EVALUABILITY ASSESSMENT OF INTERVENTIONS OF BELGIAN DEVELOPMENT COOPERATION: PRELIMINARY FINDINGS FROM RWANDA

Nathalie **Holvoet**

Liesbeth **Inberg**

Lisa **Popelier**

Institute of Development Policy and Management, University of Antwerp

Emmanuel **Gasana**

Local M&E expert

July 2015

IOB
Institute of Development Policy and Management
University of Antwerp

This report summarises preliminary findings of the evaluability review of 10 interventions of Belgian Development Cooperation in Rwanda. The report is an intermediate input into the larger study which focuses on 40 interventions in Benin, RDC, Rwanda and Belgium. This report has no official status and will not be discussed separately during meetings of the study's advisory board.

Table of contents

1. Introduction

Within development cooperation, evaluations have become increasingly important within the management cycle, not only at project level, but at higher levels as well. In addition, the role of evaluations has become clearer, with an increasing emphasis on the different objectives of evaluations: accountability for the use of the received resources, learning from experiences of the past in order to improve in the future and supporting future policy-making and –practices. Taken into account this increasing importance of evaluations within development cooperation, it is important to pay attention to the relevance and quality of evaluations and to the follow-up of their findings. A first and important (but often forgotten) step in this process is to check if all conditions to conduct a sound evaluation are fulfilled. In other words, does the evaluation meet the existing needs, is its implementation feasible in the given context and with the available means, and does the use of means match the expected benefits of the evaluation. In reality, erosion of the terminology and of the practice of evaluation has taken place; despite an increase in evaluation experience, for many initiatives called ´evaluation´ it is questionable if minimum evaluation requirements are fulfilled.

In the specific case of Belgian development cooperation, experiences with external evaluations as well as with monitoring suggest that the initial conditions to evaluate are often not or only partly present. As a consequence, it is likely that the ambitions of the evaluation exercises undertaken are not or only partly realised. Over recent years however, many actors involved in Belgian development organisations have increasingly invested in the development of their M&E systems, through which, in principle, the evaluability should have been increased.

Against the background of the above observations, the concept of evaluability has gained more interest in Belgian development cooperation, particularly in the Office of the Special Evaluator, that decided to commission 'an evaluability review'. The objective of the review is to diagnose and analyse the current evaluability of different types of interventions of Belgian development cooperation which can feed into a better understanding of evaluability and provide some suggestions towards improved evaluability in the future. More broadly, the review can as well be helpful in elaborating and improving monitoring and evaluation policies and systems.

The starting point of the review is the study conducted by the Office of the Special Evaluator which focused on the theoretical evaluability of a broad range of interventions of Belgian development cooperation. The review realised by South Research and the Institute of Development Policy and Management, focuses on three dimensions of evaluability, including 'theoretical evaluability', ´practical evaluability´ and 'influence of contextual factors'. Our review draws upon data collection related to a sample of 40 interventions in four countries: Belgium, Benin, Democratic Republic of Congo and Rwanda. This country note specifically presents the aggregated findings for the ten interventions in Rwanda. At this stage we do not disaggregate alongside intervention modality, sector, etc. as this would violate confidentiality. More detailed analysis alongside a number of potentially 'distinguishing' features will be done once data for the four different countries is merged at the end of the review.

The structure of this note is as follows: section two provides an overview of the methodology of the review while section three presents findings, discussion and suggestions.

2. Methodology

2.1. Review Framework

There is thus far no standardized framework for evaluability assessment. Different agencies are using different formats. Our own review framework draws upon various definitions of 'evaluability'. The definition of the OESO/DAC has been used most often: "*Evaluability is the extent to which an activity or project can be evaluated in a reliable and credible fashion".* Other definitions add, however, other important elements by also looking at evaluation practice and by emphasizing the need to check as well whether an evaluation is justified and feasible and able to supply useful information. This implies that the contribution of an evaluation to the improvement of the management and the results of the intervention is also verified in an evaluability assessment.

An important distinction that is often made is between ´theoretical evaluability´ and ´practical evaluability´. Theoretical evaluability refers to evaluability in principle, as derived from the intervention's design without taking the actual practice into consideration. Practical evaluability, on the other hand, verifies, for example, if data on the progress of the actions are collected effectively and if M&E systems are really being used, so that a contribution to the improvement of management and results is possible. In this way the link (practical) utility is easily made.

We have approached evaluability mainly as a continuum; no single intervention is perfectly evaluable in all its aspects, while at the same time it is also highly unlikely that interventions exist that cannot be evaluated at all. In addition, an evaluation can also entail many aspects of the intervention; the classic DAC criteria of relevance, effectiveness, efficiency, impact and sustainability and other criteria, but it is not necessary that all evaluations deal with all these criteria simultaneously. Moreover, evaluability can also be linked to the objective of the evaluation. While evaluations could have different objectives (accountability, learning, support of future policy), it is not necessary to include all of these prominently in all evaluations. If for example learning is a central objective, evaluability requirements are different than if accountability is a central objective. Finally, evaluability also depends on the evaluation policy of an organisation. In the evaluation policy a choice can, for instance, be made to assign a disproportionate amount of the available resources to the evaluation of innovative interventions.

In the elaboration of the review framework, we have been inspired by Rick Davies´ study for DFID[1]. Compared to Davies´ study, the second and most important part of the review framework is more detailed, as a result of discussions and reflections within the review team. The review framework distinguishes three key dimensions:

- Analysis of the intervention design, including the underlying theory of change (theoretical evaluability);
- Analysis of the evaluability in practice, including availability of information regarding the implementation and results of the intervention as well as the M&E system in practice (practical evaluability)
- Analysis of the evaluation context, which is considered to include all elements outside the specific intervention context (not only social, economic, cultural, political, security, but also e.g. the culture of groups and organisations which are involved in the implementation of the intervention, the financial institutions,…).

Taking into account the objectives of the review, we have tried to design a realistic review framework (not too ambitious), which is not too detailed. However, practical evaluability has many dimensions

[1] Davies (2013) *Planning evaluability assessments: A synthesis of literature with recommendations*. Report of a study commissioned by the Department for International Development.

and involved many different issues, which could not be ignored in the review framework. Therefore, a compromise between realism and completeness has been made by applying a hierarchy in the components of the review framework. In this way factors have been included under a number of central criteria (comparable to the approach in evaluations in which main questions are being operationalised through assessment criteria and assessment criteria through indicators). The table below provides an overview of the three key dimensions and 9 main components (2-digit level) of the review framework, with the maximum number of specific items being scored under each component mentioned between brackets. The figures in the columns give an overview of the number of items scored for the different OECD/DAC criteria[2]. More specifically, as regards the evaluability of effectiveness e.g., 4 of the 7 items covered under the component 'underlying analysis' are considered relevant and have been scored for evaluability of effectiveness. A full overview of the framework with the different components (2-digit level) and the items (3-digit level) is added in Annex 1. Whether a specific item is considered relevant for a specific OECD/DAC criteria is clear from the crosses in the related columns and the explanation in the grid.

Table 1: Overview of the review framework

DAC Evaluation criteria Components (number of items)	Relevance	effectiveness	efficiency	impact	sustainability
	(number of items scored)				
1. Analysis of the intervention design					
1.1 Underlying analysis (7)	(7)	(4)	(0)	(0)	(0)
1.2 The intervention logic and the theory of change (ToC) (8)	(1)	(6)	(5)	(5)	(3)
1.3 The proposed M&E system (9)	(5)	(7)	(9)	(7)	(7)
1.4. Consistency and adaptation of the intervention logic and ToC (3)	(3)	(3)	(3)	(3)	(3)
2. Practice regarding intervention implementation, intervention management and context					
2.1 Basic information regarding intervention implementation (11)	(5)	(10)	(11)	(10)	(8)
2.2 The M&E system in practice (12)	(11)	(12)	(12)	(12)	(12)
3. The evaluation- context					
3.1 Attitude of key players (9)	(9)	(9)	(9)	(9)	(9)
3.2 The broader context (3)	(3)	(3)	(3)	(3)	(3)
3.3 Practical elements (2) (°)					
Aggregate score (number of items scored)	(44)	(54)	(52)	(49)	(45)
4. Suggestions to increase the evaluability and the usefulness of evaluations in the future					
5. Feedback from the stakeholders involved on the analysis and the suggestions provided					

(°) this component has been only mentioned pro memory, because these parameters and scores are connected with the organisation of a real evaluation.

[2] We use the OECD/DAC definitions of the different OECD/DAC criteria. See OECD (DAC)(2002). *Glossary of Key Terms in Evaluation and Results Based Management*. Paris: OECD/DAC.

