

MANAGEMENT RESPONSE - BTC

Ex post Impact Evaluation of four Bilateral Cooperation projects

BTC welcomes the results of the ex post impact evaluation, which significantly contributes to the search for means to monitor and report development cooperation results, a concern and challenge that we as the implementing instance of bilateral development cooperation obviously share. Overall, this was an interesting process. It has allowed several actors to become aware of the challenges involved in impact evaluation, of its limits and of the importance of a common and shared understanding.

We want to highlight a few aspects of the findings, conclusions, recommendations and lessons learned that, according to us, need to be further discussed before we can determine how and to what degree impact evaluation can or should play a role in Belgium's development cooperation. These aspects are discussed in the last part of this management response, after our specific response to the findings on the 4 interventions and to the three recommendations.

Specific response to the impact evaluation of four interventions

Rwanda

This evaluation process was annulled due to external factors on which the actors had insufficient influence. Even though consequently no findings were formulated about this intervention *as such*, the situation provided for methodologically interesting case material to draw lessons for future impact evaluations. Therefore, this case deserves to be further analysed, since the conclusions can be interesting for determining the role of impact evaluation in Belgium's development cooperation.

DR Congo

Even before the start of this evaluation, it was expected the impact would be negative. Heavy hypotheses, which were not addressed during implementation, were weighing on the achievements at the outcome and impact level. We do agree that this intervention is to be avoided in the future. Consequently, the necessary lessons must be drawn and an adequate answer must be found on how to support the education sector in partner countries where the institutional context contains strong assumptions on the achievement of development results. Within this framework, according to us, the "post-conflict" context in which this intervention is set up deserves more attention. Also, extrapolation of results from 1 site to all sites must be interpreted with due care, since, as the evaluators actually point out, in particular the personality of the actors – *their leadership* – can make the difference.

Morocco

The original report of the evaluators as well as the synthesis report point at the poor design of the intervention but at the same time conclude that overall the outputs have been correctly delivered, that the outcome has been achieved and that a positive impact can be seen on the income of the beneficiary farmers. A generally positive impact is noticed. However, we did not find this in the summary. It emphasises the negative elements of the outcome results, while the positive results are toned down or omitted; also reference is made to the lack of effect on the rural exodus, where the question arises whether it is realistic to expect effects from this type of intervention on this macro objective, which is subject to many other factors. Even though we understand that choices have to be made when drawing up a summary, we find that the original mind-set need to be maintained so the summary still corresponds with the original final report of the evaluators and with the synthesis report and that reporting bias is excluded.

Senegal

In the reporting about the intervention in Senegal an even bigger transformation is taking place: After the consultants submitted their report a gradual shift has taken place from a positive final evaluation to a rather negative to negative final evaluation.

For instance, the conclusion of the report from the consultants submitted to the steering committee is as follows (p. 136-139):

- *Pertaining to sustainable access to water:*
 - **The main achievements of PARPEBA are still in place and continue to have a positive impact on the beneficiaries.**
 - **The good design of the infrastructure and appropriate quality of the works is noteworthy. It is still functioning and used in 2012.**
 - **PARPEBA helped to provide improved access for the population of the connected villages but also, to a lesser extent, to the population of neighbouring unconnected villages.**
 - **PARPEBA helped to provide improved sustainable access for the population of a connected village, and to a lesser extent, to the population of a village not connected to water from a borehole.**
- *Pertaining to water usage and its effects:*
 - **It is not possible to assess, based on the data collected, whether the installation of the borehole allowed households to consume more water than before. However, the field findings seem to indicate that PARPEBA has not caused an increase in water consumption per day and per person.** » *The following sub-points have also been highlighted:*
 - **Users are strongly inclined to pay.**
 - **An analysis of double-difference averages allows concluding that the access to borehole water significantly reduces water collection time to different degrees depending on the type of access and the season.**
 - **Access to the standpipe significantly reduces the difficulty of collecting water by eliminating the lifting chore.**
 - **These findings insinuate that households that can access water from a borehole can better smoothen their income throughout the year.**
 - **Effect on education. An original but imperfect test seems to indicate there is a positive effect from accessing water from a borehole on education.**
 - **Effect on the well-being and social cohesion. Access to tap water in all likelihood helped improve the well-being of the beneficiary populations, women, men and children.**
 - **Several elements show that PARPEBA is genuinely concerned with measuring its effects. However, the currently available information cannot be used towards measuring its impact. This is disconcerting. There are several causes for this.**
- *Pertaining to the quality of water and its effects:*
 - **PARPEBA gave priority to quantity over quality.** PARPEBA was embedded in the national policy that gave priority to the quantity of (public) connections (in rural zones) over the quality of the water supply. **Potability of water remains problematic from a chemical point of view.** One third of the sites concerned by the two last projects had acceptable chemical agent levels (according to WHO standards).
 - **The health data collected do not allow the assessment of a causal relation between the consumption of water from a borehole and signs of improved health.** On the other hand, the qualitative data collected indicate improved well-being of beneficiaries, less fatigue, less sickness (also owing to other factors), more cleanliness, etc.
- *Pertaining to general living conditions:* **PARPEBA contributes significantly and sustainably to improved living conditions of an important part of the population of the three regions concerned (global impact).**
- *Pertaining to the cross-cutting themes:* **The cross-cutting themes of gender and the environment have been widely addressed by PARPEBA. PARPEBA has significant positive effects on women.**

