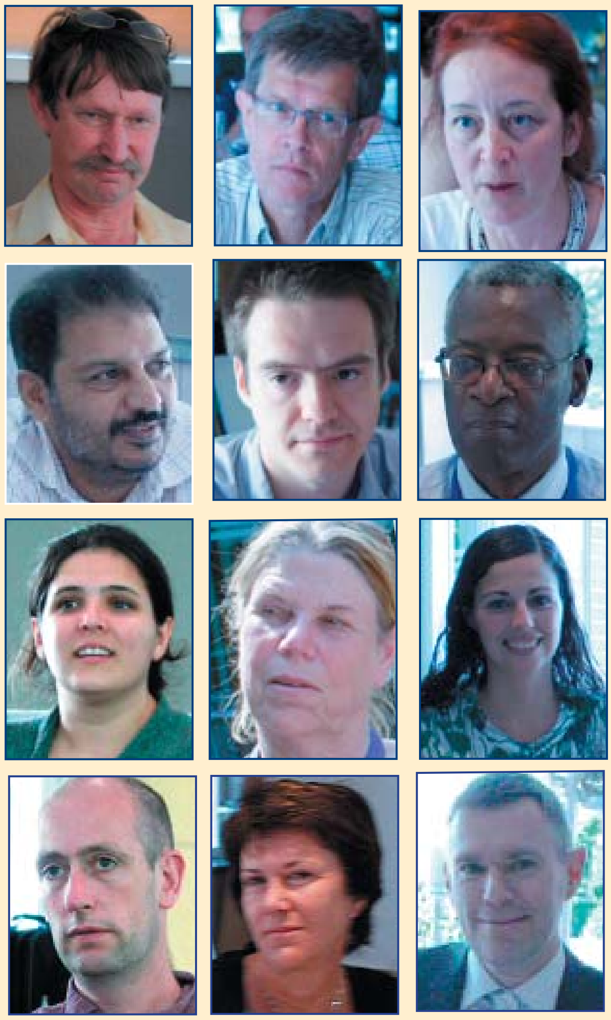
Mapping the Measures of Success

13-14 May, 2008

De Lindenhof, Delft



Mapping the Measures of Success — Who was there?

Name	Organisation
Jerry Adams	WaterAID
Teun Bastemeijer	Water Integrity Network
Kristof Bostoen	LSHTM / Independent Consultant
John Butterworth	IRC International Water and Sanitation Centre
Rachel Cardone	BMGF
Rick Davies	Independent Consultant, UK
Catarina Fonseca	IRC International Water and Sanitation Centre
Hester Foppen	ICCO
Ben Haagsma	ICCO
Carol Howe	SWITCH project
Harry Jones	ODI
Stephen Mogere	M&E Consultant, JICA
Patrick Moriarty	IRC International Water and Sanitation Centre
Murad Mukhtarov	IDEAS
Donald O'Leary	Transparency International
Sompit Punyaratabandhu	Office of Agricultural Economics, Thailand
Ratna Reddy	Centre for Economic and Social Studies, India
Richard Rheinghans	Emory University, United States
Pablo Rodriguez-Billela	RELAC, Argentina
Alistair Sutherland	National Resources Institute, UK
Rita Tesselaar	DGIS IOB
Dick C. van Ginhoven	DGIS
Rutger Verkerk	IRC International Water and Sanitation Centre
Denis Zoungrana	2IE
Round Table Coordination Team	
Deirdre Casella	IRC International Water and Sanitation Centre
Rinske Koop	IRC International Water and Sanitation Centre
Frank Little	Bureau Frank Little
Sandra Segura	IRC International Water and Sanitation Centre
Joep Verhagen	IRC International Water and Sanitation Centre

Report prepared for the IRC by Peter McIntyre, UK

Executive summary

An increasing number of development interventions in water, sanitation and hygiene (WASH) are focused on improving the governance of services and changing the way that institutions interact with each other and with communities. Such interventions are based on the premise that improving governance, particularly at the level of service planning and delivery, will lead to improvements in services and in the lives of community members. Strengthening the capacity of institutions to identify gaps in services, define quality standards, improve planning and learn from experience is seen as an essential step towards long-term sustainable services at scale.

The challenges of evaluating such interventions are considerable. It is difficult to demonstrate a clear and direct link between an intervention to improve capacity for planning and learning at the level of a district planning team with health and livelihood benefits in communities. Projects that focus on knowledge, learning, planning and communication cannot simply be evaluated by counting pipes and pumps. At the same time donors and project designers need to know about the impact of such interventions, both so that success can be replicated and so that projects that are missing their targets can be put back on track.

The IRC International Water and Sanitation convened an expert Round Table on May 13th and 14th 2008, focused on impact evaluation (IE) for strengthening governance of WASH services. The purpose of *Mapping the Measures of Success* was to identify existing practice, experiences and challenges related to evaluating the impact of development interventions in the context of governance of WASH services and institutional change. The aims were facilitate learning and sharing among WASH sector practitioners and impact evaluation experts from the South and North and identify the needs of stakeholders for relevant and appropriate methods and tools for measuring impact.

Identifying the challenges

The urgency of the need to better measure the impact of interventions and to be accountable for money that is invested in development was acknowledged. It was acknowledged that several valid definitions of impact evaluation exist, one working definition that was tabled defines it as “the analysis of the relationship between an intervention and final indicators (outcomes and impacts)” (World Bank).

However, the difficulties in achieving this are evident. Projects to improve governance typically operate in three to five year cycles, while different outcomes and impacts happen at different points in time – some during, and some long after the project. Benefits in terms of services and livelihoods may become measurable long after the project is over. The desire for quick evaluation means that the longer term impacts, and the effects on sustainability, are often neglected.

