

We are in good shape and a practical M&E-system is in reach for many of us.



Handout -notes for my presentation at the DCED Seminar on Trends and Results in PSD, Session 3, Tuesday 17 January 2012, Peter Roggekamp

As donors, consultants and implementers in private sector development we have come a long way over the last few years in developing a workable and realistic monitoring system that potentially produces credible data and that is useful as a management quality system. Under the umbrella of the DCED Standard the outline of a practical and credible monitoring and impact reporting system has evolved. Many programs are now implementing a system that is based on impact logics, impact chains, result chain or whatever people call it. Some project just started, others have been operating for a few years.

Where initial benefits of working with impact logics and applying the DCED Standard is now obvious for many, there are still valuable lessons to be drawn how to get more out of your M&E system. This can be done by integrating a DCED Standard based M&E system better with management structures in your organization.

This note lists some thoughts and is hopefully a basis for further discussions. Result chains and monitoring plans have gotten most of the attention over the last few years. This note is looking at the next step, building a full M&E management system around the logics and plans.

A note on terminology:

Long discussions can be held over the best terminology to distinguish between different functions of a system that improves designs, that monitors activities, outcomes and impacts, that analyzes these internal or external, that draws lessons, that implements these lessons and that produces all kinds of reports on impacts. In this note I try to explain that we only get to the final goal of improving interventions by understanding the real impact and to the goal of producing better impact figures if **all** elements of the system function well.

I also realize that behind every term there is a view on how a system should look like (e.g. Evaluation should be external). Getting terminology right is not the purpose of this paper; therefore please allow me to avoid this issue by inventing a term, just for this paper, while leaving the discussion to those who know better. Let me arbitrarily call it a QA&R system, encouraging you to paste the most appropriate terms. Please also forgive me when I use terms like “impact”, “evaluation”, “assessment” in a different way than you think is correct.

A bit of background first.

Market Development or Private Sector Development creates special challenges for assessing impact. We normally do not know who exactly the final beneficiaries are. We also do not know upfront exactly what our activities will be. We are expected to capture impacts on farmer or enterprise performance or even on poverty where the impact chains are complex and it often takes years for higher levels of impact to materialize.

There have been several attempts from donor organizations to design a standard QA&R system; not with much success. With every new attempt, I assume that most team leaders must have mumbled things like: *“These guys in headquarters are way too academic; they do not understand our reality”*.

Attempts to fly in international experts also have not lived up to the expectation. The independent QA&R consultants were probably able to capture elements of impact, but the socio-economic reality takes more time and tools than 2 rounds of surveys to crack. Good research has to be based on a good initial understanding. After many months or even years, the project staff may start to grasp what is really going on. Outsiders to the project, to the value chain or to the country are never given enough time to really understand the reality. Even if hiring external specialists to do the QA&R

job would work, it would only be affordable for the few large and well-funded programs.

In 2007, with these conclusions in the back of their mind, a small group decided to come up with something that should be a doable, practical QA&R system. It was intended to develop a QA&R system that:

- should be based on emerging experiences with impact logics (result chains) in a few projects,
- should be tested and adjusted by projects without initial donor involvement. The outcome does not need to be a perfect system, but a doable, a realistic one
- should be implemented by the people who understood the reality best: the project staff, but
- that would have some kind of external quality control.
- a system that would be affordable to smaller projects and organizations as well.

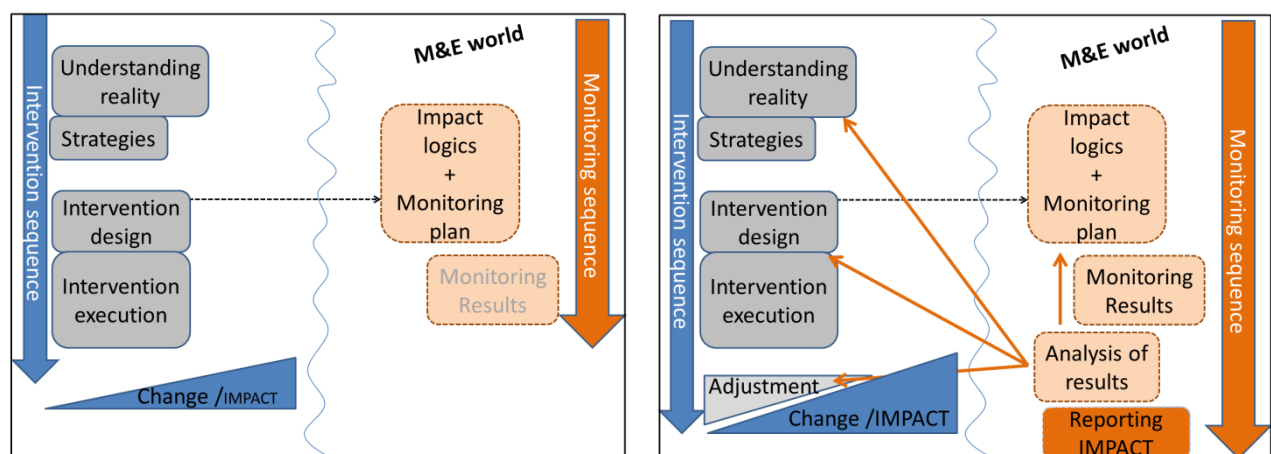
Supported by the Donor Committee for Enterprise Development, 6 or 7 projects started applying a first version of a quality standard that included not only a compliance document (*“the Standard”*), but also a system of external independent auditors.

