid ownership, harmonisation, mutual accountability, results, and alignment. This new agenda has both created and restricted the space for innovations in M&E and we argue that the M&E community in agriculture and beyond has not responded adequately to it.

Part 2 analyses the state of M&E in agriculture and finds it weak. Part 3 identifies the drivers of the weak provision of meaningful M&E as a failure to identify and capture the multiple benefits of M&E, and the closed and sensitive nature of M&E systems.

Part 4 suggests another way forward – people-centred M&E. People-centred M&E looks for ways to balance multiple accountabilities, focuses on learning within organisations and the individual incentives for learning, and finds ways to share M&E performance information more openly. In this way we argue that it can help to identify and capture multiple benefits of M&E and it reduce

information asymmetries - all of which we argue inhibit the provision and use of poverty reducing M&E.<sup>v</sup>

We argue that the achievement of this approach will allow the practice of M&E to better support improved outcomes for poor farmers while supporting donor reporting requirements. As with any proposal for change the underlying assumptions must be tested against the evidence. That is what this paper aims to do.

## **1. The changing context for M&E**

Delivering good change through public policy is challenging in any context. Headline political arguments have to be won to set a certain direction and smaller political battles have to be won for effective implementation. Demonstrating impact at scale in the real world is difficult. When there is a political and spatial dislocation between funders and intended beneficiaries the ability of the latter to hold the former to account is weakened. Poor programme performance is not sufficiently challenged, good programme performance not sufficiently incentivised. There is a broken feedback loop (Barder 2009). Situate this broken feedback loop within systems that have weaker capacities to monitor resources and are subject to frequent disruptions and it is inevitable that there is a weak learning and impact culture in aid and international development more broadly.

These concerns were crystallised in the Paris Declaration with a renewed call to improve the quality of aid in terms of ownership (countries set their own priorities), harmonisation (donors avoid overburdening countries), mutual accountability, results, and alignment (donors get behind country priorities).

In the past 5 years the quality agenda has got bogged down (Wood et al. 2008), while most of the additional attention has been on maintaining and increasing aid flows. The global financial crisis has put quantity at risk and has led to a refocusing on quality. There is now a stronger emphasis on results, transparency, accountability to taxpayers and the citizens of aid receiving countries, and value for money for all.

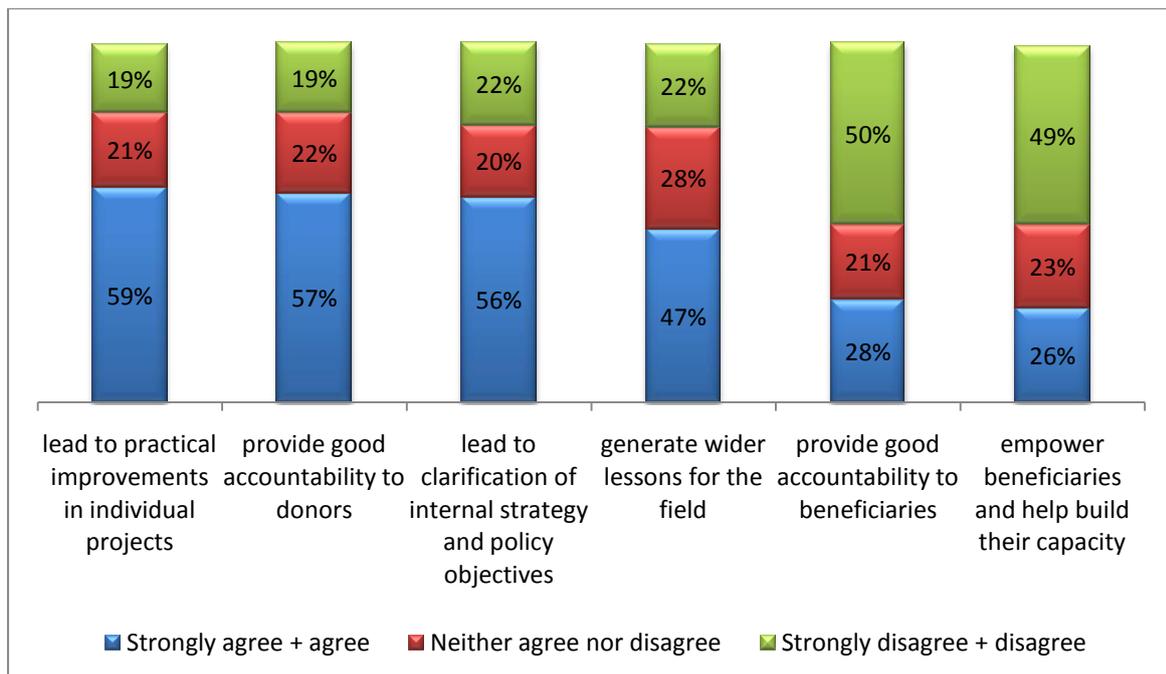
The field of agriculture has not yet, we argue, responded to this call to arms in any significant way. But now it has to, because after decades of neglect there is a renewed focus on agricultural development with an upturn in resources going to agriculture<sup>vi</sup>. There is a greater interest in initiatives that aim to make a difference to poor farmers' lives. However, with new interest and resources comes an increased need to ensure that initiatives lead to positive sustained change on the ground. Are we seeing improvements in agricultural productivity which are equitable and sustainable? Is there greater social inclusion for poor farmers? Are resources directed towards initiatives that correspond to the priorities of poor farmers? Are we seeing agriculture fulfil its potential to deliver food security, nutrition and poverty alleviation? Within the current climate of austerity the need for agriculture and agricultural M&E to respond is even greater.

Many within the agricultural development community recognise the need to strengthen their accountability to poor farmers and other stakeholders. For instance:

- The recent Global Conference on Agricultural Research for Development (GCARD) had results for poor communities and wider partnerships as key cross-cutting themes and the lead paper for the conference noted: ‘... a change is needed in the incentive structures in the national and international research community to deliver impacts for the poor. ... [S]ystems need to be more accountable to their beneficiaries.’ (Lele et al 2010: xii);
- The recent Consultative Group on International Agricultural Research (CGIAR) Independent Review Panel suggests that measurement and feedback loops appear to have a major role in sustaining reform processes in agriculture (Science Council of the CGIAR 2007);
- Established donors such as the International Fund for Agricultural Development (IFAD) and the World Bank claim to be more and more focused on people-centred approaches in their programmes (see for instance IFAD 2002 and World Bank 2005);
- new donors in the agricultural development field such as the Bill and Melinda Gates Foundation have emphasized the need to recognize the priorities of farmers (Bill & Melinda Gates Foundation 2009);
- There is a long tradition of farmer focused initiatives in agricultural research many of which have been piloted and documented by the Farmer First initiative and a few members of the CGIAR, although never wholly embraced by the powerful players within the international agriculture system (Ashby 2009; Scoones and Thompson 2009);
- An increasing number of NGOs practice more balanced stakeholder approaches to M&E within agriculture and beyond (Jacobs and Jacobs et al, in this IDS Bulletin)

However, to date, reality is falling short of this rhetoric. Our own analysis of agricultural M&E based on a survey of M&E practitioners and agricultural experts<sup>vii</sup> (see Figure 1) indicates that 50 percent of the respondents think that current M&E practice does not provide good accountability to beneficiaries. According to the 209 respondents, the strongest feature of current M&E practice is the ability to lead to practical improvements in projects, closely followed by the ability to provide good accountability to donors and clarify internal strategies and policy objectives. Current practice is also good at generating wider lessons for the field but it rarely provides good accountability to beneficiaries, rarely empowering them or building their capacity. Accounting to donors and therefore taxpayers seems to be a greater priority than exhibiting accountability to beneficiaries. These results characterise a culture of compliance rather than one of sustainable results.

**Figure 1: Responses (N=209) to statement ‘Current M&E Practice tends to’**



Source: Lindstrom 2009

## 2. What is the state of M&E in agriculture?

The broad consensus in the literature is that M&E – including agriculture – is weaker than it should be. The shortcomings of M&E make familiar reading. In a wide ranging review of the state of M&E Munce (2005) collates the standard critiques:

- A failure to specify what M&E is for and to facilitate its use
- Lack of stakeholder participation and responsiveness: ‘Effective participation requires, perhaps above all, a climate where stakeholders, including donors, see each other as partners with the common ultimate purpose of achieving development results.’ (p. 12)
- Too little attention to theories of change – too much of a focus on inputs and outputs and not enough on outcomes
- A lack of systematic capacity building for M&E: ‘Despite the term’s frequent usage and cited importance in the literature and despite the vast number of project M&E activities that are annually conducted, capacity building for M&E is frequently overlooked.’ (p. 18)

Other major criticisms include:

- Not enough resources spent on impact evaluations – i.e. studies that document the extent to which changes in the wellbeing of the target population can be attributed to the particular programme (measuring net changes) (CGD 2006 and Ravallion 2008).
- Not enough focus on the trajectories for impact - i.e. the pathways of impact, particularly in relation to the time it takes to achieve impact (Woolcock 2009).