2.2. Sample

The following criteria have been used to select the sample:

- Intervention type: inclusion of different actors in each country sample, including governmental actors (BTC), non-governmental actors (NGOs and syndicate organizations) and other actors including VVOB, APEFE, CIUF and VLIR.
- Principal sectors: inclusion of sectors with a complex and less complex theory of change. Before the field research an initial subdivision of sectors was made: health, education, water and sanitation, social economy were grouped together as 'less complex' while climate and environment, conflict and peace, governance and agriculture and rural development were considered 'more complex'. In practice, however, a number of interventions in a 'complex' sector have a more complex theory of change, for example an intervention that aims the institutional strengthening in the Ministry of Health. After the field mission, we have thus further refined the subdivision over 'complex' and 'less complex' theories of change.
- Implementation period: inclusion of interventions with considerable expenditures in 2014, with the idea that it is too difficult to determine the practical evaluability for those interventions that were closed some years ago.

These criteria led in Rwanda to the selection of the following interventions:

Table 2: overview of selected interventions in Rwanda

Title	Organisation
SPAT II: Market oriented advisory services and quality seeds	CTB/BTC
PAREF II: Projet d'appui à la reforestation dans les provinces de l'Est et du Nord	CTB/BTC
MINISANTE IV : Appui Institutionnel au Ministère de la Santé	CTB/BTC
SACB: Strategic Approach to Capacity Building; Support to the National Capacity Building Secretariat (NCBS)	CTB/BTC
LOPE: Learning Outcomes in Primary Education (2014-2016)	VVOB
Rwandan Mountain Tea	BIO
Program d'Appui Institutionnelle	ARES
PASAB II: Projet d'appui à la sécurité alimentaire au Bugesera	Caritas
Accès à et gestion de l'eau et de l'assainissement au Rwanda	Protos
Rendre la vue au Rwanda: prévention et traitements des problèmes de vues	Lumière pour le Monde

2.3. Data collection

The evaluability review draws upon a combination of secondary and primary data collection. Secondary data was provided by the organisations responsible for the implementation of the interventions and included (if relevant) the intervention proposal, baseline report, progress reports (e.g. mid term, yearly reports and end term reports), studies and evaluations, documents related to the monitoring and evaluation policy of the organisation. The collection of documents has been restricted to the same set of basic documents for each of the interventions in order to limit the time needed for the assessment to four working days/intervention.

Primary data has been collected through focus group discussions in Brussels and through interviews with important key actors on the ground. In Brussels four focus group discussions have been

organised, one with DGD actors (April 20, 2015), one with BTC actors (April 17, 2015) and two with non-governmental actors (subdivided over Flemish and French speaking actors) (April 20, 2015). The response to the invitation for the focus group discussion was quite varied; a high response from the side of non-governmental actors and BTC actors and a low response from DGD actors.

In Rwanda, 81 important key actors have been interviewed (over the 10 interventions), using a total of approximately 2 days per intervention (starting from April, 22nd to May, 13th, see Annex 2). Some of the key actors have been interviewed together, either because they were actors with comparable M&E responsibilities or positions within the intervention, or because of time constraints. After having interviewed all stakeholders involved in an intervention, the members of the review team organised a confidential feedback session in which the first preliminary findings (strengths, weaknesses, suggestions) were discussed and verified with an intervention's key actors. This allowed to fine-tune findings and to add additional data where needed. At the end of the field mission (May, 13th) a debriefing session was organised at the Belgian embassy for all stakeholders involved in the different interventions. During the debriefing session preliminary aggregated findings were presented and discussed (see Annex 3 for the power point presentation).

2.4. Analysis procedure

The different components (two-digit) and the 62 items of the review framework (3-digit questions) have been scored with regard to (some of) the DAC criteria using scores that vary between 1 and 5. In order to reduce subjectivity, scoring has been done by (at least) two persons on the basis of a grid which explains for each component, item and specific DAC criteria the different scores. In case of differences between the scores assigned by the two assessors, the scores are calibrated in order to arrive at a final score. Scoring of components (2-digit) has been done in two ways: firstly a score has been assigned on the basis of the evaluation grid, disregarding the specific scores on the different items that are related to the component; secondly an unweighted average has been calculated on the basis of scores being assigned to the items related to a specific component (e.g. scoring of 2.1. is based upon scores for 2.1.1. to 2.1.11). Some caution is warranted in interpreting scores: these cannot be interpreted as absolute numbers and are only used for comparative purposes. It allows to get a quick overview of those components, items, and OECD DAC criteria that score relatively good or poor in terms of evaluability.

Based on the scoring for the 10 different interventions, indexes have been calculated for each of the items and components, differentiated alongside the different OECD/DAC criteria (where applicable). For example, item 1.1.1. ('the beneficiaries/target groups are clearly identified and described') was considered important for the evaluability of relevance and effectiveness (see Annex 1) and thus interventions were scored on item 1.1.1. for these two DAC criteria. Subsequently, the ' relevance evaluability' index for item 1.1.1 was calculated as follows: [(number of interventions with score 1 x score 1) + (number of interventions with score 2 x 2) + (number of interventions with score 3 x 3) + (number of interventions score 4 x 4) + (number of interventions with score 5 for relevance x 5)]/maximum score (=number of interventions scored x 5). The printscreen in Annex 4 highlights that of the 10 interventions 0 have obtained a score 1, 2 have obtained a score 2, 2 score 3, 4 score 4 and 2 score 5, which finally results in an index of 0.72. Similarly, indexes were calculated for all different items, each time differentiated alongside the relevant OECD/DAC criteria. While the 10 interventions could often be scored on all items, some 'non-applicable' scores were assigned. This was amongst others the case for component 1.4 related to 'consistency and adaptation of the intervention logic and ToC' which was only scored in case changes effectively took place.

In line with the two different ways of component scoring mentioned above (i.e. unweighted average of items and direct scoring on the basis of the grid at 2-digit level) two different indexes have been calculated for each of 2-digit components.

Based on the scores, we have also ranked for each of the different OECD/DAC criteria the different items. More specifically, the 44 items that have been scored on 'relevance' have been ranked from 1 (highest score) to 44 (lowest score). A similar procedure has been followed for effectiveness (ranking from 1 to 53), efficiency (ranking from 1 to 51), impact (ranking from 1 to 49) and sustainability (ranking from 1 to 44). We have as well looked at item ranking within each of the three dimensions (theoretical evaluability, system in practice and context) while we also ranked the 8 components (2-digit level), differentiated for each of the different OECD/DAC criteria in order to distil those components on which the 10 interventions generally score better (or poor) in terms of evaluability.

3. Findings, discussion and suggestions

In what follows we first give a general overview of the findings at the level of the 8 components, differentiated alongside the 5 different OECD/DAC criteria. Subsequently, a more detailed analysis is provided subdivided over the different components. As recommendations are linked to each of the different components and items, we also immediately integrate the suggestions in this section.

In order to get a quick overview of the findings, we visualise our findings by regrouping index-scores into 4 different categories and assigned to each category a colour. This also allows comparability throughout the analysis.

Index 0.81-1
Index 0.61-0.8
Index 0.41-0.6
Index 0.2-0.4

3.1. General Overview

Table 3 which gives a general overview of the findings highlights that the **quality of the underlying analysis (component 1.1) is the component that has overall obtained the highest score** which is indicative of the fact that for most interventions the underlying analysis is clearly described and comprehensive. As the quality of the analysis is particularly important to evaluate the relevance and the effectiveness of an intervention, it particularly increases the possibility to evaluate in a sound way the relevance and effectiveness of interventions. Reversely, the component that has obtained the **poorest results is the 'consistency and adaptation of the intervention logic and ToC'**. In those cases where interventions have adapted the intervention logic in the course of the intervention, this has often not been reported or adjusted in intervention documents. This obviously hampers subsequent evaluative activities. It particularly affects the likelihood of arriving at a sound impact evaluation as it becomes difficult to disentangle implementation failures from underlying flaws in the ToC or intervention logic. More fundamentally, a gap between the intervention (and the M&E system) as it exists on paper and the reality on the ground puts into perspective the usefulness of an M&E system and makes it a ritual exercise. This is not unlikely to happen in the Rwandan **context** given the rapid changes in national policies and their implementation which obviously also affect interventions. Moreover, interventions/organisations are not incentivized to report changes given the limited flexibility in funding agencies' reporting formats and the disincentive to report changes and particularly failures in the context of the Rwandan performance culture and drive towards results.

Our findings also generally hint at the fact that it is **particularly impact evaluation that is likely to be most difficult** for our interventions under study. As further discussed below, this is related to a number of specific elements: different interventions do not elaborate a ToC up to the impact level, do not specify critical ingredients, assumptions or internal risks at the level of outcomes/impact, the majority of the interventions do not collect data on a counterfactual, etc. On the other hand, most interventions have those **elements in place to proceed with efficiency evaluation.** In many cases, resources can be allocated to specific outputs and outcomes. While there are substantial differences among our interventions under study (particularly related to the intervention channel), **our findings also point out that contextual factors, including the attitude of key players and the broader context have a number of positive effects on evaluability whereas there are also a number of possible downsides.** More specifically, many actors have an interest and demand for evaluation which drives the supply of M&E and also influences the sustainability of the M&E (system) itself. Reversely, the performance culture and the strong focus on 'accountability within the system', and particularly the drive to arrive quickly at 'positive' results, may as well undermine the quality and reliability of data collection as individuals and organisations may be reluctant to disclose 'negative' results. It may as

well negatively affect individual's and organisation's scope for experimentation and put all the emphasis and efforts on those issues and items that are in 'performance contracts' and that are being monitored at the detriment of those issues that are outside such contracts. Besides the Rwandan evaluation culture, also the fear for penalisation (loss of funds) by the funding agency tends to curtail the scope for experimentation.These downsides may obviously jeopardize evaluability and particularly the evaluation's learning function.