Now, 4 years later, the system has gained momentum. A much larger group of donors and projects have joined the movement. Specialized consultants have been trained and a number of audits have been conducted.

Many projects, programs or organizations have developed a first round of impact logics and integrated this in their project. I’m sure everybody found result chains a great tool for designing interventions and the DCED Standard helpful in this. I’m also sure that the monitoring sheets are a good basis to build a full QA&R system around, a system that produces not only credible data, but also functions as an internal quality management system. The last step is to build a full management system around the result chains and monitoring plans.

Spot the differences:

The figure below may help explaining the benefit of a full QA&R system versus only



using result chains and monitoring plans to generate data.

The sequence on the left is a simplified linear overview of steps in market development projects. When a project knows what it is going to do, it can start developing impact logics and monitoring plans. Though useful in itself, a good functioning system need other steps also well developed. It also needs:

➤ Effective monitoring.

This sounds like an open door, but is not. What indicators work? Can you really measure changes in yields or income? Are the research companies you hire any good? What is the optimal sampling? What are the right tools for each step to keep the system both manageable and credible? Courses can teach the basics, but it will need trial and error to get this one right in a specific situation. Someone needs to drive this!!

➤ Analyzing the results.

This is of course a real joy for academics! Having loads of data, case studies, observations, expert opinions, focus group results and external reports is unfortunately often the nightmare for project staff. Translating information from different sources, that often conflicts, into conclusions that make sense is difficult. But this triangulation is really at the heart of a functioning QA&R system and in the projects I worked on something very appreciated by the staff.

➤ Adjusting interventions and strategies based on early monitoring results

Projects adjust interventions when the interventions do not work. The changes are often triggered by observations of staff or managers. This is a good first step, but not the same as a solid QA&R system where all data is well analyzed and leading to a constant adjustment of strategies and interventions.

➤ Reporting results

Donor organizations cry for reliable impact data that is produced by a credible QA&R system. Unfortunately not too many organizations are able to produce these data yet. Part of the problem is that it often takes years for impact to happen. Only when the first round of impact data is produced, it becomes clear where the QA&R system is still weak and where it needs more attention.

A good number of organizations have given these issues more thoughts and produced manuals that can help others.

The best lessons come from mistakes.

I tried to make a long list of issues that can go wrong, based on what I experienced in the projects I worked on or what I saw in some other projects. All contain lessons how to do better and all can be avoided I think:

- *Only opportunistic commitment from managers.*
QA&R cannot be delegated to a dedicated monitoring expert in or outside the project. It is a management tool and the people who decide what the project does, should also drive the QA&R system.
- *No one takes the daily decisions.*
AQ&R is not only a science, but has many elements of engineering with daily decisions. The DCED standard and good manuals can reduce the decisions, but still there is a daily need. If managers are not highly involved there is a good chance that small decisions are never taken. Daily decisions cannot be left to external consultants.
- *We just monitor outcomes, others measure impact.*
This is where many may differ in opinion. I think that a program needs to monitor at least up to changes in practice of entrepreneurs or farmers. I doubt if an external impact monitoring contract ever worked in PSD.
- *No clear messages from Managers to staff that QA&R is necessary to be a successful organization; that QA&R is not the activity with the lowest priority or that QA&R that is not just needed to please donors.*
Most people do not like QA&R, “it is scary”, “a lot of paperwork” and “does not contribute to impact”. With these thoughts among the staff, it is unlikely for QA&R system to work.
- *Ability to design good research and an ability to constantly adjust the system to a system that works and that remains reliable.*
We are unfortunately not yet in a situation where experienced people can design a fool-proof QA&R system; a system that is likely to function without adjustments. This is not different from other management systems, but it needs to be done in house and driven by one of the managers.
- *Finding the right core indicators and developing realistic ways how to measure these.*
Many indicators like change in yields or employment seem good candidates but they are not always very practical. Key indicators should not be selected, but developed and tested early on in the project. This is doable, but it needs attention.
- *Management systems where senior management of the project is not part of the evaluation.*
Good lessons from QA&R will only lead to a good management system if the real decision makers are fully involved in the triangulation sessions and if the project design is flexible enough to adjust. With triangulation sessions I mean meeting with all staff involved, where progress and changes are discussed

based on all the available monitoring data and other observations. After triangulation sessions, strategies should be adjusted and interventions changed if needed.