- Not enough focus on the context of interventions – i.e. ensuring evaluation not only asks ‘what works and how’, but also ‘under what circumstances’ (Rogers 2008; Lucas and Longhurst in this IDS bulletin)
- Not enough focus on flexibility and learning – adapting to changing circumstances and learning from successes as well as failures (Oswald and Taylor in this IDS Bulletin).
- Not enough use of M&E data and findings (see Patton 1978 and Patton and Horton 2009 for a discussion). For instance, Bastoe argues that there is ‘... a growing disillusionment with conventional evaluation praxis. Many governments experience only limited use of evaluation findings. Evaluation findings do not automatically feed back into a receptive and responsive decision-making process.’ (2006: 97 quoted in Lucas forthcoming).

Reviews of M&E in agriculture tell a similar story. A scoping paper for an external review of social science in the CGIAR, summarised in Box 1 identifies some initial concerns. Weak social science in CGIAR centres reflects, in part, a lack of priority given to M&E and will, in turn, generate weak M&E.

**Box 1: Conclusions of Recent External Reviews of CGIAR Centres**

- CIAT has no core capacity to do impact analysis;
- CIMMYT social science was only working in half of the areas that CIMMYT recognized as its core tasks; and this was before it shut down what was traditionally known as one of the most successful Economics Programs in the CGIAR System;
- According to ILRI’s 2008-10 MTP, a large share of its work is social science in nature and more than 20 research scientists work in this area, yet there is almost no evidence that its social scientists are publishing in journals reviewed and recognized by their peers;
- IWMI greatly expanded its social science capacity, mostly with young and inexperienced social scientists without hiring almost any that had proven he/she could plan and direct a major social science agenda--the output of the program shows that this probably was a mistake;
- Recent External Programme and Management Reviews (EPMRs) of ICARDA, IPGRI (now Bioversity) and others have concluded that they have few systematic programs in economics, social research or general social science.

Source: Science Council of the CGIAR (2007)

The CGIAR’s own Standing Panel on Impact Assessment (SPIA) lists impact evaluations done throughout the CGIAR . Table 1 shows that out of the 761 listed by the CGIAR as having been published in the past 14 years, only 83 listed impact focusing on welfare indicators such as income or nutrition/health status and only 67 listed income. As the table shows, in the past 4 years only 5 impact studies of agricultural research on income have been published.

**Table 1: CGIAR Impact Assessment Studies**

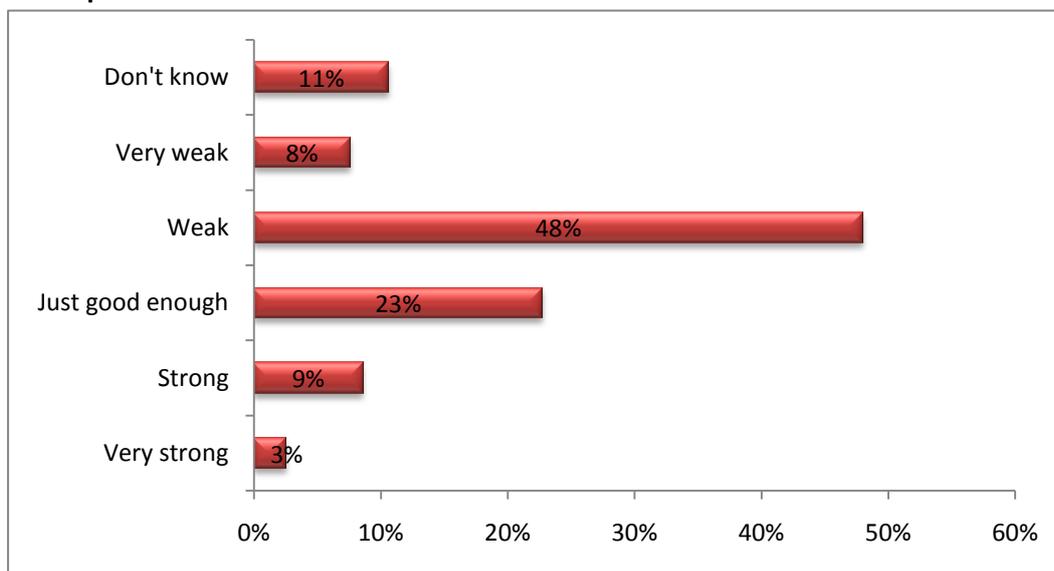
	Impact evaluations focusing on income as an outcome variable	Impact evaluations focusing on (income +nutrition/health) as an outcome variable	All Impact Evaluations
2008	0	0	
2007	1	2	
2006	4	4	
2005	0	0	
2004	4	5	
2003	5	6	
Total 1995-2008	67	83	761

As of August 29, 2009 <http://impact.cgiar.org/>

Neither the Poverty Action Lab (J-PAL) nor the International Initiative on Impact Evaluation (3ie) have undertaken or commissioned many agricultural project impact studies. As of mid 2009 the project database search at the Poverty Action Lab website shows 25 health evaluations, 38 in education, only 5 in agriculture (and these are all in Kenya)<sup>viii</sup>. And only 2 of 18 funded applications in round 1 of 3ie funding were awarded to agriculture projects (irrigation and low cost farm equipment) compared to 6 in health. Presumably this reflected some combination of low submissions and lack of quality of submissions<sup>ix</sup>.

Finally, from the stakeholder survey reported on in Figure 1, Figure 2 reports that a majority of respondents (56%) stated that in their opinions M&E in agriculture was weak or very weak<sup>x</sup>.

**Figure 2: Responses (n=198) about overall strength of monitoring and evaluation in agricultural development**



Source: Lindstrom 2009

Taken together, this evidence suggests that M&E is weak and that M&E in agriculture is no exception.

### 3. Why is M&E in agriculture so weak?

Generally, we argue that the major drivers of weak M&E in agriculture are common to many sectors and are:

- (a) a failure of the stakeholders to appreciate the multiple benefits of M&E
- (b) a failure of the investors to capture the multiple benefits of M&E as a result of:
  - the public goods aspects of M&E
  - a misalignment of implementer incentives to collect and use M&E data
- (c) the sensitive nature of M&E data which creates information asymmetries between M&E providers and users and results in underprovision of M&E

We make the case that (a) these drivers can be addressed through a focus on the multiple purposes of M&E so that the multiple benefits are identified and captured and information asymmetries are diminished, and (b) the different stakeholders involved in M&E investment, provision and use will likely have different knowledge of and preferences about these 5 purposes and so a focus on multiple purposes is best achieved by seeking the perspectives of several stakeholders. Figure 3 identifies five different purposes of M&E<sup>xi</sup> (as distinct from types of M&E) with opportunities for learning for all the stakeholders involved in M&E.

**Figure 3: The 5 purposes of M&E**



***Driver 1: A failure to appreciate the multiple benefits of M&E***

M&E can generate benefits in each of these five areas. If the benefits are understood and realised, then they will spur additional commitment, not just in terms of additional investments in time and resources, but in terms of general prioritisation of M&E. Since there are often limited resources available for M&E, the preferences of the more powerful actors will tend to dominate the way in which M&E is provided (see Oswald and Taylor in this IDS Bulletin). As we have seen in Figure 1, majorities of the surveyed stakeholders think that M&E currently generates project improvements, provides good accountability to donors and leads to clarification of internal strategies and policy objectives. These three functions are clearly of interest to those who invest in M&E. If more stakeholders were involved in project design and evaluation would a more rounded picture of benefits develop? We argue that this would be the case, thereby attracting additional commitment towards M&E.

## ***Driver 2: A failure to capture the multiple benefits of M&E***

We argue that there are two types of ‘failure to capture’:

- Investors may have an appreciation of the full range of benefits generated by M&E, but are simply unable to capture them because they are open to others for use and investors do not stand directly to benefit. This is a particular risk for private investment in M&E.
- There is a failure to capture benefits due to a lack of organisational incentives (e.g. incentives that disconnect those who initiated a programme from the results when they come in).