Even without time constraints, it is difficult to establish the impact on communities of a research project because the linkages between governance and sustainable services take place in subtle and less visible ways. There are unresolved questions about how to understand the links between the intervention and the final impact, and even about what the final impact is understood to be. These questions include: “What do improvements in governance in the WASH sector look like?” and, “Is it safe to assume that good governance leads to good WASH services?”

The question of attribution is a particular problem when interventions take place at a modest scale and there is other activity in the sector. Some longer term outcomes have to be modelled because the link from intervention to outcome is too distant. For example, in a school-based intervention it is possible to measure improved knowledge and uptake of hygiene practices, but the longer term effects on educational achievement or levels of diarrhoeal mortality in the community cannot be directly measured because of the long time lag and problems of attribution.

Even if changes are identified and can be attributed to the intervention, it may be unclear whether they are context-specific or can be replicated in a different setting.

Methods and indicators

A number of methods were explored for evaluation for development initiatives. The strengths and weakness of the Joint Monitoring Programme (JMP) for the Millennium Development Goals were discussed. Other methods examined included outcome mapping, social network models, an approach based on sustainable rural livelihoods and qualitative information systems (QIS). These were all felt to be useful in different ways – as management tools, in understanding relationships between groups, or as precursors to measures of impact on sustainability and equity – but none fully address the final impact of projects and interventions. However, it was felt that they may have extra value when used in combination, according to the indicators selected for monitoring.

Conclusions

The absence of easy or obvious solutions suggests that a variety of approaches may be needed to identify measures of success in terms of outcome and impact. A number of conclusions were drawn by working groups in the Round Table.

Clarity of purpose

Given that projects may be based on different models of change and evaluation may be done for different reasons, there is a need to identify what is being tested in the intervention and what the evaluation is setting out to show. Understanding why interventions are taking place is important, because underlying assumptions shape our understanding of impact. The intervention needs to have a clear and coherent strategy, with objectives and a planned means of getting there. Ensuring a good match between the intervention's theory of change and the impact evaluation methods selected will determine the rigour of the evaluation.

Impact evaluation may be done to demonstrate upward accountability to donors or downward accountability to beneficiaries. Donors may have a variety of objectives, from comparability measures for investments in the sector, to establishing value for money, selecting the best methods of WASH governance or simply showing that the project met its targets. Evaluation might address a number of different questions:

- Do improvements in governance provide measurable benefits to users?
- Did this project provide measurable (expected or unexpected) impact to beneficiaries?
- Did the project provide improvements in governance?

Progress towards institutional change and good governance may also be seen as a function of management for results, with a strong learning dimension geared towards measuring improvements in project and organisational performance.

Indicators of successful governance for WASH

There is a need to be clear about what burden of proof would be acceptable to demonstrate that a "good governance" process is working. The challenge is to compile an evidence base to connect demonstrations of improvements in governance to long-term impacts such as equity of benefits or sustainability.

Outcome indicators could comprise those that measure sustained functioning, access for all and use by all. Impact indicators may include social, economic and hygiene/health impacts such as time saving, changed behaviour, and reduced risk of death and disease.

Impact evaluation is not only about end-points. There is a complex, interlocking and often unpredictable cause-effect relationship between outputs, outcomes and impacts. Different indicators can be monitored at different points in the project and intermediate effects or outcomes may be steps on the road to final impact.

But is improved governance itself an impact, or is it a means to an impact? Should we be looking at the impact of governance on the institutional performance of service organisations rather than trying to measure impact on services? And what information is needed to help communities improve governance, given the gaps between what is established by legal and regulatory reform and practice on the ground, and lack of clarity about roles and responsibilities. We may only be able to measure the possibility that what we have done might lead to the outcome, especially in programmes with a short time frame. However, the evidence will eventually be found in good services and responsive organisations. It can be argued that if a water facility delivers technically sound and sustained water and sanitation services, then the quality of institutions and governance is implicit. On this basis, the following 'global' factors for good quality governance of WASH services were suggested:

- Equitable representation of the interests of the different end-user categories in services and processes,
- Equitable management of water resources (across uses and areas),
- Equitable provision of basic needs (water, food, sanitation, housing, education, health), with no undue competition between (sub)sectors
- Transparent use of investments and resources
- Good quality goods and services at equitable prices without leakage of funds or hidden extra payments
- Accountability of service providers for use of resources and service management.
- Improved demand-responsiveness of services,

Corruption can also be seen as a powerful indicator of poor governance. The Global Corruption Report 2008 has established that corruption in the water sector is widespread and preventing corruption is one of the main conditions for good governance.

Incentives for evaluation

Moving from outputs to outcomes creates risks for projects due to fear of 'failure'. Government agencies may see evaluation as 'fault finding' rather than an aid to managing change. Incentives are needed, rather than a climate of punishment, even when results are less than hoped for.

What to do with the results?

Lack of information can be seen as another symptom of weakness of governance. Data collected during evaluation need to be accessible and useful to stakeholders, including local managers, for improving the project, not lost in a 'data cemetery'. As much energy should be put into making use of data as goes into gathering it.

What follow up should come from this Round Table?

The Round Table proposed a number of ways in which these issues can be pursued and how links between those working on these problems can be strengthened. Proposals for carrying forward the discussion were accepted with developments posted on the impact evaluation weblog at http://washimpacttable.wordpress.com/. An IRC Thematic Overview Paper (TOP) on the topic will be published in late 2008 synthesising and building on some of the best practice available.

Ongoing collaborative activities will include:

- A discussion forum for minimum indicators.
- On-line email list about evaluation of partnerships
- The development of some examples of network learning or district level intervention to articulate outcomes and impact indicators.
- The development of a model of what good governance at district level looks like, including the intervention logic for partnerships and changes one might see after different periods of time.