Table 3: Overview of evaluability scores for 10 interventions (subdivided over components and criteria)

OECD/DAC criteria **Components**	**Relevance**	**Effectiveness**	**Efficiency**	**Impact**	**Sustainability**
Dimension 1: THEORETICAL EVALUABILITY					
1. 1. Analysis of the intervention design					
1.2. The intervention logic and the ToC					
1.4.Consistency and adaptation of the intervention logic and ToC					
1.3. Proposed M&E system					
Dimension 2: PRACTICAL EVALUABILITY					
2.1. Basic information regarding intervention implementation					
2.2. M&E system in practice					
Dimension 3: CONTEXTUAL FACTORS					
3.1. Attitude of key players					
3.2. Broader context					

3.2. Theoretical Evaluability

3.2.1. Analysis of intervention design

A more detailed overview of the findings related to the component 'analysis of the intervention design' is provided in Table 4.

Our findings demonstrate that for the 10 interventions under study, the linkage between the analysis and the intervention objectives is particularly well described. A similar conclusion holds for the quality of the description of the intervention's rationale and situation of the beneficiaries as well as for the linkage between the intervention and a country's sector policy. In most cases, the analysis draws upon a variety of sources including baseline studies (that are often of in-depth and of high quality), previous mid-term review and evaluation reports, workshops, etc. The quality of the analysis particularly affects the likelihood of having sound evaluations of relevance and effectiveness.

Reversely, a striking weakness is the absence of a gender analysis[3] in 7 out of the 10 interventions. This does not entirely come as a surprise and is in line with evidence from previous evaluations[4]. The absence of a gender analysis can severely undermine an intervention's evaluation because it risks covering the reasons for the failure of an intervention. More specifically, as gender relations strongly influence the behaviour of individuals, they often tend to affect the implementation of interventions as well as their results. Disregarding the influence of gender (which thus acts in many cases as an independent variable affecting implementation and results) will thus in many instances lead to flaws in the evaluation quality as well as in the related suggestions for improvement. The observations regarding the absence of gender analysis also hold for other types of disaggregated analysis which were as well largely absent.

Table 4 : Quality of the analysis of the intervention: overview of strengths and weaknesses

OECD/DAC criteria **Item**	**Relevance**	**Effectiveness**
Identification/demarcation and description of the beneficiaries	• **In 4 out of 10 interventions beneficiaries are clearly identified and described, in 2 interventions this description is particularly in-depth. In the remaining 4 interventions the beneficiaries are not clearly identified and/or the description is vague.**	
Description of the rationale of the intervention and the (problem) situation of the beneficiaries	• 8 out of 10 interventions have clearly and comprehensively described the rationale of the intervention as well as the problem situation of the beneficiaries, 2 out of the 8 have done this on the basis of a context analysis, a baseline study or another source of evidence. The remaining 2 interventions have clearly described the rationale of the intervention, but their description of the problem situation of the beneficiaries was vague.	
Description of the role of the most important beneficiaries	• In 4 out of the 10 interventions the role of the beneficiaries is clearly described, in 5 of the interventions this is done in vague way, while in 1 of the interventions information on the role of the beneficiaries is lacking.	
Description of the role of the most important stakeholders (exclusive beneficiaries)	• In 2 out of the 10 interventions there are no other important stakeholders except for the beneficiaries. In 3 out of the 8 remaining projects, the most important stakeholders as well as their role is clearly described while in the 5 other projects a clear description of the role is lacking.	
Integration of gender analysis in the analysis	• 5 out of the 10 interventions do not describe the specific situation of male and female beneficiaries, 2 interventions describe the specific situation but do not include a gender analysis while 3 interventions include a gender analysis.	
Description of the link between the analysis and the intervention objectives	• 8 out of 10 interventions have clearly and systematically described the link between the analysis and intervention objectives, 2 interventions only provide a vague	

[3] Gender analysis refers to analysis of the existing gender relations in a society and the mutually influencing relationship among these gender relations and the development interventions. It mostly involves an analysis of the time and task allocation, of the access to and ownership of material and non-material resources & benefits and of the contextual (legal, social and cultural) factors that influence men's and women's interests, needs, constraints and opportunities. See March et al. March, C., Smyth, I. and Mukhopadhyay, M. (1999) *A Guide to gender-analysis* frameworks. London: Oxfam for an overview of different gender analysis frameworks.

[4] See e.g. Caubergs, L., Charlier, S., Holvoet, N., Inberg, L. and Van Esbroeck, D. (2014) *A difficult path towards equity: gender and development in Belgian Development Cooperation: summary report*. Brussel: Special Evaluation Office of International Development Cooperation, 104 p.

	description.	
Description of the linkage between the rationale of the intervention and the sector policy of the partner country	• 7 out of 10 interventions have clearly and systematically described the link between the rationale of the intervention and the sector policy of the partner country, while for the remaining 3 this description was rather vague.	

3.2.2. The intervention logic and theory of change and possible adaptations (components 1.2 and 1.4)

Findings in Table 5 highlight that most interventions make a clear and correct distinction between outputs, outcomes and impact. Most interventions have also elaborated a ToC which is logic and realistic up to the level of outcomes. However, none of the interventions has a clearly elaborated ToC up the level of impact: while more than half of the interventions do not develop a theory of change up to the level of impact, in those cases where it is developed it is not fully clear or logic. While various interviewees highlighted that they do not see the usefulness of developing a theory of change up to the level of impact (and it is also not always required by funding agencies), elaborating a theory of change up to the impact level is particularly useful from the perspective of designing an intervention and choosing among different implementation modalities.

The lack of a theory of change up to the impact level often goes along with the absence of critical and crucial links up to level of impact while also external assumptions are in most of the cases lacking. This obviously undermines the possibility and quality of impact evaluation as it is particularly the analysis of those critical links, crucial ingredients as well as external assumptions that allows to differentiate between problems at the level of the implementation, flaws in the underlying theory of change and influence of other external factors that are beyond the intervention's control. Moreover, it is specifically the identification and testing of the links and crucial ingredients that allows to provide insight into why an intervention is generating impact or not. For project staff and policy makers these insights into the crucial ingredients of an intervention are useful from a learning perspective (as those ingredients are important to replicate in future interventions). Identifying (and testing) external assumptions can also substantially improve an evaluation's usefulness from an accountability perspective as it facilitates capturing as well the influence of factors that are beyond the control of an intervention.

The specification and testing of crucial ingredients and links between the different building blocks of an intervention is one of the elements that is generally not well covered in the different interventions and it decreases alongside the different steps in the ToC: while there are still 5 interventions that specify links up to the level of outputs, only 3 do this at the level of outcomes, while only 1 at the level of impact. As 'evaluation' (as compared to monitoring, see also below) is particularly about analysing the linkages between the different building blocks of the ToC, it does not come as a surprise that this deficiency affects the overall evaluability of the interventions.

While the way in which the interventions have elaborated their ToC and specified external assumptions negatively affects impact evaluability, findings at the level of efficiency and sustainability evaluability are more positive. All interventions include information regarding foreseen resources and in 4 cases the allocation of all resources to outputs is fully clear, while 1 intervention also integrates resources of other stakeholders in its allocation. As regards sustainability, all interventions include measures which guarantee the sustainability of intervention results: 5 of the interventions do this at the level of inputs or outputs; the other half even does this at the level of both the inputs and outputs. While the latter finding somehow comes as a surprise, it may be indicative of an increasing awareness of the importance of specifically identifying measures to stimulate sustainability. Additionally, in the specific Rwandan case, the high involvement and

ownership of local and national authorities also contributes towards sustainability (see also discussion on contextual factors below).

Adaptations in the intervention logic and theory of change are often not elaborated upon (sometimes not even reported) in intervention documents which negatively affects evaluability (irrespective of the specific OECD/DAC criteria). Where adaptations have taken place, they are also often not integrated in the M&E system. This gap between project documents and M&E systems on the one hand and the reality on the ground on the other hand does not only negatively affect evaluation quality, it also puts into perspective the usefulness of an M&E system on the basis of which erroneous conclusions can be drawn regarding the quality of interventions (related to all OECD/DAC criteria) and the factors which affect implementation and impact of interventions. The latter is not unlikely to happen in the Rwandan context of rapid changes in national policies and quick and effective implementation of such changes (see also below under contextual factors), which often have as well implications on interventions. This makes it all the more important (also for the donor administration) to leave enough room for flexibility and to stimulate interventions/organisations to document adaptations in interventions and adjust where necessary in the M&E systems being used.