On the first type of ‘failure to capture’ Ravallion (2008) argues that underinvestment in M&E (or what he calls evaluative research) happens when project practitioners and funders do not stand to capture all the benefits of their investments. Drawing on numerous examples he argues that this public goods aspect of M&E will result in a particularly large underinvestment in M&E in (a) projects that will not yield short-term results<sup>xii</sup>, (b) projects that have more diffuse impacts, and (c) projects that help us understand the relevance of results for other contexts (external validation).<sup>xiii</sup>

The ‘underinvestment due to the public goods nature of M&E’ reasoning is used in the Centre for Global Development’s report ‘When Will We Ever Learn?’ (2006). This paper, one of the outputs leading to the formation of 3ie, identifies an ‘evaluation gap’ which it asserts is costly.<sup>xiv</sup> The report classifies M&E into four categories: (a) monitoring (are projects being implemented according to plan), (b) process and institution evaluations (how and why are things going to plan or not), (c) performance assessments (being accountable to stakeholders ‘by providing information about their activities and opening their books’) and (d) impact assessments. It argues that the evaluation gap is especially serious in the latter category, impact assessments. They make the case that this is because of two factors: (1) the public goods aspect (e.g. on methods and results) leads to underinvestment and (2) the additional data that need to be collected (e.g. comparator data) that are very resource intensive.

Whereas CDG identify only the global public goods aspects of impacts assessments, there are significant public goods aspects in other forms of M&E. On monitoring, there are methods that need to be shared to identify the purpose of the monitoring. For example, Ravallion (2000) maps spending (i.e. monitoring data) by poverty rates to get some real time data on the targeting performance of programmes. There are also methods that need to be shared to allow different stakeholders to monitor different groups, e.g. methods for allowing farmers to monitor agricultural development projects. On process evaluations, there is value in sharing methods for assessing capacities ex ante (can the organisation, its implementing staff and the institutional context within which they reside sustain the project?) and ex-post (did the project leave the organisation stronger or weaker?). The significant public goods aspects and additional data needs in the other three M&E categories suggest these areas are also significantly underinvested in.

The second type of ‘failure to capture’ relates to organisational incentives. Oswald and Taylor, in this IDS Bulletin, identify four types of mis-aligned incentives:<sup>xv</sup>

- **The incentive to demonstrate impact, the disincentive to learn why:** The current trend to demonstrate impact means that there is often an incentive for organisations to focus on what is easy to measure within an M&E process and less attention is given to learning how and why an intervention has been successful or not. This argument can be linked to the debate on M&E methods (see Lucas and Longhurst in this IDS Bulletin)
- **The incentive for upward accountability, the disincentive to learn from below:** Many development organisations are unbalanced to deliver on upward accountability. This can create incentives to value the knowledge and learning that donors demand, rather than the knowledge and learning that other stakeholders (including the organisation itself) value or require<sup>xvi</sup>.
- **The incentive to ‘do’, the disincentive to ‘learn’:** In many development organisations a particular member of staff or department is responsible for M&E and others are responsible for implementation, with limited interaction and mutual learning. The incentive is for some staff to ‘do’ and different staff to ‘learn’.
- **The incentive to conform, the disincentive to take risks:** Learning involves changes in behaviour and actions and inherently involves shifting power relationships. This can create disincentives for organisations to learn, as the learning outcomes can be challenging to those in positions of authority.

### ***Driver 3: The closed and sensitive nature of the M&E system***

This is the hypothesised driver for which we have least evidence. Nevertheless it is clear that M&E is a sensitive business. It is embarrassing to finance a programme that has little impact. It is hardly career advancing if your monitoring system shows milestones being missed. This means M&E reports have less chance of being made public. This also diminishes the incentives of managers to respond to M&E data. This leads to a more closed system with two types of information asymmetry between users and suppliers of M&E:

- Organisational reputation:
  - Do M&E users have a good track record of using M&E?
  - Do M&E suppliers provide value for money?
- Quality of the services currently being provided and used:
  - Are M&E services responsive to needs of users?
  - Are M&E services being used?

Unlike for most research, donors and other users of M&E have no third party peer review process to validate the quality of most M&E. The providers do not know enough about donors’ past and potential use of M&E services. This is compounded by the different preferences and needs these groups have for M&E. Information asymmetries will diminish effective M&E supply in the medium run as users are discouraged by reception of M&E they did not anticipate and in turn providers are discouraged by the non-use of their work.

Because of these three hypothesised drivers – failure to appreciate benefits, failure to capture benefits and information gaps between M&E users and providers – we argue that M&E in general suffers from a lack of investment, provision and use, and cannot enable the needed alignment of

incentives of funders, implementers, service providers and intended beneficiaries to increase the impacts on poverty.

### Is there anything particularly special about agriculture?

Although M&E is weak in general, we suggest that there are certain features of agriculture that make it even more difficult to achieve good M&E:

- **Lack of consensus on the overall goal of agriculture:** Impacts for farm households can be both direct (e.g. higher farm income) and indirect (e.g. higher wages need to be paid for in-demand farm workers). Similarly for the wider rural/urban space, impacts can be both direct (e.g. lower food prices) and indirect (e.g. improved food security and thus political stability). This makes it more difficult to agree on the purpose of M&E (see Meizen-Dick et al 2003).
- **Long, uncertain and complex<sup>xvii</sup> causal chains:** Agricultural development interventions tend to have longer, diffuse and less certain links between interventions and human welfare outcomes. Agricultural development interventions are often classified as complex<sup>xviii</sup>. This is particularly the case for agricultural research (Millstone et al in this IDS Bulletin) and agricultural policy initiatives (Sumner et al in this IDS Bulletin). This allows more opportunity for information asymmetries to arise making M&E more difficult in a technical sense<sup>xix</sup>.
- **High level of risk:** Agricultural development interventions are sensitive to the uncertainties imposed by climate and other phenomena (Chambers 1997), accentuating the potential disconnect between individual incentives and programme impacts (see Sabates-Wheeler et al for an overall discussion of farm-level risk and Devereux and Longhurst for a discussion on seasonality in this IDS Bulletin).
- **Lack of opportunity for beneficiaries of agricultural service delivery (e.g. farmers) to organise around a facility:** Recipients of health services or education services (whether medical staff or patients, teachers or students) are less dispersed than in agriculture. There is more routine contact between provider and user, and among users. Farmers and farmer services are more spatially dispersed than their counterparts in health and education. They can less easily share innovations and less easily exert collective pressure on other actors in the system (see FANRPAN 2005). Collecting their views is more expensive. Where farmer organizations exist, these face difficulties due to their political nature and lack of long-term support.
- **Agriculture is wrapped up in a wider range of cultural and institutional aspects.** It is a way of life, an individual identity, a collective political identity, a source of food, of income, and a means of managing the environment (Cernea & Kassam 2005). Although potentially not more so than in any other sector, gender inequalities are very prevalent in agriculture (Kabeer in this IDS Bulletin).

## 4. What can be done to strengthen M&E in agriculture?

If M&E in agriculture is weak, what can be done to identify M&E's wider benefits, to capture the benefits and close information gaps about M&E users and providers?

We hypothesise that a 'people-centred' perspective on M&E can help change the dynamic by:

- finding ways to balance multiple accountabilities,
- focusing on learning within organisations and the individual incentives for learning
- finding ways to share M&E performance information more openly.

The focus is on inclusion, organisation and incentives, rather than an emphasis only on tools and methods. Our arguments are summarised in Table 2.

**Table 2: People-centred M&E**

Component of people-centred M&E	The drivers of weak M&E and examples of how people centred M&E can counter them		
	Identifying benefits	Capturing benefits	Closing information gaps about M&E users and providers
1. Balancing multiple accountabilities	Greater scoping capacity to identify benefits and unintended consequences of interventions	Enable different groups to capture benefits by building their capacity to do so	Via feedback loops from intended beneficiaries
2. Enhancing organisational learning	Greater capacity to learn from other organisations	Greater organisational capacity to capture benefits	As organisations get better at responding to M&E they become more open
3. Building wider learning	Greater capacity to identify benefits from M&E work elsewhere	Greater ability to capture more indirect benefits	Generate learning about reputations and performance of M&E users and providers

What does the evidence say about these three hypotheses?