Bartering before the Open Space session — participants decide what sessions they will attend to focus on the most important issues in impact evaluation.

Introduction

The IRC International Water and Sanitation Centre convened an expert round table in May 2008, related to evaluating the impact of development interventions in the context of governance of WASH services and institutional change. The two-day meeting aimed to map existing knowledge, practices, experiences and challenges and was attended by experts from academic departments, government agencies and and think tanks, as well as by development experts with experience of conducting evaluation in the field. All the experts are 'hands-on' and the round table focused on the practice of evaluation. It also addressed some theoretical difficulties such as how to identify indicators for interventions designed to improve governance of WASH, achieve sustainability and take successful interventions to scale and how to deal with difficulties of attribution.

Evaluation challenges have long been recognised as difficult, both conceptually and in terms of implementation. There were no unexpected breakthroughs in terms of solutions, but the round table did succeed in clarifying the challenges that experts encounter. The absence of obvious answers can itself be seen as a signpost for those undertaking evaluation, suggesting that a variety of approaches may be needed to identify progress at the levels of process, outcomes and impact.

Participating experts expressed an interest in continuing to develop sets of indicators, frameworks and guidelines to achieve rigorous and valid evaluation of institutional change and governance.

Motivation and outputs

A discussion note circulated before the meeting described the increasing demand from different stakeholders, such as policy makers, implementers, donors and politicians for reliable information on the impact of programmes focused on institutional change, and how this demand was being extended from other sectors to interventions focused on strengthening WASH governance and institutions.

The aims of the round table were to encourage:

- Peer sharing among impact evaluation experts to inform an initial mapping of the knowledge, skills and methods available
- Facilitate learning and sharing among WASH sector practitioners and impact evaluation experts from the South and North
- Capture the prevailing needs and interests of stakeholders for relevant and appropriate methods and tools for measuring the impact of interventions focused on institutional change.

In particular the round table addressed the following questions:

1. Why do we do impact evaluation? Who is the consumer of impact evaluation results and why do they need them?
2. What are the main challenges when trying to identify and measure impact of an intervention focusing on development or institutional change?
3. What can be defined as meaningful measures or indicators of success?
4. What suitable methods of impact evaluation are available to evaluate a project or programme that seeks to strengthen local governance?
5. What limitations and gaps exist in the current impact evaluation discourse?

The ambition of the round table was to facilitate the analysis of the current gaps in Impact Evaluation, by discussing the main characteristics and difficulties of interventions focus on

Institutional Change, and the methods and indicators that have been used for measuring its impact. Identifying these gaps will hopefully lead to a better match of impact evaluation approaches to the work being conducted, to avoid expenditure of resources on measuring the wrong things.

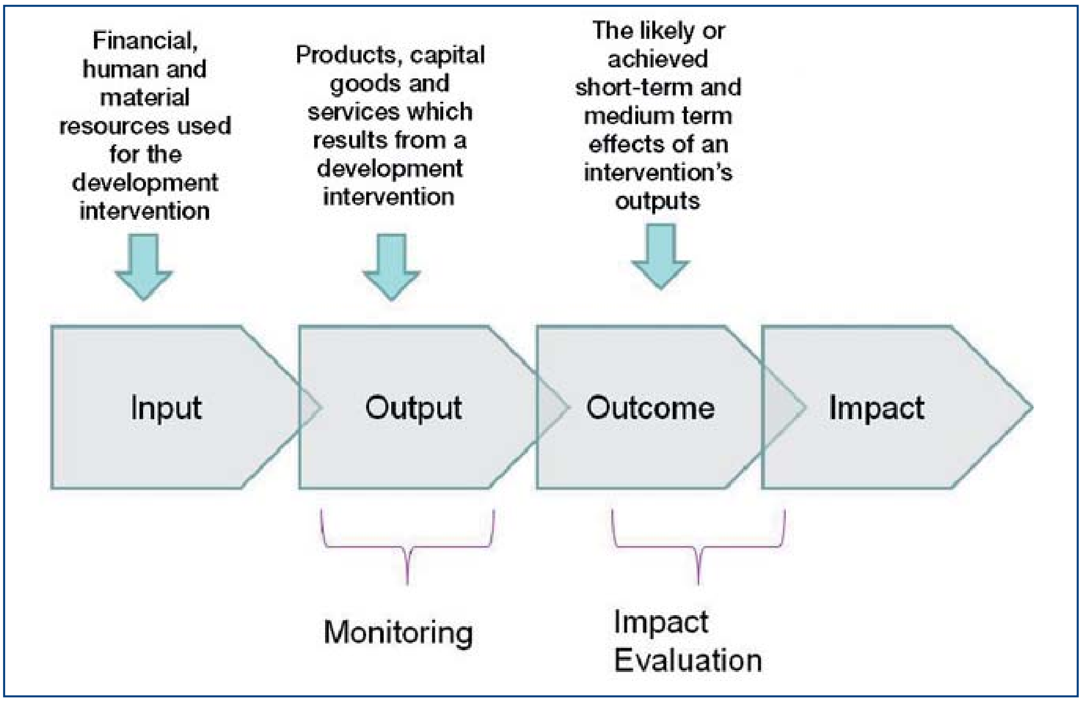
Method of working

A discussion note helped participants to focus on areas of most interest.The round table was constructed around four sessions designed to start from the experience of participants and work towards identifying future methods of work. It began by looking at current experiences, ideas on theories of change in projects, and interventions to bring about institutional change. The second session looked at some methods currently in use for measuring outcomes and impact in projects. The third session was built around two case studies looking in particular at indicators that had proved useful (and less useful). The final session opened the way for participants to address their own most urgent questions around evaluating projects aimed at improving WASH governance or bringing about institutional change.