Table 5: Quality of intervention logic, theory of change and related adaptations: Overview of strengths and weaknesses (subdivided over different criteria)

	Relevance	Effectiveness	Efficiency	Impact	Sustainability
Clear and correct distinction between outputs, outcomes and impact		7 interventions make a clear and correct distinction between outputs and outcomes, 2 interventions make a clear distinction which is however not correct while in 1 case the distinction between outputs and outcomes is not clear.		6 interventions make a clear and correct distinction between outcomes and impact, 1 intervention makes a clear but incorrect distinction, 3 interventions make no clear nor correct distinction.	
ToC clearly elaborated from outputs to outcomes and impact		2 interventions have a ToC up the level of outcomes which is fully clear, 7 interventions have a ToC up to level of outcomes but it is not fully clear.		None of the interventions has a clearly elaborated ToC up to the level of impact. For 6 interventions the ToC from outputs to impact is not elaborated at all, for 4 interventions it is only partially clear.	
Toc is logic and realistic		5 interventions have a ToC (up to the level of outcomes) which is fully logic and realistic, 4 interventions have a logic ToC which is however not fully realistic while the ToC of one intervention is not logic nor realistic.	6 interventions have a ToC (up to level of outputs) which is fully logic and realistic, 4 interventions have a logic ToC (up to the level of outputs) but it is not fully realistic.	For 2 interventions the ToC is logic and realistic (from outcomes to impact), for 5 interventions it is logic but not realistic, for 3 interventions it is not logic nor realistic.	
Critical and crucial links in the ToC are identified and tested		For 6 interventions critical and crucial links (up to the level of outcomes) are not identified and cannot be tested, for 3 out of the 10 interventions links are identified but difficult to test, for 1 intervention links are identified and can be tested.	For 5 interventions critical and crucial links (up to the level of outputs) are not identified and cannot be tested, for 3 interventions links are identified (up to level of outputs) but difficult to test, for 2 interventions links are identified (up to level of outputs) and they can be tested.	9 of the interventions do not identify critical and crucial links up to the level of impact, 1 intervention identifies critical and crucial links but it is difficult to test them.	
ToC includes concrete measures which guarantee sustainability of the intervention results					5 interventions include measures to guarantee sustainability at the level of inputs *or* outputs, while another 5 interventions include measures at the level of inputs *and*

					outputs.
Internal risks are clearly identified and explored/estimated		For 7 interventions internal risks (up to the level of outcomes) are clearly mentioned and further explored, for 2 interventions risks are mentioned but not explored, for 1 intervention internal risks are not explored.	For 7 interventions internal risks (up to the level of outputs) are clearly mentioned and further explored, for 3 interventions risks are mentioned but not further explored.		For 7 interventions internal risks (up to the level of outcomes) are clearly mentioned and further explored, for 2 interventions risks are mentioned but not explored, for 1 intervention internal risks are not explored.
External assumptions are clearly identified and explored/estimated	For 2 interventions external assumptions (up to the level of outcomes) are clearly mentioned and further explored, for 5 interventions external assumptions are mentioned but not further explored, for 3 interventions external assumptions are not mentioned.		For 2 interventions external assumptions (up to the level of outputs) are clearly mentioned and further explored, for 8 interventions external assumptions are mentioned but not further explored.	8 interventions do not clearly mention external assumptions (at the level of outcomes and impact), 1 intervention identifies external assumptions at this level but does not explore them, while 1 intervention identifies and explores the external assumptions.	For 2 interventions external assumptions (related to sustainability) are clearly mentioned and further explored, for 7 interventions external assumptions are mentioned but not further explored, for 1 intervention external assumptions are not mentioned.
Allocation of resources foreseen to outputs is clear			All interventions include information regarding the foreseen resources. For 1 intervention the allocation of resources to outputs is not clear, for 5 interventions the allocation of part of the resources is clear, for 4 interventions the allocation of all resources to outputs is clear. Out of the latter 4, one intervention also allocates inputs of other stakeholders to outputs.		
Changes in the intervention logic and the ToC are clearly reported	In 3 interventions changes have taken place in the intervention logic that potentially affect effectiveness and relevance. For 1 of these interventions, changes have not been reported, for one intervention changes are reported but not justified and for one intervention changes are reported and justified.		In 6 interventions changes have taken place that potentially affect the efficiency of the intervention. For one of these interventions, changes have not been reported at all, for 3 of the interventions changes have been reported but not justified, in 2 cases	In 2 interventions changes have taken place that potentially affect impact. In one of the cases changes have been reported but not justified, in one of the case changes have been reported and justified.	In 4 interventions changes have taken place that potentially affect sustainability. In 1 of the cases these changes have not been reported, in 2 cases changes have been

		changes have been reported and justified.		reported but not justified, in 1 case changes have been reported and justified.
Information is available about the vision and the opinions regarding the changes	For none of the interventions where changes took place, information is provided regarding the vision and opinions of most important stakeholders regarding these changes.			
Changes in the intervention logic and ToC are adequately integrated in the M&E system	In 2 out of the 3 cases where changes have taken place, these have not been integrated in the M&E system, in 1 of the cases the change has been integrated in the M&E system but not in an adequate way.	2 of the 6 interventions where changes took place that affect efficiency have not integrated these changes in the M&E system, 2 interventions integrated the change in the M&E system up to the level of outputs but not in an adequate way, 2 interventions have integrated changes in the M&E system in an adequate way.	In 1 out of the 2 cases where changes have taken place, these have not been integrated in the M&E system, in 1 of the cases the change has been integrated in the M&E system but not in an adequate way.	In 2 out of the 4 interventions were changes took place that affect the sustainability, changes have not been integrated in the M&E system. In one of the cases changes have been integrated in the M&E system but not in an adequate way, in 1 intervention the change has also been integrated in the M&E system in an adequate way.

3.2.3. Proposed M&E system (component 1.3)

Table 6 highlights that our findings regarding the proposed M&E system are generally in line with findings at the level of the two previous components. More specifically, the lower (e.g. gender disaggregation, specification of external assumptions) scoring-items in the analysis and ToC components are also those that score lower at the level of the proposed M&E system. The lower evaluability of impact is also reconfirmed.

The analysis of the proposed M&E systems of the 10 interventions under study shows that most of the interventions have translated the intervention logic and ToC up to level of outcomes into an M&E system and in more than half of the cases this was also done in a consistent way. Moreover, 6 out of the 10 interventions have consistently translated the sustainability related results into an M&E system. In line with this, the most important results at the level of outputs and outcomes have also been made operational. In most of our 10 interventions, more than half of the indicators are smart, particularly at the output level. Also the sustainability-related elements of interventions have been made operational and in most of the cases sustainability indicators are smart. In line with the above findings, the impact level is less adequately dealt with: 4 interventions do not capture this level in the M&E system. Similarly, the lack of gender analysis mentioned above, also translates into a lack of disaggregation of indicators alongside gender or other relevant parameters. Disaggregation according to gender and other relevant parameters is particularly important from a learning perspective as it may highlight in which circumstances, for which part of the population, an intervention might work or not. It is obviously the selection of parameters alongside which disaggregation is done that is crucial. One *suggestion* is to use the baseline study to identify those parameters that are likely to influence the implementation and results of the intervention. These parameters are often related to basic characteristics of the target population which position them differently in society (such as gender, age, income, disability, rural/urban, ethnicity, …) and also affect their participation in the intervention as well as the probability of achieving the intended impact. Failing to capture such differences in the target population in M&E exercises risks to lead to erroneous conclusions regarding implementation and results of interventions and creates a serious learning deficit.