In the synthesis report only the following elements have been retained:

- The effects of water usage.
 - Overall the project had a **positive effect on the beneficiaries and to a lesser extent the non-beneficiaries.**
 - The study shows that the use of tap water has resulted in **significant time and energy savings, especially for those having access to a private tap** (self-financed).
 - Another effect of the new boreholes highlighted by the microeconomic survey is **an increase in well-being and improved social cohesion** (a reduction in conflicts between women, who are primarily responsible for collecting water). It would be worth studying the effect on education in greater detail, particularly a reduction in the gender gap in school attendance. One of the great merits of the household surveys conducted in Senegal is that they highlight a certain number of unexpected indirect effects of enterprise cooperation.
 - **PARPEBA helped to offer improved water access to at least 15.5% of the population in the affected region.** Unfortunately, this percentage is divided by 3, if only boreholes offering water with a good chemical composition are taken into consideration.
- Water quality and its effects.
 - Although the positive effects of access to tap water were demonstrated, **water quality in terms of both bacterial and chemical contamination continues to be a real problem.** Indeed, over two thirds of the sites that were constructed or rehabilitated through PARPEBA have harmful fluorine and salt levels. In addition, there are many water contamination sources between the point of supply and point of use. The population is generally unaware of these water quality problems. Such an observation calls for urgent corrective measures to prevent the risk of serious health problems among the rural population in the near future.

Of these elements only the following are retained in the summary of the synthesis report:

- The effects of water usage.
 - The evaluation concluded that **the water supply from boreholes through taps has saved significant amounts of time and energy, especially for beneficiaries with access to a private tap.**
 - Another effect of the new boreholes was **the improvement of well-being and the strengthening of social cohesion** through a reduction in conflicts between women, who are primarily responsible for collecting water.
- Water quality and its effects.
 - **Nevertheless, water quality in terms of both bacteria and chemical contamination remained a real problem. More than two thirds of the sites that were built or rehabilitated through PARPEBA had harmful fluorine and salt levels (WHO standards).** In addition, there are many water contamination sources between the point of supply and point of use. The population was unaware of these problems. Such an observation calls for urgent corrective measures to prevent the risk of serious health problems among the rural population in the near future.

We notice that the number of positive elements retained in the synthesis report has strongly dropped, that these are mostly methodologically questioned, that the negative aspects of water quality are zoomed in on compared to the analyses made and the context, and that allusions are made about the fact that this intervention leads to serious health problems. We therefore conclude reporting bias has occurred in this case.

We do not question the gravity of the consequences of excessive fluorine levels or the fact that most of the sites of this intervention have too high a fluorine level in the water basin. Placing these elements in context as well as mentioning the positive ex post evolutions would provide for a different image and would provide a better footing for formulating alternatives and policy choices. Since an impact evaluation by definition measures the effects of an intervention, the report misses a chance when concluding there is a health hazard, without showing the measurable health effect after years of water usage. Even so, the summary of the synthesis report gives the impression that the population has not been sensitised about the consequences of the poor quality of the water, where the initial wording said “not sufficiently” sensitised. Another “shift” has occurred due to the omission of the fact that this intervention has applied the Paris Declaration principles by aligning the intervention with

Senegal's water strategy – which, in that phase, did put quantity above quality. The synthesis report also does not regard the fact that later water interventions in Senegal targeted on improving the water quality and that in other words this issue has not been neglected by the Belgian Development Cooperation. On the contrary. The partner country was successfully influenced. It has adapted its water policy and given water quality higher priority. This is an example of the ownership principle put to practice. The geo-hydrological situation in the intervention region is complex and insufficiently known. In large parts of this zone there are no productive aquifers of good water quality (62% of all boreholes in the intervention region have too high a fluorine level). To avoid irreparable ecological damage geo-hydrological studies must be conducted before the usage of aquifers with good water quality can increase. Also in this area progress is made: Belgium finances studies to better map the complex water situation in Senegal and promote evidence based decisions in view of a sustainable supply of water of good quality to the population. As short-term solutions the Belgian Development Cooperation also installs (reverse osmosis) water treatment units that supply good quality drinking water for drinking purposes only.