Outputs

There are five outputs from this round table:

- This report in Word format and as a pdf version (this version)
- A fuller account of the meeting (in Word) available on request
- An executive summary report of about 1,600 words (already included in this version)
- An open discussion public weblog at http://washimpacttable.wordpress.com/ set up to encourage exchanges of information and opinions. This will remain active through September 2008 and beyond, depending on demand.
- An IRC Thematic Overview Paper (TOP) on the topic which will build on the outcomes for the round table and be published on paper and online in late 2008.



Intervention chain described in the discussion note prepared for the Round Table

Session 1

Setting the scene

The purpose of the first session was to create a common starting point for discussions, acknowledging differing definitions and experiences of impact evaluation and governance-focused interventions. The main questions addresssed were:

1. Why do we do impact evaluation? Who is the consumer of impact evaluation results and why do they need them?
2. What are the main challenges when trying to identify and measure impact of an intervention focusing on development or institutional change?

Hopes and expectations

In opening the round table, Ben Lamoree IRC director, reminded participants of the need to develop knowledge and share experiences: "The need to better measure the impact of interventions has become more and more urgent and called for in the light of political accountability of donor organisations and result orientation of actors in the sector." This need was underlined as participants explained what they hoped to get from the meeting.

Ben Lamoree: "The need to measure impact has become more urgent"

One challenge is that projects typically operate in three to five year cycles, but the benefits of improving governance in terms of better services may take much longer to become apparent and measurable. Meanwhile those conducting projects and programmes need to know whether they are on track or how they need to change direction. Interventions may also have effects that are broader than the project objectives. How will those be measured especially in projects about 'softer' issues involved in WASH governance, such as improving information and knowledge, planning, decision making and communication?

The question of what donors want from impact evaluation was also an issue. They might be seeking to establish comparable measures to judge investments in the sector by, or to find the best methods of WASH governance or simply want to know if a project or programme has been well run and met its targets. Often donors are seeking to link costs to impacts, to assess the value for money of interventions, and (under some definitions of impact evaluation) to demonstrate whether projects eventually show an economic rate of return on the investments. The cost of evaluation, and who pays for it, is also a factor.

Information collected during evaluation needs to be accessible and useful to different actors, and it was said that as much energy should be put into making use of this data as goes into gathering it. Data should not only be useful to donors and planners, but also to local managers who need to see if they are on the right track.

Voices of experience

This articulation of experiences, issues and concerns was followed by four interviews, conducted by Joep Verhagen, which sought to highlight the diversity, and complexity, of intervention theories that exist around issues such as capacity development, institutional change and strengthening governance.

Stephen Mogere

Stephen Mogere, who advises the Japanese donor agency JICA on evaluation in Kenya, underlined how motives may diverge even between partners. JICA see evaluation as a need to establish the cost effectiveness of their support, but Kenyan Government partners may see it more as an audit. There was a need for the joint team to avoid approaching evaluation as a fault finding exercise, and for government officials to see it as an aid to managing change. One area of JICA support is directed towards capacity strengthening, but it has proved difficult to link training at top level institutions to changes at community level.

Sompit Punyaratabandhu has 20 years experience in evaluation in the Office of Agricultural Economics, in Thailand. There is often pressure to show outcomes quickly, as a project is coming to its close, before problems of sustainability have emerged. She would like more opportunity to look at impacts over a longer time frame to discover, for example, whether farmers still had access to water and an improved income many years after an irrigation project had been implemented.

Donal O'Leary, from Transparency International, works with the Water Integrity Network (WIN), raising awareness, networking with 600 member organisations in 90 countries, supporting projects undertaken by southern NGOs and providing capacity development. He emphasised that donors who support this work are very keen to see if they are getting value for money. IE is crucial for following the impact of reductions in corruption on the sustainability of water supply and sanitation and improvements in people's health.

Donal O'Leary

Patrick Moriarty, from IRC, talked about the difficulties in measuring outcomes from the EMPOWERS project to support the improvement of Integrated Water Resources Management (IWRM) by strengthening local governance Egypt, Jordan and Palestine. The objective was to create a self-reflective learning process through learning alliances to engage service providers and policy makers at the meso level and link them with problem solving at community level. However, impact would be felt over a decade or so, and the project life was much shorter. It would also be very difficult to separate out the impact of this project (in one governorate in each country) when so much else was happening the sector. 'Evaluation' often simply assesses whether a project has completed the tasks it has set itself, rather than measuring impact. Process documentation carried out for the project was probably more productive in tracking changes in thinking and interactions.

Patrick Moriarty

Summarising the challenges

Clearly, the need for impact evaluation is based upon a variety of priorities. The four interviews highlighted a number of challenges faced by a range of stakeholders in different contexts and settings. Although the theory of change, or intervention logic, differed from programme to programme, a paramount concern of the interviewees is the need to ensure impact evaluations and their outputs are reliable, useful and contribute to learning not only about 'what works', but, perhaps more importantly, 'why something works' and how to 'do things better' in future.

Following these contributions, participants outlined the challenges that need to be met, if impact evaluation is to be effective, meaningful and useful in improving the quality and effectiveness of interventions. These can be summarised as follows:

What you know about the intervention?

- The intervention needs to have a *clear and coherent strategy* if impact evaluation is going to be useful – with objectives and a planned means of getting there.
- Understanding why interventions are taking place is important, because *underlying assumptions shape our understanding* of impact.

What are the challenges? Denis Zoungrana of 2IE, Burkina Faso and Sompit Punyaratabandhu from the Office of Agricultural Economics, Thailand, give their views.

What you want from the evaluation?