For most interventions it is possible to allocate expenditures to specific outputs and activities, but in only four cases the reporting mechanisms is embedded in a real MIS. Moving towards a real Management Information System would allow to bring data from different sources together which would foster the feasibility of subsequent comprehensive analyses. Most interventions also describe their method(s) for M&E. Only 1 intervention fails to provide any information on the proposed method, while (again) methods for M&E at the level of impact are not described in half of the interventions. Focusing more closely at the proposed M&E methods highlights that the 'monitoring' component is in most of the interventions much better covered than the 'evaluation' component. This focus on 'monitoring' at the 'detriment' of evaluation is also confirmed by findings related to other components of the review and it does not really come as a surprise. First, monitoring is often considered a necessary first step towards evaluation. Second, evaluation is a much more demanding undertaking than monitoring. While monitoring focuses on the follow-up of the different 'blocks' in the ToC, evaluation deals with analysing the linkages between the different blocks which requires more methodological expertise. It is also from the latter 'evaluative' perspective that the lack of proposed methods for M&E of the external assumptions and internal risks is particularly 'worrisome'. More specifically, none of the interventions proposes methods for the follow-up of external assumptions and internal risks which makes it highly unlikely that there will be M&E of assumptions and internal risks on the ground, which eventually undermines sound evaluation. Third, monitoring is in many instances also politically less challenging than evaluation which studies the underlying reasons for failing implementation and results. Fourth, individuals might also be incentivized towards data collection and monitoring activities (as compared to analysis of data) as the latter often involve

field visits which usually generate per diems. Fifth, as evaluation is also often interesting beyond the boundaries of a specific intervention (as compared to monitoring that is more confined to specific interventions), evaluation has more characteristics of a public good which might as well explain its underinvestment (this may partly be solved by a move towards experiential learning, see also below). Sixth, the tendency for monitoring to crowd out evaluation is also in line with a general trend within development cooperation towards results-based management and more specifically the conflation of results-based management with M&E. In many cases 'management for results' tends to be misinterpreted as 'management by results'. This generally leads to a focus on quick wins and easily measurable results and it might lead to a penalisation of interventions/organizations that focus on complex and long-term changes such as gender equality, and/or a sidelining of interventions/organizations that specialise in policy advisory work and lobbying and where the ToC are more complex. Somehow related to this is the excessive reporting burden (particularly in those cases where interventions/organisations are financed by multiple donors who often use different reporting formats) which is often put upon interventions/organisations which risks to take resources away from evaluation while it is particularly analysis and evaluation that allows to feed into learning.

While most interventions describe the human and financial resources involved in M&E, it is mostly in a fragmented way (information on human *or* financial resources), only 2 interventions clearly describe the human and financial resources involved. In a number of interventions the local stakeholders and beneficiaries play an important role in M&E activities, even to that extent that it can be considered as additional activities with specific results at the level of the project (and the evaluability) but also at the level of the stakeholders and beneficiaries themselves (i.e. the empowering effect of monitoring and particularly 'evaluative' activities). It is recommended for these projects to also explicitly mention these additional resources (which also influences efficiency evaluations).

Finally, most interventionmost (7) of the interventions do not describe if and how the M&E system is aligned to the country's local or national system, nor is a justification provided for non-alignment. As will be discussed under 3.3.2, this does not imply there is no alignment to the system on the ground.

Table 6 : Proposed M&E system: Overview of strengths and weaknesses

OECD/DAC criteria Items	Relevance	Effectiveness	Efficiency	Impact	Sustainability
Most important results are made operational in adequate way		1 intervention has not made results operational up to the level of outcomes operational while the other 9 made results operational. For 4 out of 9 less than half of indicators are smart, for 2 more than half and for three all are smart.	1 intervention has only partly made results operational up to the level of outputs, while the other 9 have made all results operational. For 2 out of 9 less than half of indicators are smart, for 4 more than half and for 3 all are smart.	7 interventions have not made results operational at the level of impact, 1 intervention has made some of the results operational, 1 has made all results operational but indicators are not smart, for 2 interventions indicators are smart.	1 intervention has not made results related to sustainability operational, while the other 9 have made all results operational. For 2 out of 9 less than half of indicators are smart, for 4 more than half of the indicators, for 3 all indicators are smart.
Where necessary, indicators are disaggregated by sex or other relevant characteristics		2 out of 7 interventions have not disaggregated indicators up to level of outcomes, in 1 intervention less than half of the indicators are disaggregated, in 2 interventions half of the indicators, in 1 intervention more than half of the indicator and in 1 intervention all indicators.	3 out of 8 interventions have not disaggregated indicators up the level of outputs, in 1 intervention less than half of indictors are disaggregated, in 2 interventions half of the indicators, in 1 intervention all indicators.	6 out of 7 interventions have not disaggregated indicators at level of impact, 1 intervention has disaggregated all indicators at the level of impact.	3 out of 6 interventions have not disaggregated indicators related to sustainability, 1 intervention has disaggregated less than half of the indicators, 1 intervention more than half of the indicators, and 1 intervention all indicators.
Proposed M&E system is a consistent translation of proposed intervention logic and ToC	In 4 out of 10 interventions no M&E system is proposed which allows to M&E of relevance, 1 intervention proposes an M&E system but it is not entirely clear, 5 interventions propose an M&E system that is entirely clear.	In 6 out of 10 interventions the M&E system is a consistent translation of the intervention logic and ToC (up to level of outcomes), two interventions propose an M&E system but it is not a consistent translation, 2 interventions do not propose an M&E system (on paper)	In 7 out of 10 interventions, the M&E system is a consistent translation of the intervention logic and ToC (up to the level of outcomes), 1 intervention proposes an M&E system but it is not a consistent translation, 2 interventions do not propose an M&E system.	In 5 out of 9 interventions no M&E system at level of impact is proposed, 1 intervention proposes an M&E system but it is not entirely clear, 4 interventions propose an M&E system at level of impact that is entirely clear.	In 6 out of 10 interventions the sustainability related results are consistently translated into an M&E system, in 2 interventions an M&E system is proposed but it is not entirely clear, in 2 interventions no M&E system is proposed.
The method for M&E of the	2 interventions do not describe the method	1 intervention does not describe method for M&E (up	1 intervention does not describe method for M&E (up to level of	5 interventions do not describe the method for M&E	1 intervention does not describe the method for M&E of the

intervention results and their sustainability is clearly described	for M&E that allows to do M&E of relevance, 4 interventions only describe the method for monitoring, 4 interventions describe a method for both monitoring and evaluation of which one is not clear.	to level of outcomes), 3 interventions only describe the method for monitoring, 3 interventions describe method for monitoring and evaluation but it is not entirely clear, 3 interventions clearly describes the method for M&E.	outputs), 3 interventions only describe the method for monitoring of which 1 is not clear, 3 interventions describe method for M&E but it is not entirely clear, 3 interventions clearly describe the method for M&E.	(up to the level of impact), 4 interventions only describe the method for monitoring, 1 intervention clearly describes the method for M&E up to the level of impact.	sustainability indicators, 4 interventions only describe the method for monitoring, 5 interventions describe the method for monitoring and evaluation but in 2 of the cases it is not entirely clear.
Method to monitor assumptions is clearly described	8 interventions do not describe the method for monitoring assumptions (up to level of outcomes), 2 interventions describe the method but it is not entirely clear.		7 interventions do not describe the method for monitoring assumptions (up to the level of outputs), 3 interventions describe the method but it is not entirely clear.	9 interventions do not describe the method for monitoring assumptions (up to level of impact), 1 intervention describes the method but it is not entirely clear.	8 interventions do not describe the method for monitoring assumptions (related to sustainability), 2 interventions describe the method but it is not entirely clear.
Method to follow up internal risks is clearly described	8 interventions do not describe the method for monitoring internal risks (up to the level of outcomes), 2 interventions describe the method but it is not entirely clear.		7 interventions do not describe the method for monitoring internal risks (up to the level of outputs), 3 interventions describe the method but it is not entirely clear.	9 interventions do not describe the method for monitoring internal risks (up to level of impact), 1 intervention describes the method but it is not entirely clear.	8 interventions do not describe the method for monitoring internal risks (related to sustainability), 2 interventions describe the method but it is not entirely clear.
Human and financial resources involved in M&E system are clearly described			1 intervention does not describe the human and financial resources involved in M&E, 3 interventions only describe human *or* financial resources and it is not entirely clear, 4 interventions describe the human *and* financial resources but it is not entirely clear, 2 interventions clearly describe the human *and* financial resources.		
Proposed MIS allows an allocation of expenditures to specific outputs and intervention			2 interventions do not propose an MIS, in 4 interventions an MIS is proposed but it does not allow an allocation of expenditures to specific outputs and activities, in 4 interventions such an allocation		

components		is possible.		
The way in which the M&E system is aligned to the national/local system is clearly described	7 interventions do not describe the way in which the M&E system (to do M&E of relevance) is aligned to country system, 3 interventions provide a description, of which 1 is not clear.	6 interventions do not describe the way in which the M&E system (up to level of outcomes) is aligned to the country system (or provide a justification for non-alignment), 4 interventions provide a description of which 1 is not clear.	7 interventions do not describe the way in which the M&E system (up to level of impact) is aligned to country system, 3 interventions provide a description, of which 1 is not clear.	6 interventions do not describe the way in which the M&E system (related to sustainability indicators) is aligned to the country system (or provide a justification for non-alignment), 4 interventions provide a description of which 2 are not clear.

3.3. Practical evaluability

3.3.1. Basic information regarding intervention implementation

Table 7 highlights that in most cases those documents are available that can reasonably be expected to be available. These include intervention proposals, DTF, baseline reports, progress reports. However, in 1 intervention two of these documents were lacking, in another intervention the last progress report was not available, whereas the baseline study was lacking in a third intervention. In most cases also the information regarding the progress in implementation up to the level of outcomes is available. Positively, in more than half of the cases, a comparison is made with the initial plan and possible reasons for deviations are described. The latter is particularly valuable information to move beyond pure monitoring. Such information is however mostly not available up to the level of impact. The fact that there is no proposed method for M&E of external assumptions (see 3.2) obviously affects the availability of this information: in most of the interventions no information is available regarding the evolution of external assumptions (which seriously affects evaluability, particularly impact evaluability). Compared to information regarding external assumptions, there is more information available regarding internal risks and potential consequences. In an exceptional case, an intervention has elaborated a table which gives an overview of the different external assumptions and internal risks, the probability of the realisation of these assumptions and risks, the way in which the risks are managed and the implications of the (non)realisation of assumptions and risks for the implementation and effect of the intervention. This is obviously particularly valuable information from an (impact) evaluability perspective.