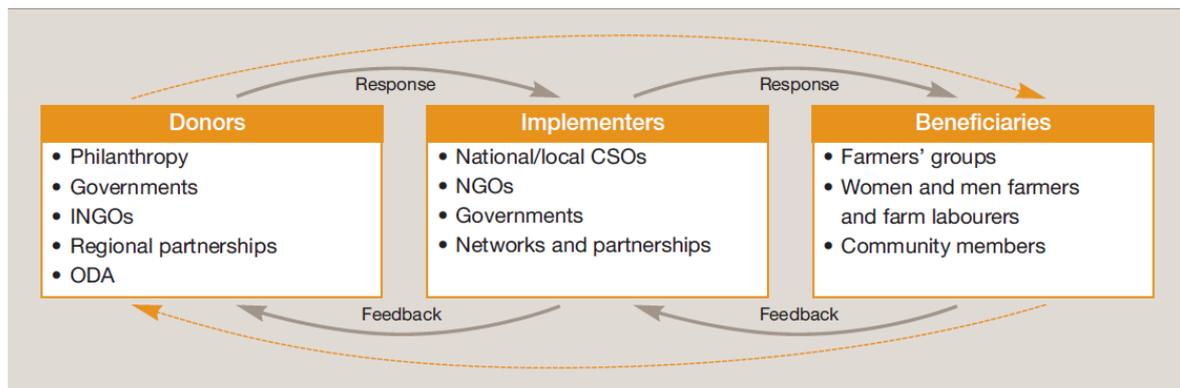
#### 4.1 Balancing Multiple Accountabilities

As we argue above, in order to reap the full benefits of M&E, there is a need to balance multiple accountabilities and align the preferences of all stakeholders. In practice this means greater involvement of beneficiaries in M&E activities, along with all other stakeholders. We argue that greater involvement of beneficiaries in M&E activities is a key leverage point to strengthen M&E provision, in addition to being the 'right' thing to do.<sup>xx</sup>

What does the evidence have to say about the value of balancing multiple accountabilities? In the absence of directly relevant literatures, we focus on papers that ask (a) does participation in project design (upstream M&E) enhance outcomes? and (b) do social accountability mechanisms (downstream M&E) improve outcomes?

In terms of the feedback loop picture below, we are interested in the evidence of whether the arrows going from right to left improve outcomes.

**Figure 4: The Feedback Loops Between Donors, Implementers and Intended Beneficiaries**



### ***Does participation improve project outcomes?***

Mansuri and Rao (2004) provide a review of community-driven projects in terms of the impacts of participation on targeting, service delivery and sustainability. The review finds that targeting of the poor is improved by the participation of intended beneficiaries, but tending to this only when preferences are egalitarian, decision making is open and transparent, and the rules for classifying households as poor are clear. The authors conclude ‘some studies have shown an association between the level of some index of participation and project effectiveness, but the direction of causality is unclear.’ Nevertheless, a number of multi-project studies that pay attention to causality in fact find that participation improves delivery of water projects in Indonesia (Isham et al 1994)<sup>xxi</sup>, public works in South Africa (Hoddinott et al 2001), and the maintenance of rural support programmes in Northern Pakistan (Khwaja 2001)<sup>xxii</sup>. Many more studies are cited which pay less attention to causality and there the results are mixed<sup>xxiii</sup>.

Another review paper by Platteau (2007)<sup>xxiv</sup> summarises the difficulty with these studies: ‘measurement of the impact of participation on development project outcomes is methodologically complex, and, at this stage, there are still few conclusive statements that can be made about the importance and the *modus operandi* of this impact’ (p 38).

The Platteau paper restricts itself to quantitative outcomes in a cause and effect framework. Taking a more qualitative systematic approach a recent paper by Gaventa and Barrett (2010) reviews a non-random sample of 100 case-studies of citizen engagement in 20 countries conducted over a 10 year period by one research programme consortium. It finds that citizen participation produced positive effects across developmental and democratic outcomes in 75 per cent of the outcomes<sup>xxv</sup>. The study highlights the importance of looking at intermediate outcomes; it is difficult to draw a straight line between participation and development outcomes without looking at the process of engagement which might facilitate empowerment outcomes and contribute to development outcomes.

We conclude that there is plenty of evidence that participation can make a positive difference to development outcomes but that it certainly does not have to. The impacts depend on a wide range of context factors, often relating to norms around hierarchy and power.

### ***Do social accountability mechanisms improve outcomes?***

Many mechanisms for making government and private sector actors more accountable to citizens have been trialled over the last decade. These include citizen report cards (CRCs), budget tracking, community score cards, citizen juries, project monitors, community expenditure tracking and social audits of commitments and realities on the ground (Arroyo and Sirker 2005). The studies of their impacts are not methodologically strong, at least in terms of accounting for causality.

Paul (1998) was one of the first to grapple with the measurement of the impact of the citizen report card on changes in the quality and responsiveness of service providers. In the context of evaluating the response of Bangalore's public services to report card implementation, he concludes:

'there is some evidence that public awareness of these problems has increased as a result of the experiment. Civil society institutions seem to be more active on this front and their interactions with public agencies have become better organized, purposive and continuous. As a result, some public agencies in Bangalore have begun to take steps to improve their services. The paper concludes that public feedback ("voice") in the form of a report card has the potential to challenge governments and their agencies to become more efficient and responsive to customers' (p: 2).

Ravindra (2004) also undertakes a review of the Bangalore public service experience with the report card for the World Bank's Operations Evaluation Department. The paper concludes:

'on the whole, the impact of the CRCs has been positive. They helped to increase public awareness of the quality of services and stimulated citizen groups to demand better services. They influenced key officials in understanding the perceptions of ordinary citizens and the role of civil society in city governance. Bangalore has witnessed a number of improvements following the CRCs, particularly the second one. The state government and public agencies launched a number of reforms to improve the infrastructure and services in the city, including via property tax reform through a self-assessment scheme, the creation of the Bangalore Agenda Task Force (BATF), and streamlining of agencies' internal systems and procedures. There is now greater transparency in the operations of government agencies and better responsiveness to citizens' needs. While a number of other factors have also contributed to the transformation of Bangalore, the CRCs acted as a catalyst in the process. ' (p: iii)

Conducting further research on the report card experience in Bangalore and Jaipur, Deichmann and Lall (2003) question the very theory of change of citizen scorecards of service provision. Do perceptions of quality received bear any resemblance to actual quality received? They find that scores are indeed influenced by the quality of services provided but scores are also influenced by a number of household characteristics, including the quality of services provision received by peers.

Brix (2009) assesses that the impact of the use of a citizen scorecard survey helps Chinese citizens influence urban service provision design and implementation, helping policymakers to 'reveal weaknesses and monitor progress in public service delivery' (p: 2)

McNamara (2006) reviews US and developing country evidence on the impacts of health provider report cards on accountability to citizens and in the quality of health care provision. She points out that the idea of provider-specific reports is not new:

'I am fain to sum up with an urgent appeal for adopting this or some uniform system of publishing the statistical records of hospitals. If they could be obtained ... they would show subscribers how their money was being spent, what amount of good was really being done with it, or whether the money was doing mischief rather than good' (Florence Nightingale 1863).

The paper concludes that 'there is evidence that provider-specific comparative reporting, and in particular public reporting, enhances provider accountability and prompts improvements in quality of care.' (p: 106).

One of the most rigorous studies in the classical sense is a randomised field trial of community-based monitoring of public primary health care providers in Uganda. It finds evidence that the implementation of 'citizen report cards' reduced child mortality by 33 per cent as well as generating other health benefits. A year after the intervention, treatment communities were more involved in monitoring the provider and health workers appeared to exercise greater effort in serving the community. The study documents large increases in utilisation and improved health outcomes. Within the experiment, the cost per child death averted was \$300, well below the average of \$887 for 23 other interventions (Björkman and Svensson 2009). This study seems to support our analysis that beneficiary feedback can drive up performance by ensuring that implementers become more responsive to the needs of beneficiaries.

Another rigorous evaluation shows little improvements in public sector behaviour from bottom-up feedback mechanisms. Olken (2007) undertakes a randomized field experiment on reducing corruption in over 600 Indonesian village road projects. The paper finds that:

'increasing government audits from 4 percent of projects to 100 percent reduced missing expenditures, as measured by discrepancies between official project costs and an independent engineers' estimate of costs, by eight percentage points. By contrast, increasing grassroots participation in monitoring had little average impact, reducing missing expenditures only in situations with limited free-rider problems and limited elite capture.' (p: 200)

A randomized control trial by Banerjee et. al. (2008) in India, on community mechanisms to hold primary schools accountable found that giving villagers information about the state of their schools was not enough – it required encouragement and training, in small groups, to turn this information into an intervention that improved learning outcomes. They conclude that 'it seems clear that the current faith in participation as a panacea for the problems of service delivery is unwarranted' (p: 25).