- *Reasons for evaluation differ*, including upward accountability to donors or downward accountability to beneficiaries. IE can evaluate whether the underlying hypothesis of the intervention is valid, or look at whether the intervention is working. ('Are we doing the right things? or 'Are we doing things right?').
- There are several definitions of impact evaluation; those embarking on IE should be clear about *which definition they are using.*.
- *Learning how to do things better* is a key purpose.
- Impact analyses can be perceived as audits you have to do for someone else. How can you *change mindsets* so that stakeholders become engaged in the process and see it as a process of *learning rather than 'fault finding'*?
- Donors fund the process and do not want bad news. This can lead to *biased reporting*. This was challenged. There was a counterview that donors have good reasons for wanting evaluation, that donors themselves are accountable, and they should not be stereotyped as unsympathetic to project realities.

What are the problems about measuring impacts?

- There are *problems of attribution* – how do you know what is the impact of your invention and what is the result of other inputs? The project does not control the sector and has a limited time frame.
- There is a question over whether *control (counterfactual) groups* are possible or useful for interventions directed at governance and institutional change.
- It is hard to get *evidence based documentation*.

What can you measure?

- Impact can be measured as *sustainability in local governance change processes*.
- If you are seeking institutional change then the evidence will eventually be found in *good services and responsive organisations*. In the shorter term, it is possible to monitor intermediary outcomes.
- We may only be able to measure the *possibility* that what we have done might lead to the outcome, especially if the programmes have a short time frame.

- Impact evaluation in *not only about end-points*: different indicators can be appropriate at different points in the project.
- Every change recorded during a project can be seen as *a step effect* on the way to a final impact.
- *Unplanned outcomes* (positive and negative) should also be measured.

What do you do with the results of evaluation?

- Impact evaluation should be ongoing and *fed back into improving the project*; not a post-mortem.
- Results should be *accessible*, not end up in a 'data cemetery'.

Session 2

What methods do we use to measure impact?

There is an ongoing discussion not only about whether impact evaluation should be conducted, but also what might be the 'most suitable' methods for measuring impact of interventions of all varieties. This debate especially is heard in the development cooperation sector where the main distinction is the difference between 'experimental' methods, referring to randomised controlled trials, and 'quasi-experimental' methods which consists of all other methods. The objective of the second session of the round table was to create space for learning and sharing about a range of methods that have been applied to interventions focused on various aspects of institutional change, to derive lessons from these experiences (recognising that it would not be possible to present an exhaustive selection of available impact evaluation methods).

Presentations on evaluating specific development initiatives

Five short presentations therefore outlined methods of evaluation for specific development initiatives.

Kristof Bostoen

1. Dichotomy between global and local monitoring

Kristof Bostoen, London School of Hygiene and Tropical Medicine (LSHTM)

Kristof Bostoen outlined the history of global sector monitoring and in particular the Joint Monitoring Programme (JMP) for the Millennium Development Goals. His theme was that global monitoring and local monitoring require different approaches in methodology and use. JMP data is more robust at global level than at local level but its data is often misused for national planning. His conclusion was that there is no such thing as the holy grail of a universal indicator that is useful at a national and local level. Data collected for local sector programmes require more detail on the level of access, causes, consequences and the enabling environment.

Harry Jones

2. Focus of Outcome Mapping

Harry Jones, Overseas Development Institute (ODI), UK

Harry Jones put the case forward for outcome mapping, against the caricature of using log frames etc. His main message was that that measuring impact is often not realistic in complex development projects and there is a need to adjust planning, monitoring and evaluation accordingly. Outcome mapping recognises the limits of the influence of projects and is sympathetic to the difference between attribution and contribution.

3. Social Network Models of Process Design

Rick Davies, Consultant , UK

Social network models are derived from social network analysis and are about better representations of what you try to do and what happens subsequently. They have some overlap with outline mapping. The assumption is that if we have better models, (i.e. better descriptions of what we are trying to do) then it will be easier to assess impact. Social network models help actors to conceptualise what is happening in relationships between local organisations.

Rick Davies

Rick Davies manages the Monitoring and Evaluation News website, http://mande.co.uk/

Ratna Reddy

4. Sustainable Rural Livelihoods Approach

Ratna Reddy, Centre for Economic and Social Studies, India

Changes in livelihoods can be understood as changes in financial capital, physical capital, natural capital, social capital and human capital. Strong sustainable livelihoods occur when these are constant or grow. Measuring changes gives an idea of the impact of interventions on livelihoods. The Sustainable Rural Livelihoods Approach seeks to integrate sustainability into monitoring frameworks, offering a broader definition of impact than simply focusing on income generation or employment.

Deirdre Casella

5. Qualitative Information Systems (QIS), Poverty- and Gender-Sensitive M&E for Improved Service Delivery

Deirdre Casella, IRC, The Netherlands

The Qualitative Information System (QIS) is a flexible system to capture and manage qualitative information which can be tailored to suit local conditions and needs, yet still generate uniform and comparable qualitative information that can link assessment, learning and action in a loop for project management and communities. What is innovative about the method is its ability to include demand-responsive service and gender considerations and equity considerations, such as user participation, in the same framework as issues of system quality, financing, management and institutional and policy support. Indicators underpinning the picture of sustainability can be tailored to match the logic of an intervention. Qualitative as well as quantitative information can be captured and managed thereby making the results useful at project and community level as well as in terms of policy influence.