- *Management system where the “Doing QA&R” later is accepted.*
In most projects, time is not on our side. Understaffed projects that need to meet deadlines tend to postpone QA&R a bit. There are always more urgent things.
- *Enough funding and staff*
Even though QA&R is likely to improve the activities so much that it pays for itself, organizations still need to be allowed to set aside resources for QA&R. Surveys can be a bit expensive. All staff should probably spend 20% of their time on thinking about impact chains, monitoring and adjusting. Dedicated QA&R expertise should be in house. (10% of the staff?)
- *Segregation between implementation staff and QA&R team*
This probably does not need much explanation, but the more efforts to avoid this.
- *Availability of outside research capacity for larger surveys.*
It is not always easy to get reliable data from surveys that are outsourced. These surveys must be of good quality to be useful. In too many countries the majority of the survey companies serve their clients with what they want to hear or just lack the people to do a proper job.
- *No culture of internal criticism, one where failure of activities is really OK.*
And linked to this: a “spirit of honest inquiry”. For both staff and managers there are very good argument not to be too critical of internal failures. If a culture of internal criticism is not established, good QA&R as a management tool will probably never flourish.
- *Unrealistic or unclear expectations between donor and implementer.*
Donors are good in putting projects under pressure for data that the project just does not have and probably cannot have at that stage of the project with threads like: “I need to know this impact next week or.....” Being forced to lie is not a good basis for honest data. Complex log-frames that do not reflect the program too well also do not help.

All organizations suffer to some extent from some of the above. The list is really more a checklist to identify areas for improvement for existing programs or things to consider for new organizations. Donors can make elements of the list part of the agreements and implementers can integrate some of the lessons in the offer, in the inception phase and in the first one or two years of the program. I’ll try to chalk some ideas of some of the things both implementers and donors could do, but first **some overall lessons:**

- *QA&R should be integrated* with the other management systems in the program.

- *Big boss needs to drive QA&R* and really implement management systems that incorporate monitoring outcomes in decision processes. This need to be enforced.
- *Management actively needs to develop AND MAINTAIN a culture of honesty* and QA&R. A negative culture is easily established and hard to reverse. For us managers the benefit of not telling everything is only too clear. The benefits for staff to stay quiet are often less obvious, but in many cultures a critical attitude is not seen as a virtue but rather as lack of respect. Staff is also often a bit blinded by trying to appear successful and the consequences of this for his or her salary increase and career.
- *Key indicators and the aggregate indicators need to be developed early in the project.* Developed means that they need to be tested on feasibility, reliability, replicability and significance.
- *The system needs permanent maintenance* to keep the right balance between *simple and credible*. There needs to be a driver of this and this has to be a senior manager.
- *All professional staff should be involved* with clear roles and responsibilities.
- *Most programs will also need external support.* There is not that much expertise available at the moment so some efforts are needed to get the good ones.
- *Starting early on with periodically triangulation processes.* A triangulation culture needs to be established and to have it as part of the management systems, it should be developed and tested early on. You will not get it right the first time.
- *Early on there should be a realistic agreement with the donor* on what level of impact data can be expected and when.
- *As QA&R is always the least urgent activity to do,* it is necessary to build in events in the calendar where one catches up. Where one updates result chains, monitoring plans, strategies, reports and checks that all planned research is actually done.

When designing a program the following tips could help:

You cannot have it both ways. If there is a need for credible impact figures and a desire for the project or program to have a mechanism that leads interventions to success, you have to make it part of the design and budget. Implementers normally have no choice in accepting QA&R requirements and may think: “*Just say yes and we deal with it later*”. So the QA&R requirements need to be realistic with the right resources available. Proposals could also contain a specific plan to have a system up and running early on while donors have a monitoring check list that early on checks if all elements of the system are in place. More specifically:

Resources:

Just as a first rough idea, a program is unlikely to be successful in QA&R if fewer than 1 in 10 staff is dedicated to QA&R. The other professionals may spend 10 to 20% of their time on thinking through impact chains, on monitoring and on evaluating the

finding. There must be enough money to conduct some surveys and budget lines for training and external support.

Monitoring:

For a 4 to 5 year program, the QA&R systems should be up and running after a year. After a year, a QA&R monitoring mission could check:

- Is the team leader or key manager driving the process or is it left to a more junior person?
- Is there a manual that is known by the staff?
- Are there good strategies of the markets, sectors or value chains?
- Are there impact logics of the main activities? Are these impact logics backed by solid market assessment / research?
- Are the accompanying monitoring plans realistic and are the resources to monitor in place?
- Have the key indicators been developed and tested and is it feasible to monitor these?
- Can all staff explain what is expected of them and is there a culture of being critical?
- Has all staff been involved in triangulation meetings and are there good records of this?
- Is there an agreement between donor and implementer of when what kind of impact data will be available?