The importance of a properly formulated M&E system (component 1.3) for the availability of information (component 2.1) is also evident rom the limited presence of disaggregated baseline information, particularly when moving upwards in the ToC, and from the fact that intervention expenditures are mostly well documented (an item that was also well dealt with in the proposed M&E system), which benefits the efficiency evaluability.

Focusing on baseline information, a distinction is made between availability of information with respect to the target group and a counterfactual. Only in one of the interventions, the notion of 'counterfactual' was introduced and in that case baseline information for indicators at impact level was collected. During discussions at the moment of individual project's debriefing it became clear that most interventions conceive of counterfactuals in an overly narrow way. While it is obvious that the first best option of 'an RCT-type of control group' is often not feasible[5], or even not desirable, in the interventions under study, there are many second-best options. Depending on the availability of resources, other options include e.g. selection of non-intervention group through matching, selection of a similar group that receives another intervention (this allows to evaluate *differential* impact), use of statistical controls, generic controls or some more complex types of before-after such as panel or time series[6]. In various of the interventions the identification of a sound second-best counterfactual would have been possible without much additional effort.

[5] Interventions referred to the fact that in many cases a similar non-intervention group does not exist because many interventions are simultaneously active in specific settings, others highlighted that the selection of a control group is not ethically acceptable, etc.

[6] See Rossi P.H., Lipsey M.W. and H.E. Freeman (2004). *Evaluation: a systematic approach*, 7th edition, Thousand Oaks: Sage, Bamberger M., J. Rugh, M. Church and L. Fort (2004). "Shoestring evaluation: designing impact evaluations under budget, time and data constraints", *American Journal of Evaluation* 25 (1): 5-37.

While the absence of information regarding a counterfactual did not really come as a surprise, information regarding the target group/institution's baseline status on different relevant indicators mentioned in the ToC was also not omnipresent. In most cases a baseline study had taken place which provides interesting information on the context of the intervention but in many instances this baseline study does not provide specific information on the target group/institution's initial status on relevant indicators. More specifically, in 3 out of 10 interventions no baseline information was available up to the level of outputs, this further increases to 5 out of 10 for information at outcome level and to 6 out of 10 for information at impact level. The absence of baseline information obviously puts into perspective the evaluability of the intervention as it hampers a comparison with the take-off situation. Moreover, it also puts into perspective an analysis of the degree of inclusion/participation of the initial target group at different stages of an intervention. In most cases an analysis of the drop-out rate provides useful information regarding the way in which an intervention reaches its target group and the degree to which the intervention is designed in such a way as to accommodate the specific target group. The absence of information on drop-outs also seriously affects the evaluability of the intervention: in most cases it leads to an overestimation of the intervention's effect.

In the majority of the cases there is some information available regarding the data collection process. In 3 out of 10 interventions there is complete information including details on what, who, frequency, coverage and methods of data collection for indicators up to the level of outcomes, but only for one case such detailed information is also available regarding collection of data at the impact level. In the other cases there is mostly some (fragmented) information available, particularly up to the level of outputs, while there is no information available for data collection up to the level of outcomes and impact in 3 out of 10 and 9 interventions respectively. As regards the reliability of the data collection process, in 3 out of 10 cases the most important conditions for reliable data collection are satisfied for effectiveness and efficiency evaluation, as regards impact, sustainability and relevance evaluation this is only applicable to 2 interventions. In one case there are serious doubts about reliability of data collection up to the level of outputs, whereas in 3 and 4 cases similar doubts exist regarding data collection up to the level of outcomes and impact respectively.

Table 7: Basic information regarding intervention implementation: overview of strengths and weaknesses

OECD/DAC criteria Items	Relevance	Effectiveness	Efficiency	Impact	Sustainability
The basic documents are available	In 6 out of the 10 interventions all basic documents that can reasonably be expected (intervention proposals, DTF, progress reports, baseline reports, ...) are available. In 1 intervention two or more of the documents were not available, in 1 intervention all documents are available except for baseline, in 2 interventions all documents are available except for the last progress report (which was not available three months after the foreseen deadline)				
Baseline information with regard to target group is available		In 5 of the 10 interventions no baseline information was available for the beneficiaries up the level of outcomes, in the 5 other interventions baseline information was available for all relevant indicators up to outcome level.	For 3 out of 10 interventions no baseline information was available for the beneficiaries up to the level of outputs, in 1 case baseline information was available for some of the relevant indicators, for 6 interventions baseline information was available for all relevant indicators up to the level of outputs.	For 6 out of 10 interventions no baseline information was available for the beneficiaries at the level of impact, for 4 interventions baseline information was available for all relevant indicators at the level of impact.	For 6 out of 10 interventions no baseline information was available for the beneficiaries related to sustainability indicators, for 4 interventions baseline information was available for all relevant sustainability indicators.
Baseline information with regard to counterfactual is available		In none of the interventions information was available regarding the counterfactual (up to information at level of outputs & outcome)		In 1 out of 7 interventions information was available regarding the counterfactual at the level of impact	In none of the interventions information was available regarding the counterfactual (related to sustainability indicators)
Disaggregated baseline information according to gender (or other relevant parameter) is available		In 4 out of 6 interventions disaggregated baseline information (up to level of outcomes) was not available, in 1 intervention disaggregated baseline information is available for some relevant indictors, in 1 intervention for all relevant indicators.	In 3 out of 6 interventions disaggregated data up to the level of output was not available, in 2 interventions disaggregated data is available for some of the relevant indicators, in 1 intervention for all relevant indicators.	In 6 out of 7 interventions disaggregated data was not available at the level of impact, in 1 intervention for all relevant indicators at impact level.	
Information available regarding progress in implementation		For 7 out of 9 interventions information is available regarding the progress of implementation (up to level of outcomes) and	For 8 out of 10 interventions information is available regarding the progress of implementation (up to level of outputs) and comparison is	For 5 out of 6 interventions no information is available regarding the progress of implementation (up to level of impact), for 1 intervention a	

		comparison is made with initial plan and reasons for deviations are described, in 1 case no information is available regarding progress, in 1 intervention information is available regarding progress but no comparison is made with what was initially foreseen.	made with initial plan and reasons for deviations are described, in 1 case information is available, comparison is made but reasons for deviations are not described, in 1 intervention no information is available regarding progress.	comparison is made with initial plan and reasons for deviations are described	
Information available regarding participation of initial target group	In 3 out of 10 interventions no information is available on participation of initial target group, in 4 interventions some information is available, in 3 interventions there is complete information on participation and discussion of reasons for changes and drop-outs.		In 2 out of 10 interventions no information is available on participation of initial target group, in 5 interventions some information is available (up to level of outputs), in 3 interventions there is complete information on participation and discussion of reasons for changes and drop-outs.	In 3 out of 10 interventions no information is available on participation of initial target group, in 4 interventions some information is available, in 3 interventions there is complete information on participation and discussion of reasons for changes and drop-outs.	In 3 out of 10 interventions no information is available on participation of initial target group, in 4 interventions some information is available, in 3 interventions there is complete information on participation and discussion of reasons for changes and drop-outs.
Information available regarding data collection process for the indicators	For 3 out of 10 interventions no information is available regarding the data collection process for indicators (relevance), for 2 others there is reasonable amount of information, for 2 interventions there is relatively complete information, for 3 interventions there is complete information (who, what, frequency, coverage, methods, etc.)	In 1 out of 10 interventions, no information is available regarding the data collection process for indicators up to the level of outcomes, for 2 interventions there is only fragmentary information, for 2 others there is reasonable amount of information, for 2 interventions there is relatively complete information, for 3 interventions there is complete information (who, what, frequency, coverage, methods, etc.)	For all interventions information is available regarding the data collection process for indicators up to the level of outputs, for 2 interventions there is only fragmentary information, for 3 others there is reasonable amount of information, for 2 interventions there is relatively complete information, for 3 interventions there is complete information (who, what, frequency, coverage, methods, etc.)	For 3 out of 9 interventions no information is available regarding the data collection process for indicators at the level of impact, for 1 intervention there is only fragmentary information, for 3 others there is reasonable amount of information, for 1 intervention there is relatively complete information, for 1 intervention there is complete information (who, what, frequency, coverage, methods, etc.)	In 1 out of 10 interventions, no information is available regarding the data collection process for sustainability indicators, for 2 interventions there is only fragmentary information, for 2 others there is reasonable amount of information, for 2 interventions there is relatively complete information, for 3 interventions there is complete information (who, what, frequency, coverage, methods, etc.)
Data collection allows in principle reliable	For 3 out of 9 interventions there is no	For 3 out of 9 interventions there is no proposal for	For 1 out of 9 interventions there is no proposal for data	For 4 out of 9 interventions there is no proposal for data	For 2 out of 9 interventions there is no proposal for data