Conclusions

These presentations were examples of the kinds of methods available for evaluating a variety of outcomes and relationships. In the discussion that followed, it was agreed that although they are all useful in different ways, none fully addresses the impacts from projects and interventions. Outcome mapping can be seen as a very useful management tool rather than an impact evaluation tool. It is not clear, for example, how to use outcome mapping to measure whether a culture of learning has been embedded long term into a sector or institution. Social network models were seen as useful in understanding how relationships between groups and organisations match with household level changes. It could be used in conjunction with a household survey to explain outcomes. QIS was seen as useful for revealing information about outcomes that are precursors to saying something about the impact on sustainability and equity. The methods could usefully be used in combination, each more appropriate for evaluating different indicators, for different reasons.

The discussion brought out a number of different understandings of 'impact evaluation'. The critical point was raised that the various definitions need to be articulated more explicitly. This may shift the focus of discussions about appropriate methods away from seeking to identify or achieve unanimity about a "universally accepted best method" and towards greater clarity about the underlying needs and expectations of a particular evaluation. Greater clarity about the boundaries will make it possible to better address the different needs of different stakeholders for information about long-term outcomes and about intermediate impacts or outcomes.

Session 3:

What do we want to measure?

By the end of session 2, participants had begun to clarify some of the issues and problems around evaluation and had become more aware of the difficulty in moving in practical ways from monitoring outcomes to impact evaluation. The objective of this session was therefore to look in more detail at what kind of indicators can be used focused on project outcomes. The discussion was based around two keynote contributions.

1. Measuring Success of School WASH:

Evaluating Impacts or Evaluating for Impact

Rick Rheinghans Hubert Department of Global Health Emory University

Rick Rheinghans

Rheinghans described a school based water, sanitation and hygiene project in Western Kenya seeking to determine whether schools are an effective, sustainable and scalable point of entry for improving WASH. The project will test the hypotheses that schools are an important entry point for the community and that children can be agents for behaviour change in the household.

The team at Emory University is working on how to evaluate this five year project for impact, with project partners, CARE, Water Partners International (WPI), Sustainable Aid in Africa (SANA International) and Great Lakes University (GLUK).

The evaluation team will look at the incremental benefits of various school based approaches, looking at:

- Improved knowledge and uptake of behaviours such as water treatment and handwashing
- Diffusion of behaviour change into the community
- Health impact in the schools and the community
- Educational impact in terms of absenteeism and test scores
- Cost effectiveness.

Long-term outcomes such as diarrhoeal mortality, educational impact and increases in productivity cannot be measured and will need to be modelled. The evaluation will also seek to assess the potential for sustaining and scaling school based WASH interventions.

Rheinghans concluded that moving from measuring outputs to measuring outcomes creates risks for projects and fear of failure Incentives are needed to undergo evaluation, rather than a climate of punishment.

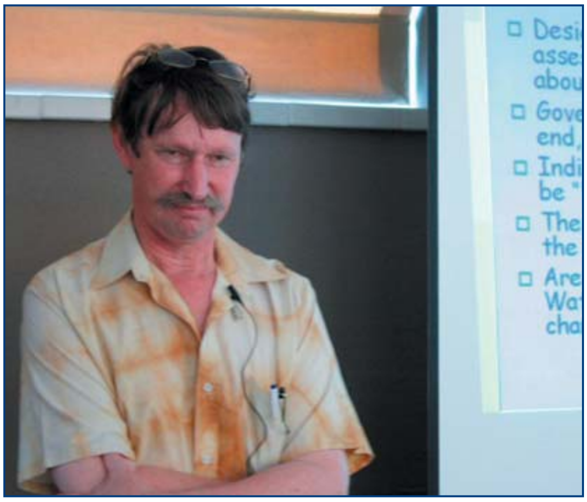
"We cannot make any inference from this study about impact but it does tell us about the things that need to be in place in order to have impacts.."

Rick Rheinghans

2. Identifying Indicators

Example from Agricultural R&D and thoughts about WASH Governance Indicators

Alistair Sutherland, National Resources Institute



Alistair Sutherland

The NRI was funded by DFID to evaluate a 15-year agricultural programme in Kenya and Tanzania, which was seeking to bring research results into use by farmers through an extension programme. It was difficult to show the impact of research, before the results were in use. Participatory action research was undertaken with three teams (two in Kenya and one in Tanzania) over two years, working at district or regional level.

Some of the issues were about who sets the agenda for research, equity of access to research findings, value for money (cost-effectiveness), and the problems of lesson sharing if policy makers are not engaged.

Sutherland noted that simple indicators – such as the age and gender of those taking part in an initiative – could tell a story about coverage and inclusiveness. Knowledge attitudes and practice (KAP) surveys were also useful. Indicators worked best when they were based on a clear model of change linked to a specific intervention. Indicators generated by communities tended to emphasise material benefits as outcomes rather than abstract advantages, like greater knowledge. Researchers need to spend time with communities if they are to capture learning and acquire contextual information to make sense of the data.

Learning points

- If the main aim is to look at effectiveness of different approaches to knowledge sharing a lot of time is needed with communities working through how they are going to capture that learning.
- You need contextual information to make sense of the data.
- Donors do not want to wait 4-5 years for answers so it is important to identify leading indicators of impact which can be given quickly.

Implications for WASH and Governance

Both presenters reflected on some of the key issues evaluating interventions focused on WASH governance arising out of their own experiences.

Rick Rheinghans concluded that the challenge is to compile an evidence base to connect demonstrations of improvements in governance to long term impacts such as equity of benefits or sustainability. He suggested that evaluation might address a number of different questions:

- Do improvements in governance provide measurable benefits to users?
- Did this project provide measurable impact to beneficiaries?
- Did the project provide improvements in governance or is likely to do so?