In Kenya, in another randomized control trial, Duflo et. al. (2008) find that hiring supplementary contract teachers may be an effective way of meeting the demand for teachers in Kenya thereby reducing absenteeism and promoting learning. They find that 'the biggest gains come when local school committees are empowered to effectively monitor these teachers and when extra classes are structured so as to target instruction to students' initial achievement level.'<sup>xxvi</sup>

These studies are far from conclusive: some show a positive impact of feedback mechanisms on development outcomes, some show no impact, and some are not set up to conclusively

demonstrate impact, just plausibility. The studies that examine the ‘how’ questions confirm that implementation methods, norms of fairness and expectations matter.

## 4.2 Enhancing organisational learning

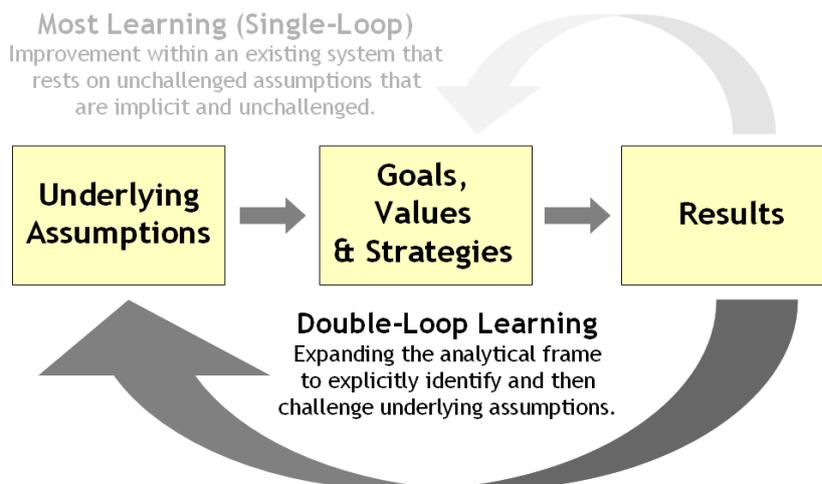
The beginning of section 4 argued that enhancing organisational learning can help strengthen M&E. Learning organisations would, by definition, have a greater capacity to (a) learn from other organisations and therefore identify multiple benefits (b) have a greater organisational capacity to capture these multiple benefits and (c) become more open to others to learn from thereby contributing to the reduction of information gaps between M&E users and providers.

But how are learning organisations built and incentivised? There is a large set of literatures on how organisational incentives and wider institutional incentives affect learning. There are two major strands: on organisational learning (Huber 1991; Easterby-Smith 1997) and on learning organisations (Senge 1990; Roper and Pettit 2003). The paper by Oswald and Taylor (in this IDS Bulletin) reviews much of this literature.

There is a great deal of overlap between the two approaches<sup>xxvii</sup>. Characterising both at their extremes, we can describe organisational learning as understanding how organisations learn about themselves and how they make decisions. Here learning is very much a means to an end – better performance defined as improved project outcomes. At the other extreme, we have learning as an end in and of itself as well as being a means to an end. The focus is more openly normative about the values it espouses (diversity, non-hierarchy, boundary-spanning, critically self-reflective) and, importantly, attempts to link the organisation strongly into a wider social system (Easterby-Smith et al 1999).

Again at their extremes, the two approaches can be characterised as in Figure 5: organisational learning focuses more on single loop learning from results to shaping goals values and strategies. The learning organisation approach focuses more on the double loop learning whereby underlying assumptions are challenged (Argyris and Schon 1978). These distinctions are fluid. As Roper and Pettit (2003) note, the learning organisation approaches can be very pragmatic, emphasising knowledge creation linked to organisational goals, linked to action. Similarly many organisational learning approaches are designed to examine the underlying assumptions and risks about their projects, organisation and context.

**Figure 5: Single and Double Loop Learning**



Source: [http://edbatista.typepad.com/edbatista/images/2006/12/DoubleLoop\\_Learning\\_2\\_Large.gif](http://edbatista.typepad.com/edbatista/images/2006/12/DoubleLoop_Learning_2_Large.gif)

Oswald and Taylor in this Bulletin argue that some of the organisational disincentives that drive a lack of learning from M&E could be changed with self-reporting systems and participatory forms of M&E. Linking such systems to incentive structures within organisations could to be a key starting point (Pasteur and Scott-Villiers 2006).

Can beneficiary feedback support the realignment of incentives within organisations and promote single loop or even double loop learning? Brett contends that ‘because agencies are often providing services (for example, the empowerment of rural women) whose output cannot be priced and whose impact is very difficult to observe....these agencies will perform effectively only where workers are genuinely committed and where managers know that citizens or donors are able to monitor what they are doing, and able to withdraw support should they fail’ (2001: 18). A shift toward such participatory feedback approaches linked to appropriate incentive structures is not easy and it is clear that organisations must understand the potential benefits for their own learning (Abbot and Guijt 1998). Thompson (1995), reviewing the literature on why government bureaucracies in the South have shifted towards participatory approaches, shows there are four main drivers<sup>xxviii</sup>:

- 1) **Expediency:** attempts to do more with less have forced agencies to find new ways of implementing programmes. In the current climate of austerity it is likely that organisations will need to continue to find ways to be more efficient. There is some evidence to show that feedback systems and other forms of participation can be cost effective in terms of identifying rapid corrections to projects (World Bank 2009) and making projects more relevant for beneficiaries (Magione et al 2005), although in certain circumstances they may also increase logistical and transaction costs of monitoring (World Bank 2009; Parkinson 2009). Abbot and Guijt (1998) argue that further study on costs is needed.
- 2) **Pressure from the donor community:** where donors have demanded greater involvement of beneficiaries in decision-making this has been an important factor (Johnson 2001). However, such demand is often accompanied by greater focus on accountability in general, which may

have the effect of increasing pressure on agencies to perform to donor-defined standards (Thompson 1995). Nevertheless, donors are the key agents of change. If they show the kind of leadership that sets up feedback systems for the greater good, even if this may prove inconvenient for them at a later date, then they can begin to create real partnerships.

- 3) Recognition of limits of top-down, blue-print development: these approaches have been seen to be ineffective in addressing the complex realities of poor people (see Rodrik 2008). Widening the set of stakeholders consulted reduces the likelihood of reverting to blueprints.
- 4) Stories of successful application of participatory approaches by other organisations have been convincing in persuading others to try.

For those who advocate greater organisational attention to beneficiary involvement and feedback in M&E, there are a number of drivers to engage with and to leverage. The main driver however will be the answer to the questions: does it work, when and how? As we have seen in section 4.1 the evidence base is patchy. A systematic review, especially one drawing on rich country experience, is badly needed.<sup>xxix</sup>

### 4.3 Building wider learning

Finally, a people centred perspective to M&E can enable the sharing of M&E performance information more openly to support wider learning about what works, how, by whom, and under what circumstances. It does this by creating and sharing public goods. Table 3 shows some of the mechanisms and their hypothesised impacts on identifying benefits, capturing more indirect benefits and learning, and closing information gaps between M&E stakeholders.

**Table 3: Activities to support wider learning**

To:	Sharing data about innovations, programmes, users, providers and tools	Conducting systematic public research on M&E	Creating comparator data
Greater capacity to identify benefits from M&E work elsewhere	X	X	X
Greater ability to claim benefit for more indirect benefits	X		X
Learning about reputations and performance of M&E users and providers	X		X

**Sharing data:** As Thompson (1995) argues, highlighting success stories is one way of encouraging organisations to adopt new practices. Examples include the opening up of the World Bank’s database on development statistics<sup>xxx</sup> and new efforts to make health research data available<sup>xxxi</sup>. Radelet and Siddiqi (2007) show the value of making M&E project data available. They use data from the first 134 evaluated programmes funded by the Global Fund to Fight AIDS, TB and Malaria and

correlate performance grades with programme characteristics. They find, for example, that projects which receive initial 'accept' scores are associated with better performance some years later, suggesting – but not proving – a return to initial investments in thinking through the project theory of change and implementation issues. There are many project databases available which could be brought into the public domain and analysed for public goods purposes and for organisational learning.