data collection of indicators	proposal for data collection (for relevance related indicators), for 4 interventions the proposal for data collection satisfies two of three aspects of reliable data collection, for 2 interventions all aspects of reliable data collection are satisfied.	data collection (up to level of outcomes), for 3 interventions proposal for data collection satisfies two of three aspects of reliable data collection, for 3 interventions all aspects of reliable data collection are satisfied.	collection (up to level of outputs), for 5 interventions proposal for data collection satisfies two of three aspects of reliable data collection, for 3 interventions all aspects of reliable data collection are satisfied.	collection (up to level of outputs), for 3 interventions the proposal for data collection satisfies two of three aspects of reliable data collection, for 2 interventions all aspects of reliable data collection are satisfied.	collection (sustainability indicators), for 5 interventions proposal for data collection satisfies two of three aspects of reliable data collection, for 2 interventions all aspects of reliable data collection are satisfied.
Information regarding monitoring if internal risks is available and potential consequences highlighted		For 4 interventions no information is available regarding monitoring of internal risks (up to level of outcomes), for 2 interventions information is available but potential consequences are not documented, for 4 interventions also potential consequences are documented.	For 2 interventions no information is available regarding monitoring of internal risks (up to level of outputs), for 2 interventions information is available but potential consequences are not documented, for 6 interventions also potential consequences are documented.	For 9 interventions no information is available regarding monitoring of internal risks (up to level of impact), for 1 intervention information is available and potential consequences are documented as well.	For 3 interventions no information is available regarding monitoring of internal risks (related to sustainability dimension), for 2 interventions information is available but potential consequences are not documented, for 5 interventions also potential consequences are documented.
Information regarding monitoring of assumptions and consequences is available	For 9 interventions no information is available regarding monitoring of assumptions (relevance related), for 1 intervention information is available and potential consequences are documented as well.	For 6 out of 8 interventions no information is available regarding monitoring of assumptions (up to level of outcomes), for 1 intervention information is available but potential consequences are not documented, for 1 intervention also potential consequences are documented.	For 7 out of 8 interventions no information is available regarding monitoring of assumptions (up to level of outputs), for 1 intervention information is available but potential consequences are not documented, for 2 interventions also potential consequences are documented.	For 9 interventions no information is available regarding monitoring of assumptions (up to level of impact), for 1 intervention information is available and potential consequences are documented as well.	For 8 out of 10 interventions no information is available regarding monitoring of assumptions (related to sustainability), for 1 intervention information is available but potential consequences are not documented, for 1 intervention also potential consequences are documented.
Intervention expenditures well documented and linked to outputs			For 5 interventions expenditures are well documented and linked to outputs (but there are expenditures that are recorded elsewhere), for 5		

			interventions expenditures are well documented and linked to outputs and no expenditures are recorded elsewhere		

3.3.2. The M&E system in practice

Table 8 highlights that in practice the M&E system functions relatively well. This particularly holds for the 'monitoring' component of the system and to a much lesser extent for the more 'evaluative' activities. In some of the cases where 'monitoring' systems were not clearly described on paper, they tend to function better than expected on the ground. This does not hold for the 'evaluative' component of M&E systems which is in all interventions much less developed in terms of vision, clear allocation of responsibilities, time, expertise, level of independence, etc... It then does not come as a surprise that M&E systems are mainly geared towards accountability rather than learning as it is particularly the learning objective that depends to a much higher extent on the analytical (evaluative) component of the system. The existing M&E systems are overall less well apt to provide inputs for sound impact evaluation and evaluation of 'relevance'.

Most interventions (8 to 9) describe the vision on and the role of M&E, however only for 1 intervention the vision and the role are also formally elaborated in one document which is also endorsed by the main stakeholders. In about half of the other interventions several elements are described but these are not necessarily coherent or coordinated. In all interventions the vision and role of monitoring is more clearly described then the one on evaluation. The notions of 'monitoring' and 'evaluation' are also often used interchangeably and where M&E is used as a notion, the focus is in reality on 'monitoring'. In 6 out of 10 interventions the most important stakeholders also agree on the proposed M&E system, while in 2 out of the 10 cases there are disagreements among important stakeholders regarding the proposed M&E system. In the latter cases, the disagreements were mainly regarding the direction and intensity of information flows. More specifically, local beneficiaries or partner organisations who are (extensively) involved in data collection and upwards channelling of information considered the feedback from the more centrally located actors too weak. While at this stage the beneficiaries and local partner organisations are not (yet) demotivated to collect data, this may become a real possibility in the future as it is particularly feedback and the use of collected data that incentivize people and organisations to invest in and improve data collection. Without such feedback data collection may turn into a kind of 'ritual exercise' which may also undermine the reliability of the data collected and the sustainability of the system in itself. This observation also holds true for unbalanced information flows between actors at higher levels in the chain (field – HQ, non-government actor – DGD, etc.) and hints at the importance of investing as well in 'downwards' feedback.

Whereas many interventions did not clearly describe if and how their M&E system is aligned to the local or central M&E system (see component 1.3), in reality, more than half of the interventions are to a varying extent aligned to the local or central M&E system or can justify why there is no alignment. This alignment of the system in practice does not come as a surprise, because it is indicative of the Rwandan government's drive towards country-ownership. The fact that there is alignment to the system in practice does however not imply that an explicit description of the alignment (or non-alignment) is useless. A first step in any alignment exercise is a diagnosis of the quality of the existing system, its major strengths and weaknesses. Such a diagnosis might highlight that (full) alignment is not (yet) feasible (and maybe not desirable) while it might as well be a first step in capacity building efforts of existing M&E systems. Strengthening national/local M&E systems might have an effect on the evaluability of interventions, particularly in those cases where ex-post evaluations are taking place (and which are often difficult to realise because of the breakdown of intervention M&E systems at the end of an intervention). Alignment to national M&E systems does

however not imply that additional (independent) complementary evaluative exercises are useless; to the contrary, triangulating among various information streams may increase an evaluation's quality.

Responsibilities and procedures regarding the collection and analysis of information are more clearly defined for collection of data (particularly up to level of outputs) than for analysis of data and this is particularly true as regards procedures for data collection and analysis at the impact level. Similarly, resources (time, financial and human resources) for M&E are in most interventions much more geared towards 'monitoring' rather than 'evaluative' activities. In most of the interventions large amounts of data are being collected that are never analysed (or used). In about half of the interventions resources (time, financial and human resources) are in principle sufficient (up to the level of outcome), but they are mostly not well spread over 'monitoring' on the one hand and 'evaluative' activities on the other hand. In many cases the staff involved is at least moderately competent and/or experienced for monitoring activities but much less for evaluative activities. However, in 1 to 2 cases there were serious concerns about the experience and competence levels of the staff involved in M&E and/or their degree of independence. As regards M&E at impact level, this even holds for 3 interventions. Given the size of some of the interventions/organisations, it is sometimes understandable that there is no specific staff available for M&E. In such circumstances, it might be worthwhile to envisage the possibilities to explore opportunities to set up a kind of pooled fund with different organisations to finance M&E expertise with each organisations having a number of expertise days available (kind of drawing rights). Other possibilities might be to look for cooperation with local universities (which could as well be linked to projects of university development cooperation). This might in particular be an interesting route for evaluation as in many cases evaluation is also interesting beyond the boundaries of the own intervention, whereas monitoring is much more closely related to specific interventions and thus less easy to share among different interventions. Another way to stimulate evaluative analysis within the interventions may be to use it earlier in the intervention which may also increase its learning potential. An interesting route to explore is the practice of 'structured experiential learning'[7] (M**e**E) which refers to evaluative activities ('e') between 'monitoring' (M) and 'evaluation' (E) that are considered particularly useful for innovative interventions (with complex programme theories) where there is not enough knowledge yet regarding the effectiveness and efficiency of different implementation modalities. As discussed in Pritchett et al. (2013)[8], in such circumstances a more flexible type of ex-ante design which underscores the 'unknown' and allows for experimentation with different implementation modalities and close monitoring and comparison of these different implementation modalities is advisable. During the debriefing, it was highlighted that such experiences already exist in some of the interventions, yet these are not documented.