Studying the indicators of success — Top, Rita Tesselaar, Catarina Fonseca and Rutger Vervenk. Below, Stephen Mogere and Ratna Reddy

Alistair Sutherland posed some similar questions for impact evaluation when investigating interventions concerned with WASH governance:

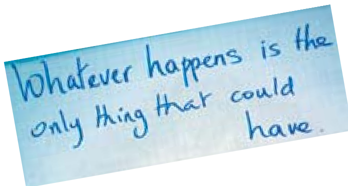
- Is it safe to assume that good governance leads to good WASH services or do we need to ask about the conditions under which this will happen?
- What do improvements in governance in the WASH sector look like?
- How can the results (intended and unintended outcomes and impacts) of improvements in governance be measured and assessed in a cost-effective way?
- What exactly do we mean by governance? Governance issues need to be unpicked and stated clearly in the project strategy.
- Should we be looking at the impact of governance on the institutional performance of service organisations rather than trying to measure impact on services?
- Is improved governance an impact, or a means to an impact? If it is a means, then the indicators of good governance are more likely to be "leading" indicators of impact – ie. indicators that show you whether you are moving in the right direction.

These presentations and the discussion that followed highlighted the need to develop indicators which could be seen as reliable predictors of impact, or be intermediate outcomes leading to the final impact ("leading outcomes").

Session 4:

Added value of impact evaluation

What does this all mean?



The rules of the Open Space session:

- Whoever comes are the right people
- When it starts is the right time
- Whatever happens is the only thing that could have
- When it is over, it is over.

Open Space Session Topics
1. What should we be measuring for WASH Governance?
2. Identifying indicators of governance for WASH services
3. Achieving sustainable and scaleable impact
4. Management for results
5. Impact assessment as a process. Making data live after the event.
6. How can corruption be a measure of outcomes towards impact of governance?
7. How can partnerships be monitored and is this important?
8. What follow up should come from this Round Table?

Open Space

The final session was an Open Space opportunity for those attending the round table to set their own agenda, to pick up the most critical issues and work on them. Each participant decided what they wanted to focus on and participants signed up to the sessions that most interested them. A number of critical issues were addresed in the groups

Summaries of group discussions

What should we be measuring for WASH governance?

What should we be measuring for impact evaluation for WASH governance, when we are developing methodological change for good governance at district level?

Getting to grips with 'impact' is a particular challenge in the context of learning project when the primary unit of intervention may be one to three districts, and the objective is to test, develop and start to prove tools and methods.

IRC's WASHCost project seeks to collect and make available data on whole-life unit costs for WASH interventions, working at global, national and district levels in four countries. It will seek to develop tools/approaches/methods to be used in decision making in a small number of pilot districts/programmes.

Open Space session discussing what to measure for WASH governance — Christine Sybesma, Kristof Bostoen, Ratna Reddy and Patrick Moriarty.

The heart of the discussion was what combination of research tools (including IE) can be used to best understand the change process being catalysed, and the real changes this leads to in service delivery to users. There are several issues to consider:

- Different outcomes and impacts will happen at different points in time – some during, and some long after the project.
- The sphere of intervention and theory of change: Where do we intervene and how do we understand the links between the intervention and final impact? What do we understand the final impact to be?
- Attribution: How far is our intervention responsible for changes seen?
- Scalability: Are the changes context-specific or can they be replicated in a different setting?

One aim of WASHCost is to improve governance processes, including planning, budgeting, transparency and user satisfaction. Outcome mapping could be used to test the theory of change in the districts. Some form of quantitative impact assessment can also be used. How to do it requires further work. There is a need to be clear about what burden of proof would be acceptable to demonstrate that the process is working.

The extent of complexity and uncertainty means that establishing a control (comparison) group would neither be possible nor meaningful. A better control would be a before and after comparison within a district.

The WASHCost project will use a facilitated learning alliance, consisting of stakeholder platforms at national and district level. Evaluating the ability of the learning alliance to deliver change will be an interesting challenge.

Identifying indicators of Governance for WASH services

What are the key governance factors that influence outcomes and impacts of WASH services?

The starting point was a suggestion to look especially at factors for resilience, which help recovery rather than account for failures. WASH governance indicators are difficult to standardise in widely diverging and changing settings. However, indicators are needed for three levels of governance: national (policy and law making setting), intermediate level (dealing with implementation), and user level (sometimes conflicting interests in communities and households).

Outcome indicators comprise those that measure sustained functioning, access for all and use by all. Impact indicators include social, economic and hygiene/health impacts such as time saving, changed behaviour and reduced risk of death and disease.

Impacts include those from programme processes, as participation in planning and management organisations and decisions, various types of training, etc. have an impact in addition to the impact from improved infrastructure.

The following 'global' resiliency factors for evaluating governance of WASH services were identified:

- Equitable representation of the interests of the different end-user categories in services and processes,
- Equitable management of water resources (across uses and areas),
- Equitable provision of basic needs (water, food, sanitation, housing, education, health), with no undue competition between (sub)sectors
- Transparent use of resources
- Everyone gets good quality goods and services at equitable prices without leakage of funds or hidden extra payments
- Service providers account for their use of resources and service management.

Achieving sustainable and scaleable impact

What essential questions should we be asking to achieve sustainable and scaleable impact with respect to WASH governance?

Despite legal and regulatory reform, there are large gaps between the framework (decentralisation, private sector participation, clarity in roles and responsibilities) and practice. Learning questions include:

- How do people practice the law in different contexts?
- How can demand-responsiveness be improved?
- What information does a supporting organisation need to help communities improve governance?
- How can we learn from positive and negative experiences?

Ideas put forward included:

- There are taboos in the WASH sector about the role of the private sector and financial sustainability, which are out of sync with the views of users on ability or willingness to pay.
- Evaluation of WASH governance is challenging in a sector where broader management of water resources is fragmented.
- Nobody funds research to deconstruct 'horror' stories in the sector or to research why things go 'right'.