**Conducting systematic research on M&E:** There are very few systematic reviews on M&E and even fewer with features of 'people-centred' M&E. One such systematic review was conducted by Miller and Campbell (2006) evaluating the claims made by proponents of 'empowerment evaluation', an approach to M&E which casts the evaluator as an agent of social change. Note that this approach is distinct from our hypothesis about voice, which says that a more balanced set of voices will help establish a consensus Theory of Change and help planning for impact and learning from the intervention, and make grantees and funders more responsive to beneficiary needs. Miller and Campbell evaluate 46 evaluations that call themselves 'empowerment evaluations'. They classified them according to approaches to empowerment (Socratic, structured guidance and participatory) and then they assessed the claims made by the authors against the stated goals of empowerment evaluation. They did not evaluate whether the approach strengthened the impacts of the projects in terms of outcomes such as income. Nearly all evaluations drew on community knowledge but only 19 percent of them reported on evidence based strategies that appreciated the value of scientific evidence. This research clearly shows the value of shining a light on the difference between rhetoric and reality in M&E.

**Creating comparator data:** We argue that to deal with the information asymmetries in M&E provision, there is significant value in collecting comparator data on both M&E suppliers and users. There are many mechanisms that rate non-profits and charities (<http://greatnonprofits.org/> and <http://www.charitynavigator.org/>). There are mechanisms that evaluate private foundations (Center for Effective Philanthropy Grantee Perception Reports) and international networks, social investors and international development NGOs (Keystone Accountability's Performance Surveys). There are indices that rate the effectiveness of aid agencies (Center for Global Development's Commitment to Development Index). There are also programmes that rate think tanks (<http://www.sas.upenn.edu/irp/>). We cannot find any mechanism that rates M&E providers. Nor can we find any site or programme that rates the capacity to use M&E information for improved performance. Both the African Evaluation Association and the American Evaluation Association have publicly searchable databases of evaluators, with neither organisation endorsing the list in terms of quality of provider or performance assessment by third parties. Could this be done through a reputation market?<sup>xxxii</sup> As Jackson (2009) reports, Dellarocus (2002) suggests that online reputation markets are most likely to succeed where traditional trust building mechanisms (e.g. state enforced contractual guarantees) fail. The characteristics of these successful online reputation markets include: lack of repeated interaction (e.g. one off trades are most common), large numbers of small players, geographically dispersed actors, and relationships are built on easily changeable pseudonyms. Jackson notes that these features (with the exception of geographic dispersion) do not seem to characterise the marketplace for agricultural M&E. Instead he suggests effort should be

made to develop criteria for verifying the quality of the information provided by M&E organisations / practitioners to existing databases.

## 5. Conclusions

The context for M&E is changing. Impact and value for money are the *mantra du jour* in the age of austerity. But M&E as we currently know it is threadbare. It does not attract enough investment. It is viewed as an enabler of compliance rather than of competence. When it is done well, it is done to satisfy donors, not intended beneficiaries. Agriculture does not escape this conclusion, and it may even exacerbate it. M&E in agriculture is woeful. Why is M&E so weak? We argue that investment and interest is low because the multiple benefits of good M&E are not identified and when they are, they cannot be captured. The fact that so much M&E goes on undercover allows this situation to persist.

What can be done? We suggest a new type of M&E is needed, one that is people-centred. People-centred in the sense that it focuses on wellbeing outcomes, and in the sense that it asks people about what they need and what they think is working. What are some of the components of this approach? It has three. First, it balances multiple accountabilities through greater participation in programme design, implementation and evaluation. The literature on the impacts of these approaches has grown in the past 10 years and shows more successes than failures. Second, it focuses on enhancing organisational incentives for learning. What needs to change for organisations to engage in single and double loop learning? Beneficiary feedback systems represent one such incentive and new donor requirements would provide another. The third feature of this people-centred M&E is that it seeks to build wider learning about M&E, its users and its providers. The semi-closed nature of M&E is inhibiting the sharing of learning about what works and who does it well.

We believe that M&E in agriculture has to be improved. This paper has presented some hypotheses for how it can be. The evidence is stronger for some of the hypotheses than others. We hope the paper opens up a debate about the best ways to strengthen M&E in agriculture. If M&E in agriculture is not improved then we will have wasted the political opportunity represented by the current high interest in food and agriculture. We will have no excuses when the budget axe is eventually aimed at food and agriculture and we will have failed to meet our obligations to the current and future generations of hungry and malnourished people.

## Cited references

Abbot, J. and Guijt, I. (1998) *Changing views on change: participatory approaches to monitoring the environment*, London: IIED

Argyris, C., & Schon, D. (1978) *Organisational learning: a Theory of Action Perspective*. Reading: Addison-Wesley.

Arroyo, D. and Sirker, K. (2005) 'Stocktaking of Social Accountability Initiatives in the Asia and Pacific Region, Community Empowerment and Social Inclusion Learning Program', World Bank Institute, Washington: World Bank

Ashby, J. (2009) 'Fostering Farmer First Methodological Innovation: Organisational Learning and Change in International Agricultural Research' in Scoones, I. and Thompson, J. (eds) (2009), *Farmer First Revisited: Innovation for Agricultural Research and Development*, Oxford: ITDG Publishing

Banerjee, A., Banerji, R., Duflo, E., Glennerster, R. and Khemani, S. (2008) 'Pitfalls of Participatory Programs: Evidence From a Randomized Evaluation in Education in India' No 14311, NBER Working Papers from National Bureau of Economic Research, Inc

Barder, O. (2009) *Beyond Planning: Markets and Networks for Better Aid*. CGD Working Paper 185. Washington: CGD

Bastoe, P. O. (2006). *Implementing Results-Based Management*. In R. C. Rist & N. Stame (Eds.), *From Studies to Streams: Managing Evaluative Systems* pp. 97-110). London: Transaction Publishers.

Bill & Melinda Gates Foundation Global Development Program (2009) *Agricultural Development: Strategy Overview*, Seattle: Bill & Melinda Gates Foundation Global Development Program, available <http://www.gatesfoundation.org/agriculturaldevelopment/Documents/agricultural-development-strategy-overview.pdf> [accessed 16 July 2010]

Björkman, M and Svensson, J. (2009) 'Power to the People: Evidence from a Randomized Field Experiment on Community Based Monitoring in Uganda', *The Quarterly Journal of Economics*, Vol 124: 2, pp 735–69

Brett, E. A. (2000) 'Understanding organisations and institutions', in Robinson, D. et al. (eds.) *Managing Development: Understanding Inter-Organizational Relationships*, London: Sage for the Open University

Brix, H (2009) 'China Urban Services and Governance', The World Bank East Asia Region Poverty Reduction and Economic Management Department Policy Research Working Paper 5030

Center for Global Development (2006) 'When Will We Ever Learn? Improving Lives Through Impact Evaluation', Report of the Evaluation Gap Working Group, Washington: CDG

Cernea, M. M. & Kassam, A. H. (eds) ( 2005): *Researching the Culture in Agri-Culture: Social Research for International Agricultural Development*, CABI.

Chambers, R. (1997). *Whose Reality Counts: Putting the Last First*. London: Intermediate Technology Publications.

Deichmann, U. and Lall, S.V. (2003) *Are You Satisfied? Citizen Feedback and Delivery of Urban Services*, Development Research Group, World Bank

del Carpio, X. (2009): *Measuring the Impacts of Agricultural Projects: A Meta-analysis and Three Impact Evaluations* May 25, 2009. Independent Evaluation Group, World Bank.

Dellarocus, C. (2002): *The Digitization of Word-of-Mouth: Promise and Challenges of Online Reputation Mechanisms* Centre for eBusiness Working Paper #170, MIT, Cambridge, USA

Duflo, E., Dupas, P. and Kremer, M. (2008) 'Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya', NBER Working Paper No. 14475

European Commission (2007) *The Guide, Evalsed: the resource for the evaluation of socio-economic development*, available at

[http://ec.europa.eu/regional\\_policy/sources/docgener/evaluation/evalsed/guide/index\\_en.htm](http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/guide/index_en.htm) [accessed 13 July 2010]

Easterby-Smith, M. Burgoyne, J and Araujo, L. (1999) *Organisational Learning and the Learning Organisation*. London: Sage.

Easterby-Smith, M. (1997). Disciplines of organisational learning, contributions and critiques. *Human Relations* 50/9: 1085.

FANRPAN (2005): *Report on SWOT, capacity and communication needs of farmer organisations in the SADC region*, available at: <http://www.fanrpan.org/documents/d00049/> [accessed 12 August 2009]

Gaventa, John and Greg Barrett (2010) *So What Difference Does it Make? Mapping the Outcomes of Citizen Engagement*, IDS Working Paper.