The DCED Standard audits have most of these elements in it, but go further with issues like a well-organized system of data collection. I would advise donors to make this more transparent in the contract with the contractor.

For so long implementers could claim that there is no suitable QA&R system to apply. We now have an opportunity with the new DCED Standard to turn the tide. This system may be far from perfect, but there is not too much else that has a good chance of success.

For implementers:

Much is said above. I will add some loose comments:

- ❖ *“Activities first, then QA&R” will not do it.* It should be: Strategies, QA&R system, staff, culture and indicators first, then activities impact logics and monitoring.
- ❖ *Build in regular QA&R weeks where staff triangulate,* update the paperwork and recommend changes to the value chain / market strategies and the implementation.

- ❖ *I can see good change, but “no impact yet”* is not a problem. Impact takes time and the donor may need to be convinced upfront about this. Agreeing when donors get what kind of data, based on a realistic schedule, is essential. There is another blessing: projections. Donors are often much helped by early realistic projections of the kind of impact that might be achieved. Impact logics can help with this. There is also a great risk in this. I’ve added some thoughts that could reduce some of the risks: Only predict for the donor 70% of your actual projections to be on the safe side; aggregate the impact as much as you do not know who the real winners and failures in the project will be; Communicate in writing while very clearly saying that these are early projections with all the limitations and only give these to your donor. There is a to take figures out of context where projections and actual impact data can be mixed up.
- ❖ Most externally hired short term consultants will says they *can do it*. This does not mean that they will come up with something reliable. You have to be in charge and you should not outsource blindly.
- ❖ Measuring reduction in poverty levels, increase in income, company profits or even yields and attributing this (plausibly) to your activities is not too easy for most market development projects that do not know upfront with whom they work and who will benefit. The constant efforts to keep QA&R simple enough to implement while credible enough to sell means that you really need to sort this out early on. The table below gives some ideas how it can be done. This is stuff for discussion.

Stage of impact chain	Measuring how	Measuring what
1) Activities (outputs)	Easy...observing	What really happened
2) Changes in the support systems	Much observation, small simple surveys	Data on outreach, indications of sustainability (investments, profit etc) and attribution. Signs of indirect effects on markets.
3) Changes in farmer / entrepreneur behavior and practices	Do this as accurately as you can with cleverly developed indicators. This is the key level to give attention and to spend money on.	Changes in practices, changes in behavior, investments etc. The changes must both be clearly linked to your activities and easily linked to farm or enterprise performance.
4) Enterprise / farmer performance, (Income)	Measure if easy, otherwise: calculate from no 3)	General trends to see if calculated values make sense
5) poverty	Do a descriptive explanation how poverty is affected with some categories and outreach numbers.	General trends to see if story makes sense

- ❖ Do intervention based logics and not market level ones. This may look as more work but is much simpler.
- ❖ There are those logics that have a very limited impact on the overall numbers....Only do simple monitoring. There are also interventions where the attribution will be questioned by many, I suggest that, in those cases, you just measure outreach and not try to measure impact.
- ❖ QA&R is probably
 - 50% hard work, common sense by manager
 - 40% having the right local staff and culture
 - 10% the right external support
- ❖ Hiring QA&R staff is not just hiring someone who “did QA&R before”. In many existing projects or NGOs, what is called QA&R is actually only counting how many trainings have been done. The QA&R person or team needs to have the following skills:
 - Strong analytical skills to design and update impact logics.
 - Ability to analyze large surveys, ability to triangulate
 - Well organized to keep good records
 - Ability to design small surveys and manage large ones, some experience is very useful here.
 - Reasonably good writer for mini-cases etc.
 - Real team player who can work well with other specialists.

If there are 2 positions, one could consider hiring a really clever analytical thinker and someone familiar with conducting surveys. The clever thinker could be someone working at the university with a PHD.
- ❖ At the start of projects there is normally the intention to do QA&R well and once the project or program is in full operation, QA&R is often considered less urgent and is therefore postponed. Especially when the project needs to triangulate and some of the data just does not make sense, people tend to look the other way. As a manager you need to plan QA&R events clearly in advance. Make sure all managers are actively involved and make sure there is enough analytical capacity and critical minds in each team.

These are of course just some personal thoughts and not necessarily those of my employer AusAID. They are based on my work in ILO, Katalyst and CAVAC, but also on many talks with development colleagues. I hope some points are useful. Please edit, copy and paste freely. Constructive feedback is appreciated.

peterroggekamp@cavackh.org