Factors to protect against risk (factors for resilience) can be strengthened through demand responsive approaches in a decentralised environment. This requires flexibility and openness to a variety of approaches, which supply-oriented stakeholders often lack.

Questions about local government capacity centred on accountability (to whom and for what?), capacity building (for what? how to sustain over time?) and financing (what financing systems for planning, budgeting, and reporting?).

Management for results — Ben Haagsma and Carol Howe

Management for results

How do we manage for results for institutional change and good governance?

Management for results is geared towards constantly improving project and organisational performance, rather than towards hard scientific evidence of results. It implies a strong learning dimension, based on collecting information continuously on emerging changes at beneficiary (or client organisation) level. It promotes greater ownership by staff, downward accountability towards clients and upward accountability to donors. It promotes greater sustainability of results.

There is no simple time related sequence between outputs, outcomes and impacts, but rather a complex, interlocking and often unpredictable cause-effect relationship. Changes can be seen as effects, outcomes and/or 'impacts', which are indications of the final impact, but emerging much earlier. Information generated by this internal M&E mechanism can be usefully fed into external evaluations.

It can be argued that if a water facility delivers technically sound and sustained water and sanitation services, according to contextual standards, then the quality of institutions and governance is assured.

Impact assessment as a process – Making data live after the event

How do we keep valuable data up to date, relevant and in use?

Impact assessments happen as an event, but impact takes place over time. Data therefore needs to be kept alive, up-to-date and accessible, rather than being consigned to the straightjacket of a report ('data cemetery'!).

Lack of information can be seen as a symptom of weakness of governance. For data to be kept up to date it was essential to have the trust and buy-in of key players. Means of locating existing data were given e.g. www.alexa.com (the Wayback Machine) and MandE News. NONIE (Network of Networks on Impact Evaluation) is a tool for sharing information.

Jerry Adams of WaterAid — looking for ways to keep data alive

Corruption

How can corruption be a measure of outcomes towards impact in governance?

The Global Corruption Report 2008 has established that corruption in the water sector is widespread and preventing corruption is one of the main conditions for good governance. Impact may be achieved by diminishing the amount of bad water governance and water services management can be improved through preventing corruption.

- Corruption risk assessments can be made and linked to incentives to make measures to prevent corruption attractive for stakeholders.
- More use can be made of existing lines of action to do with access to information and reporting complaints and tools such as metering systems for irrigation
- Corruption requires secrecy, so measuring transparency and access to information can provide a useful indicator. Budgets need to be presented so that civil society organisations and other stakeholders can understand them. Corruption and transparency should also be taken up in evaluations
- The short term risks should be acknowledged – breaking traditional patterns of corruption or patronage may diminish the sustainability of services in the short term.

Combating corruption — Teun Bastermeijer from the Water Integrity Network

Monitoring and measuring in partnerships for research and development

How can partnerships be monitored and is this important?

With an increasing number of partnerships for research projects, and as many organisations adopt multiple roles, evaluation of partnerships is becoming more important. These may be formal or informal arrangements for a common purpose and time. There is a need to be clear about the value added from partnerships, and also the risks. Differences of power between partners should not be allowed to override the principle of working together. If a project is being implemented by several agencies, it is important to be clear about who will do what.

For evaluation in the context of strengthening organisational capacity – how can we know when an organisation's partnership capacity has improved? Indicators can be decided on at a diagnosis workshop with key partners. If potential partners are reluctant to engage, their attendance at a workshop can be a progress marker.

Rick Davis agreed to set up and moderate an e-group – using Google Groups – to develop the concept of evaluating partnerships.

What follow up should come from this Round Table?

The inputs to this open space session were sorted into two main categories:

Collaborative activities

- A discussion forum on minimum indicators.
- On-line email list about evaluation of partnerships established via MandE News

It's a round table — with facilitatrs Joep Verghagen (left) and Frank Little and organiser Deirdre Casella

- Harry Jones will adapt existing work Improving IE for Coordination & Uptake to be WASH sector specific.

John Butterworth suggested:

  - some examples (WIN network learning & WASH district level intervention) are developed that articulate what expected outcomes and impact indicators could / should include.
  - Two or three projects develop a model of what governance at district level could look like. John will write a description of what networks and partnerships could look like for WIN, including the intervention logic and what should be measured at different points, and changes one might see after different periods of time up to ten years.

- Rick Rheinghans will post a model of a district level project on the website, and encourage a range of experts to take part in a dialogue about it.
- WIN is interested in taking a strategic lead for its members on participatory community management & other tools that can be adapted to examine social auditing. WIN could also prepare corruption risk analysis for WASH projects and provide an explanatory note on how to undertake such an analysis.
- IDEAS has an upcoming conference on governance –WASH people should be in Quebec in October 2008. IDEAS will explore a follow up Conference on WASH sector evaluation.

Joint knowledge and information management products

- Track the impact of the Open Space process on communication between people after the workshop.
- The weblog should continue, so as to include perspectives from those not attending the round table.
- NGOs need to have common access indicators clearly spelled out. The link between access indicators and institutional change also needs to be articulated.
- A briefing note is needed, detailing gaps in knowledge and the interests expressed.
- Translate the executive summary of the final report into other languages.

Sandra Segura, reporting back from the blog — at http://washimpacttable.wordpress.com

Finally

With the weblog still live and a number of commitments for following up this work, the round table clearly provided an impetus to efforts to clarify issues around both outcome evaluation and impact evaluation in relation to WASH governance and institutional change. Evaluation of the round table showed that participants had valued the open and friendly style of the discussion and would have liked to spend more time on these issues. The round table served as a significant opening to the discussion, yet one that will need to continue, with concerted efforts to engage other stakeholders in the dialogue.

We did it our way. Participants use the talking stick to channel the discussion