Hoddinott, J., Haddad, L., Besley, T. & Adato, M. (2001): 'Participation and poverty reduction,' FCND discussion papers 98, International Food Policy Research Institute (IFPRI). Later published as Adato, M., Hoddinott, J. & Haddad, L. (2005): 'Power, politics, and performance: community participation in South African public works programs,' Research reports 143, International Food Policy Research Institute (IFPRI).

Huber, G. (1991) *Organisational Learning: the contributing processes and the literature*. *Organisational Science*, 2/1: 88-115

IFAD (2002) 'Managing Impact in Rural Development – A Guide for Project M&E', Rome: International Fund for Agricultural Development (IFAD)

International Labour Office (2002) 'Extending Social Protection in Health Through Community Based Health Organizations: Evidence and Challenges', Geneva, ILO, Discussion Paper Universitas Programme.

- Isham, J. & Narayan, D. & Pritchett, L. (1994): 'Does participation improve project performance : establishing causality with subjective data,' Policy Research Working Paper Series 1357, Washington: The World Bank.
- Jackson, C. (2009) 'IPAL Service Provider Online Marketplace Feasibility Report', unpublished report
- Johnson, C. (2001) Towards accountability: Narrowing the gap between NGO priorities and local realities in Thailand, ODI Working Paper 149, London: ODI
- Khwaja, A.I (2001) 'Can Good Projects Succeed in Bad Communities? Collective Action in the Himalayas', Harvard University - John F. Kennedy School of Government, John F. Kennedy School of Government Faculty Research Working Papers Series RWP01-043
- Lele, U., Pretty, J., Terry, E. Trigo, E with assistance from Klousia, M (2010) Transforming Agricultural Research for Development, Global Authors Team Report, The Global Forum for Agricultural Research (GFAR) Report for the Global Conference on Agricultural Research (GCARD) 2010,28th – 31st March, 2010,  
<http://www.fao.org/docs/eims/upload/273759/GAT%20Report%20for%20GCARD%202010%20-%20Version%2011.0.pdf> [accessed 16 July 2010]
- Lindstrom, J. (2009): What is the state of M&E in agriculture? Findings of the ALINe online consultation survey, October 2009, available: <http://www.aline.org.uk/pool/what-is-the-state-of-mande-in-agriculture.pdf> [accessed 16 July 2010]
- Lucas, H. (forthcoming) Evaluation in Theory and Practice, unpublished report
- McNamara, P. (2006); Provider-specific report cards: a tool for health sector accountability in developing countries, *Health Policy and Planning*, 21(2):101-109
- Mansuri, G. & Rao, V. (2004) 'Community-Based and -Driven Development: A Critical Review,' *World Bank Research Observer*, vol. 19, no. 1, pp. 1-39.
- Mansuri, G. & Rao, V. (2006) 'Update Note on Community-Based and -Driven Development,' Washington: World Bank.
- Mackay, K. (2007) *How to Build M&E Systems to Support Better Government*. Washington: World Bank
- Mackay, R & Horton, D. (2003): Expanding the use of impact assessment and evaluation in agricultural research and development. *Agricultural Systems*. 78: 143-165
- Magione, D., Senni, S., Puccioni, M., Grando, S. & Ceccarelli, S. (2005) The cost of participatory barley breeding, *Euphytica* (2005): 1–17
- Meinzen-Dick, R., Adato, M., Haddad, L. & Hazell, P. B. R. (2003) 'Impacts of agricultural research on poverty: findings of an integrated economic and social analysis,' EPTD discussion papers 111, Washington: International Food Policy Research Institute (IFPRI).

- Miller, R. L. & Campbell, R (2006): Taking Stock of Empowerment Evaluation, *American Journal of Evaluation*, Vol. 27, No. 3, 296-319
- Munce, K. (2005): As good as it gets? Projects that make a difference: the role of monitoring and evaluation, University of Western Sydney, Australia
- Olken, B. A. (2007) Monitoring Corruption: Evidence from a Field Experiment in Indonesia. Benjamin A. Olken. *Journal of Political Economy*, vol. 115, no. 2
- Parkinson, S. (2009). Power and perceptions in participatory monitoring and evaluation. *Evaluation and Program Planning*, 32, 229–237.
- Pasteur, K. and Scott-Villiers, P. (2006) ‘Learning about Relationships in Development’, in R. Eyben (ed.) *Relationships for Aid*, London: Earthscan: 94–112
- Patton, M. Q. (1978). *Utilization-focused evaluation*. Beverly Hills, CA: Sage.
- Patton, M. Q. and Horton, D. (2009) Utilization-focused evaluation for agricultural innovation. ILAC Brief No. 22. Rome, Institutional Learning and Change (ILAC) Initiative.
- Paul, S. (1998) ‘Making Voice Work: the Report Card on Bangalore’s Public Service’, World Bank Policy Research Working Paper No. 1921, Washington: World Bank
- Platteau, J-P. (2007): Pitfalls of participatory development, CRED Working Paper 2007/03, Centre for Research in Economic Development, University of Namur
- Power, G., Maury, M. and Maury, S. (2002) 'Operationalising bottom-up learning in international NGOs: barriers and alternatives', *Development in Practice*, 12: 3, 272 — 284
- Radelet, S. and Siddiqi, B. (2007) “Global Fund Grant Programs: Which are Most Likely to Receive High Evaluation Scores?” *The Lancet* 369: 1807–13
- Ravallion, M. (2000): Monitoring Targeting Performance When Decentralized Allocations to the Poor Are Unobserved, *World Bank Economic Review*, VOL. 14, NO. 2: 331-45
- Ravallion, M. (2008): “Evaluating Anti-Poverty Programs.” In *Handbook in Economics 9: Handbook of Development Economics*, Vol. 4, ed. T. Paul Schultz and John Strauss. Amsterdam: North-Holland.
- Ravindra, A (2004) An Assessment of the Impact of Bangalore Citizen Report Cards on the Performance of Public Agencies, ECD Working Paper Series - No. 12, World Bank OED
- Rodrik, D. (2008) *One Economics, Many Recipes: Globalization, Institutions, and Economic Growth*, Princeton University Press
- Roper, L., & Pettit, J. (2003). Development and the Learning Organisation: an introduction. In L. Roper, J. Pettit, & D. Eade, *Development and the Learning Organisation* (pp. 1-21). Oxford: Oxfam.
- Science Council of the CGIAR (2007): *Stripe Review of Social Science in the CGIAR System*, Scoping Paper1, SC Secretariat, available at

[http://www.sciencecouncil.cgiar.org/fileadmin/user\\_upload/sciencecouncil/Monitoring\\_and\\_Evaluation/Stripe\\_Review\\_scoping\\_paper\\_Dec\\_2007.pdf](http://www.sciencecouncil.cgiar.org/fileadmin/user_upload/sciencecouncil/Monitoring_and_Evaluation/Stripe_Review_scoping_paper_Dec_2007.pdf) [accessed 12 August 2009]

Scoones, I. and Thompson, T. (eds.) (2009) *Farmer First Revisited – innovation for Agricultural Research and Development*, Oxford: ITDG Publishing

Senge, P. (1990). *The Fifth Discipline: the art and practice of the learning organisation*. New York: Currency DoubleDay.

Thompson (1995) Participatory Approaches in Government Bureaucracies: Facilitating the Process of Institutional Change, *World Development*, Vol. 23, No. 9, pp. 1521-1554

Victora, C.G (1995) 'A Systematic Review of UNICEF-Supported Evaluations and Studies, 1992-1993', No. 3, New York, UNICEF, Evaluation & Research Working paper Series.

Wood, B., Kabell, D., Sagasti, F. and Muwanga, N. (2008) 'Synthesis Report on the First Phase of the Evaluation of the Implementation of the Paris Declaration', Copenhagen, July 2008, [http://www.diis.dk/graphics/Subweb/paris\\_evaluation\\_web/files/pdf/eng/ENG-Report-chap1-5.pdf](http://www.diis.dk/graphics/Subweb/paris_evaluation_web/files/pdf/eng/ENG-Report-chap1-5.pdf) [accessed 30 July 2010]

Woolcock (2009) 'Towards a Plurality of Methods in Project Evaluation: A Contextualised Approach to Understanding Impact Trajectories and Efficacy', BWPI Working Paper 73, Brooks World Poverty Institute, University of Manchester

World Bank (2005) 'The Effectiveness of World Bank Support for Community-Based and -Driven Development – An OED evaluation', Washington: The World Bank Operations Evaluation Department

World Bank (2009) 'Gender in Agriculture Sourcebook: Module 16: Gender Issues in Monitoring and Evaluation', Washington: World Bank

---

<sup>i</sup> We thank Chris Barnett for his insightful comments. All errors are ours.