According to us, all these elements give a different insight in the results of this evaluation as presented in the synthesis report. It remains unclear whether Belgium should have made other choices than working progressively from investing in quantity to influencing the partner in view of implementing a sustainable water policy.

These shifts between the various reports go further than detail difference. It changes the final conclusion a reader may draw. The shifts do not contribute either to drawing lessons that would be relevant for policy-making in the water sector.

Specific response to the recommendations

1. Define a global evaluation strategy where impact evaluation has a place

We endorse the principle of these recommendations, which is to bring about a reasoned system for the Belgian Development Cooperation's accountability, steering and learning functions. However, we find this still requires further research, in order to have the objective of impact evaluations and the methodologies used for the purpose assessed in a realistic way. As pointed out below, there are unanswered questions that require a strategic policy answer. It would be rash to implement impact evaluation measures that are based on interventions that do not suit the current development cooperation paradigm and that only make up a fraction of today's portfolio of bilateral development cooperation interventions.

2. Conduct high-quality rigorous impact evaluations

We endorse the principle of conducting high-quality rigorous outcome and impact evaluations as well as the importance of a mixed methodology and having a good results framework along with a good intervention strategy. However, we find it too early to already put this down in detailed steps, because: i) The nature of the results as well as the strategy of current interventions differs, because of their programme-broad support, very much from the interventions evaluated; ii) Interventions do not have the necessary means available to be able to measure the impact during implementation, and iii) The use of partner country M&E systems – or supporting these where they are not sufficiently developed – requires the Belgian Development Cooperation to make strategic choices.

3. “Management towards outcomes and impact” for more realistic projects targeting the achievement of tangible effects

We endorse the importance and the possibility of monitoring targeted outcome results. MoRe Results, BTC's M&E system, is currently targeting the outcome level and also lists the assumptions that determine the change process between the output and outcome level, which is essential for drawing

lessons from interventions. This approach suits the requirements inherent to impact evaluation at the level of outcome measurement.

We endorse the importance of M&E at the impact level. However, this requires policy choices that have not been made yet (see the general response to impact evaluation below) as well as further research into the most adequate and efficient methods to measure this level, taking into account the possibilities of the partner countries and the high costs of parallel data collection systems at this level.

However, we want to expound on conclusions that are at the basis of this recommendation. The design and consequently the M&E framework of the interventions that were conducted between 2004 and 2008 were defined around 2002. This was a long time before the Paris agreements of 2005, which would bring change in the matter. Before Paris, achieving objectives, not results, was paramount. Paris brought about significant change. Monitoring results was limited to the output results; outcome indicators were mainly used for "explaining" the objective targeted, not to redirect it. This focus on the operational level of interventions has fortunately changed since the Paris Declaration, which has also been transposed in BTC's M&E system. The recommendations, even though they are pertinent in a generic sense, do not take into account the current situation. This is not insignificant if we want to make the recommendations operational.

Concerning the internal evaluation system, the synthesis report states "The internal evaluation system is challenged. Indeed, almost all the conclusions issued by the final project evaluations have been contradicted or slightly modified by the findings of the ex post evaluations."ⁱⁱ This "lesson learned" is not included in the report that was submitted to the steering committee. We were ourselves only notified about this opinion via the publication of the synthesis report. Moreover, this argument is flawed. No final evaluation was conducted for DRC and Moroccoⁱⁱⁱ. We therefore regret that opinions can be included in a synthesis report without prior consultation of the members of the steering committee, especially since the opinions are incorrect.

Pertaining to Senegal, the conclusions of the final evaluation on the impact of the intervention generally do converge with the conclusions in the initial report drawn up by the evaluators of this impact evaluation. The difference is in the fact that there is no measurable evidence in the final evaluation for this. This is interesting for drawing lessons on the "value for money" of impact evaluations, since the final evaluation can be conducted with a significantly smaller budget, while leading to similar conclusions.

General response to the ex post impact evaluation

Policy dimension of impact evaluation: alignment, ownership and accountability

Since the Paris Declaration, interventions are to support the results of the partner countries and to monitor these results partner country M&E systems have to be used. In concrete terms, this means that the impact level is a result of the partner country, while the outcome level indicates how Belgium supports the partner in achieving this impact level. Determining and measuring the impact level, and to a lesser extent the outcome level, is consequently done with partner country measurement systems. The evaluation does not start from an approach where measuring the impact level is the partner's competence. This means the following issues are still not addressed:

- The evaluation of the intervention in Rwanda has exposed the sensitive issue of data collection at the impact level. Not every partner country is open to the idea of sharing data about these strategic results, which are developed and managed under its ownership. In 1 of the 4 cases this was an issue. This issue shows a policy position is needed.
- The question also arises to what degree Belgium can take over the partner's role in measuring the impact. Is preference given to compliance with the commitments made in Paris

on using partner country systems at the expense of accountability (to Belgian stakeholders), or is internal accountability more important than alignment and ownership?