<sup>ii</sup> The full speech can be accessed at <http://www.gatesfoundation.org/speeches-commentary/Pages/bill-gates-2009-world-food-prize-speech.aspx>

<sup>iii</sup> We could not find any studies that estimated the value added of strong versus weak M&E. This is not surprising - one would need good M&E to assess the extent of foregone benefits from weak M&E. For such an estimate to be credible each project would have to have two randomly allocated M&E systems in place – a good one that assessed the inputs, outputs and outcomes of the project in a reliable way and one that did not. We could not find an experiment like this and in any case it is ethically questionable.

<sup>iv</sup> This is why we resist the temptation to delink monitoring and evaluation. As two sides of the same coin they both contribute to the multiple purposes of M&E.

<sup>v</sup> The suggested corrective measures in this article emerge mainly from a partnership between IDS and Keystone Accountability to strengthen M&E in agriculture development. Keystone has been a pioneering advocate for systems of M&E that are grounded in dialogic learning for improvement between development agencies and their intended beneficiaries.

<sup>vi</sup> After a decline over the past 30 years, over the period 2003-08, bilateral aid to agriculture increased at an average annual rate of 13% (real terms) according to the OECD. See <http://www.oecd.org/dataoecd/54/38/44116307.pdf>. A rapid assessment by FAO in 2009 of financing to

agriculture in SSA (based on data from the OECD, major bilateral and multilateral financing agencies, as well as two private foundations active in the agriculture sector), concluded that ODA to the agriculture sector has increased and there is more financing for agriculture in SSA than usually assumed if aid flows from private foundations are included. There are also a number of new commitments for agriculture have been made by various important donors (World Bank, EU, AfDB, IsDB and IFAD). See <http://www.fao.org/docrep/012/al144e/al144e.pdf>

<sup>vii</sup> Note that statistical analysis of the stakeholder data did not show any significant associations between answers and participant characteristics.

<sup>viii</sup> [www.povertyactionlab.org](http://www.povertyactionlab.org)

<sup>ix</sup> <http://www.3ieimpact.org/page.php?pg=round1>

<sup>x</sup> This is a self-selected perception survey and so will be subject to bias. We invited people to take part in the following way: ‘We are writing to ask for your input to help shape the design of a new initiative to improve M&E practices in agriculture.’ It may be the case that those who were more likely to respond to the survey felt that M&E in agriculture is weak, although there is no ex-ante reason to expect that the relative weights given to any one cause of its weakness would be biased. A multiple regression analysis of these data showed that NGOs as a group tended to rate M&E in agriculture more strongly than other groups.

<sup>xi</sup> This is similar to other frameworks, such as the purposes identified by the European Commission (2007) (see Lucas and Longhurst in this IDS Bulletin) and McKay (2007).

<sup>xii</sup> Woolcock’s review (2009) suggests that this dimension of project performance – impact trajectories – is chronically underfunded. Different types of projects will have different types of trajectories. The impacts of conditional cash transfers may be quite proportional to the resources poured into them. But the same may not be the case for projects with a long gestation (e.g. science and technology interventions) or projects that rely on a certain intensity of participation (e.g. market access projects or technology adoption) or projects where adaptation to the intervention can occur (e.g. pest control). Randomisation *per se* does not help here because the impact estimated will depend on when the follow up surveys are conducted. What is needed is an explicit time dimension to the theory of change and the use of monitoring data to track impact trajectories, especially when there is significant behaviour change involved.

<sup>xiii</sup> He notes that randomisation for impact assessment will not, *per se*, help with the context issue, because those selected to receive a ‘treatment’ (i.e. participate in the programme) will not be the same as those who seek to participate in a programme in practice.

<sup>xiv</sup> For example: a systematic review of the United Nations Children’s Fund (UNICEF) estimated that 15 percent of all its reports included impact assessments, but noted that ‘[m]any evaluations were unable to properly assess impact because of methodological shortcomings’ (Victora 1995). Similarly, a review of 127 studies of 258 community health financing programs found that only two studies were able to derive robust conclusions about the impact on access to health services (ILO 2002).

<sup>xv</sup> Their discussion is primarily focused on the incentives for organizational learning and thus primarily concerned with implementing organizations, but these are mirror the overall mis-alignment of incentives in the aid system (Barder 2009).

<sup>xvi</sup> The dilemma of the role of implementing agencies in this situation is the argument behind the concept of the alien-hand syndrome (Power et al 2002). They argue that the success of the implementing agencies is dependent on satisfying the donor, not the community it serves. This disconnect results in ‘success’ – that is empowering communities and giving them voice – possibly being detrimental to the success of the implementing agency because it creates a direct accountability link which may threaten the organisation’s methods, mission or focus if a community has the ability to question the organisation’s choices or approaches. Power compares non-profits to the private sector where there is an incentive to engage in organisational learning that values the knowledge and experiences of the customer, as this will help determine success. The dilemma is unique to the aid system – in countries whose public services are financed by taxation, those who deliver public services are accountable to politicians, who in turn are accountable to tax payers – the same people who are meant to benefit from public services (Barder 2009).

<sup>xvii</sup> Complex is here distinguished from complicated. For a more detailed discussion, see Rogers 2008

<sup>xviii</sup> Lucas and Longhurst in this IDS Bulletin agree that this is certainly the case in comparison to health where there is a wide variety of well understood basic health interventions that are generally regarded as both effective and inexpensive.

<sup>xix</sup> del Carpio (2009) notes how difficult it is to undertake impact evaluations in agriculture and how hard the evaluators have had to work to construct a plausible comparator group, often through the innovative use of non-M&E data.

<sup>xx</sup> In many ways this argument is more instrumental than the participatory M&E advocates would be comfortable with. Beneficiaries have access to knowledge, expertise and relationships that other actors do not. But these views may not be the most accurate in the system, nor the most cost-effective to access, nor any less politically-motivated or subject to capture than other actors in the system (see for instance Parkinson 2009). As such our approach can be seen as a combination of a rights and a results-based approach. See Jacobs and Jacobs et al. in this IDS Bulletin for a detailed discussion.

<sup>xxi</sup> The study provides strong statistical evidence that increasing beneficiary participation leads directly to better project performance – a 10 per cent increase in participation of the rural poor in these projects resulted in a 2 per cent increase in overall performance.

<sup>xxii</sup> The Khwaja study is methodologically the most rigorous, comparing a random selection of community driven projects with similar but non-community driven projects in the same village. This concludes that community participation significantly improves outcomes when the community is involved in non-technical decisions, but worsens them when the community is involved in technical decisions that require expert knowledge.

<sup>xxiii</sup> An update note by the same authors a few years later (Mansuri and Rao 2006) is noteworthy because there are so few additional studies on the impact of participation on project outcomes (there are a few more on the distribution of benefits which give similar results to the initial paper).

<sup>xxiv</sup> This study is suspect because of the way in which they report the same Hoddinott et al. paper (2001). Platteau states ‘Hoddinott et al. (2001) (who) studied the effects of participation in public work programs in the Western Cape Province in South Africa. Their results indicate that participation has no effect whatsoever on any of the (employment) outcome variables that they have considered.’

<sup>xxv</sup> Gaventa and Barrett note that 25 per cent of the outcomes in the sample of case studies were negative.

<sup>xxvi</sup> See <http://www.povertyactionlab.org/evaluation/peer-effects-pupil-teacher-ratios-and-teacher-incentives-kenya>

<sup>xxvii</sup> This review is based on an earlier version of Oswald and Taylor in this IDS Bulletin.

<sup>xxviii</sup> See Jacobs et al in this IDS Bulletin for a further discussion on the drivers of adopting participatory M&E.

<sup>xxix</sup> As Bonbright and Power note in their article in this volume, the business sector has long established the correlation between customer satisfaction (the business analogue to beneficiary participation in M&E) and performance improvements, profits and shareholder value. Businesses invest billions of dollars in eliciting customer feedback annually.

<sup>xxx</sup> See

<http://web.worldbank.org/WBSITE/EXTERNAL/NEWS/0,,contentMDK:22547256~pagePK:64257043~piPK:437376~theSitePK:4607,00.html>

<sup>xxxi</sup> See <http://www.who.int/bulletin/volumes/88/6/09-074393/en/index.html>

<sup>xxxii</sup> Online reputation markets trade access to reviews of goods and services that rank and comment on quality.