- An additional consideration is whether the resources should rather be used, in accordance with Article 18 c of the *Busan Partnership for Effective Development Cooperation*, to systematically enhance partner country capacity so their statistics and M&E systems can monitor the impact level?

Impact evaluation as a means to support policy

Impact evaluations are expensive. The evaluation report assumes that these costs disappear when M&E systems of the interventions measure the outcome and impact. For the outcome current M&E systems are designed in such a way that this should be the case. Measuring the impact level, as the evaluation points out, is a complex matter that requires the necessary resources, including competencies. If one wants to measure impact in the current situation, this comes down to setting up parallel measuring systems at the intervention level, with a shift of the expenditure for data collection of the impact evaluation to the interventions. The following question arises in this respect:

- To what degree can resources for impact-level data collection be withdrawn from investments in services and supplies for the partner country? Again, this is a choice that must be made by the Belgian Development Cooperation, i.e. is it feasible to negotiate with the partner countries about higher M&E costs and extra international technical assistance for the interventions to measure impact?

Impact evaluation can be determining for policy support. However, this conclusion comes with the necessary questions and constraints.

- This impact evaluation has evaluated 4 interventions while the total number of interventions of bilateral development cooperation currently amounts to approximately 150 interventions. Considering the cost and the efforts required as well as the actual availability of data at the impact level collected by the partner countries themselves, it would be impossible to do so for all interventions.
- To what degree can the conclusions of these interventions be extrapolated to similar interventions in a different context? It would be rash to assume that the conclusions of an impact evaluation on a water intervention in Senegal apply to the whole water sector, irrespective in which country the intervention is conducted.
- The question how to faster produce information that is pertinent to policy-making and reliable to come to reasoned decisions at critical moments, according to us, requires further comparative research that takes into account the actually limited resources and limited data collection systems in the partner countries. We agree with the finding that attention must be paid to further fine-tune how information can be streamlined between the implementation phase and the programming and identification phase in order to strengthen the negotiation position of the Belgian actors.

The importance of realistic interventions, tangible results and a results-based M&E system

- The interventions evaluated, which were concluded in 2008, are atypical compared to current interventions, which usually follow a programme approach in support of the partner country's sector strategy. The difference is that today there is a need for evaluation methods that can measure the impact of this specific sector support and that can speak out about cooperation with other donors.

- The evaluation studied interventions that were concluded in 2008 and that were developed before the Paris Declaration. The Paris Declaration has brought about major changes pertaining to results-based management, with the outcome and impact level not being considered an “objective to be achieved” any longer, but instead a “result to be achieved”. The outcome and impact level of the interventions evaluated was consequently formulated following the former paradigm (objective) but evaluated following the latter paradigm. This has consequences on the applicability of the conclusions and recommendations to today's context.

General conclusion

We share the finding that impact evaluations can be an answer to the demand for rigorous accountability about the achievement (or not) of development results, for support to decision-making and for learning lessons from past interventions. An impact evaluation can also answer the attribution question that is raised at the outcome-impact level, as well as policy-related questions. Here, a value for money assessment is required, as well as further research into methods to deliver policy-supporting information on time so evidence based policy choices can be made.

We also share the opinion that an impact evaluation requires a robust methodology, which ideally is supported by sound monitoring and evaluation (M&E) systems of interventions. Doing so, we pointed out the fact that the cost of data collection is not disappearing, but shifted to the M&E systems of interventions.

We regret that reporting bias occurred in the synthesis report and the summary of that report with respect to Senegal, and to a lesser degree with respect to Morocco, as well as in the statement about internal evaluation systems. This should be avoided in future in order not to disseminate incorrect information.

Pertaining to the recommendations, we have drawn attention to the policy choices required for extending impact evaluations and to the need to further research certain aspects, within a framework that compares the methodology used with other methodologies as well as with the current M&E system of interventions which monitors the outcome and contribution to the impact.

ⁱ Original text in French, translation by BTC.

ⁱⁱ Synthesis report, page 67.

ⁱⁱⁱ This is a departure from the way the system used to work in the past: Usual practice was that in case the mid-term evaluation was conducted in the second half of the intervention period, the final evaluation was omitted. BTC's current M&E system has made such practice impossible.