In 8 out of 10 interventions, there is a clear interest in M&E (which may also be related to the context, see below) and an internal drive for strategic management and learning. In half of the interventions this is also supported by the HQ of the key players. Turning to the use of M&E findings, it is clear that there is in the large majority (8 to 9) of the cases not really a systematic approach towards the use of M&E for accountability and learning. The ad-hoc use of M&E does, however, lead to a focus on accountability rather than on learning which is entirely in line with the predominance of

[7] See Pritchett, L., S. Samji and L. Hammer (2013)."It's all about MeE: Using structured experiential learning to crawl the design space", *Working Paper 322*. Washington D.C., Center for Global Development.

monitoring at the expense of evaluation. As mentioned above, the latter is not unique for the 10 interventions under study but indicative of a general tendency within development cooperation to consider 'results based management' as a synonym of 'evaluation'. Moving towards the level of meta-evaluation (evaluation/review of the evaluation (system)), 4 interventions regularly and systematically review the quality of the M&E system and adjust accordingly, whereas 4 interventions have never done such an exercise, and the remaining two review their system but not in a systematic way.

Table 8: The M&E system in practice: Overview of strengths and weaknesses

OECD/DAC criteria Items	Relevance	Effectiveness	Efficiency	Impact	Sustainability
Vision on and role of M&E are clear		In 1 intervention the vision and the role of M&E (up to level of outcomes) are not described, in 2 interventions different visions are described but these are not clear and not internally coordinated, in 2 interventions several elements are described which are clear but not coherent, coordinated or aligned, in 4 interventions there is a high degree of consensus and 1 intervention the vision and role are formally elaborated in one document which is endorsed by the main stakeholders.	In 2 interventions different visions on role of M&E up to level of outputs are described but these are not clear and not internally coordinated, in 3 interventions several elements are described which are clear but not coherent, coordinated or aligned, in 4 interventions there is a high degree of consensus and 1 intervention the vision and role are formally elaborated in one document which is endorsed by the main stakeholders.	In 2 intervention the vision and the role of M&E (up to level of impact) are not described, in 1 intervention different visions are described but these are not clear and not internally coordinated, in 2 interventions several elements are described which are clear but not coherent, coordinated or aligned, in 4 interventions there is a high degree of consensus and 1 intervention the vision and role are formally elaborated in one document which is endorsed by the main stakeholders.	In 1 intervention the vision and the role of M&E (related to sustainability) are not described, in 2 interventions different visions are described but these are not clear and not internally coordinated, in 2 interventions several elements are described which are clear but not coherent, coordinated or aligned, in 4 interventions there is a high degree of consensus and 1 intervention the vision and role are formally elaborated in one document which is endorsed by the main stakeholders.
Most important stakeholders agreed upon the proposed M&E system	In 6 interventions the most important stakeholders agreed upon the proposed M&E system (up to level of outcomes), in 2 interventions some of the stakeholders agreed and in 2 interventions most of the important stakeholders disagreed upon the proposed M&E system.		In 6 interventions the most important stakeholders agreed upon the proposed M&E system (up to level of outputs), in 3 interventions some of the stakeholders agreed and in 1 intervention most of the important stakeholders disagreed upon the proposed M&E system.	In 6 interventions the most important stakeholders agreed upon the proposed M&E system (up to level of impact), in 2 interventions some of the stakeholders agreed and in 2 interventions most of the important stakeholders disagreed upon the proposed M&E system.	In 6 interventions the most important stakeholders agreed upon the proposed M&E system (for sustainability indicators), in 2 interventions some of the stakeholders agreed and in 2 interventions most of the important stakeholders disagreed upon the proposed M&E system.
There are sufficient	In 4 interventions	In 4 interventions there	In 5 interventions there are	In 3 interventions there are	In 4 interventions there are

resources provided to allow for adequate functioning of M&E system	there are sufficient resources for adequate functioning of M&E system (for relevance dimension), in 3 interventions one or two of the resources are insufficient, in 3 interventions there are insufficient resources.	are sufficient resources for adequate functioning of M&E system (up to level of outcomes), in 5 interventions one or two of the resources are insufficient, in 1 intervention there are insufficient resources.	sufficient resources for adequate functioning of M&E system (up to level of outputs), in 5 interventions one or two of the resources are insufficient.	sufficient resources for adequate functioning of M&E system (up to level of impact), in 4 interventions one or two of the resources are insufficient, in 3 interventions there are insufficient resources.	sufficient resources for adequate functioning of M&E system (for sustainability indicators), in 5 interventions one or two of the resources are insufficient, in 1 intervention there are insufficient resources.
The responsibilities and procedures regarding the collection and analysis of information are clearly defined	In 3 interventions responsibilities are clearly described (for M&E related to relevance), in 4 interventions responsibilities are only fragmentarily described, in 3 intervention it is unclear who is responsible for collection and analysis of data.	In 3 interventions responsibilities are clearly described (up to level of outcomes), in 6 interventions responsibilities are only fragmentarily described, in 1 intervention it is unclear who is responsible for collection and analysis of data.	In 4 interventions responsibilities are clearly described (up to level of outputs), in 6 interventions responsibilities are only fragmentarily described.	In 2 interventions responsibilities are clearly described (up to level of impact), in 1 interventions responsibilities are only fragmentarily described, in 7 interventions procedures for M&E at level of impact are not described.	In 3 interventions responsibilities are clearly described (sustainability indicators), in 6 interventions responsibilities are only fragmentarily described, in 1 intervention it is unclear who is responsible for collection and analysis of data related to sustainability indicators.
The responsibilities and procedures regarding decision-making based on M&E information are clearly defined	In 9 interventions the responsibilities and the procedures for decision-making are established and clear (up to level of outcomes and from the perspective of relevance). In 1 intervention these responsibilities are not clear.		In 10 interventions the responsibilities and the procedures for decision-making (up to level of outputs) are established and clear.	In 4 intervention the responsibilities and the procedures for decision-making are established and clear (up to level of impact). In 6 interventions these responsibilities are not clear.	In 9 interventions the responsibilities and the procedures for decision-making are established and clear (related to sustainability). In 1 intervention these responsibilities are not clear.
Staff responsible for M&E is competent and independent	In 1 intervention staff responsible for M&E (related to relevance) is highly competent, experienced and independence is formalised, in 5 interventions staff	In 1 intervention staff responsible for M&E (up to level of outcome) is highly competent, experienced and independence is formalised, in 6 interventions staff has	In 1 intervention staff responsible for M&E (up to level of output) is highly competent, experienced and independence is formalised, in 1 interventions staff has strong experience and competences and independence can reasonably be assumed, in 6 interventions staff	In none of the interventions staff responsible for M&E is highly competent, experienced and independence formalised for M&E at impact level, in 6 interventions staff has moderate experience and	In 1 intervention staff responsible for M&E (sustainability indicators) is highly competent, experienced and independence is formalised, in 6 interventions staff has moderate experience and competences and independence can reasonably be

	has moderate experience and competences and independence can reasonably be assumed, in 2 interventions staff has limited experience and competences and independence is limited, in 2 interventions staff has no experience and competence and independence is not guaranteed.	moderate experience and competences and independence can reasonably be assumed, in 2 interventions staff has limited experience and competences and independence is limited, in 1 intervention staff has no experience and competence and independence is not guaranteed.	has moderate experience and competences and independence can reasonably be assumed, in 2 interventions staff has limited experience and competences and independence is limited.	competences and independence can reasonably be assumed, in 1 intervention staff has limited experience and competences and independence is limited, in 3 interventions staff has no experience and competence and independence is not guaranteed.	assumed, in 2 interventions staff has limited experience and competences and independence is limited, in 1 intervention staff has no experience and competence and independence is not guaranteed.
M&E system is aligned/related to the national M&E system	In 6 interventions the M&E system is aligned to local/national M&E system or it is justified why it is not aligned, in 2 interventions M&E system is not aligned and it is partly justified why it is not, in 2 cases it is not aligned and it is not justified why not.		In 6 interventions the M&E system is aligned to local/national M&E system or it is justified why it is not aligned, in 3 interventions M&E system is not aligned and it is partly justified why it is not, in 1 case it is not aligned and it is not justified why not.	In 5 interventions the M&E system is aligned to local/national M&E system or it is justified why it is not aligned, in 2 interventions M&E system is not aligned and it is partly justified why it is not, in 3 cases it is not aligned and it is not justified why not.	In 6 interventions the M&E system is aligned to local/national M&E system or it is justified why it is not aligned, in 2 interventions M&E system is not aligned and it is partly justified why it is not, in 2 cases it is not aligned and it is not justified why not.
There is an internal drive for strategic management and learning	In 5 interventions there is real and strong interest of all stakeholders in M&E findings (up to level of outcome and from relevance perspective) for purposes of strategic management and learning which is also supported by HQ of key players, in 3 interventions there is some interest of stakeholders in M&E findings, in 2 interventions there is no real interest (M&E is perceived as a threat or an necessary evil)		In 6 interventions there is real and strong interest of all stakeholders in M&E findings (up to level of outputs) for purposes of strategic management and learning which is also supported by HQ of key players, in 3 interventions there is some interest of stakeholders in M&E findings, in 1 intervention there is no real interest (M&E is perceived as a threat or an necessary evil)	In 5 interventions there is real and strong interest of all stakeholders in M&E findings (up to level of impact) for purposes of strategic management and learning which is also supported by HQ of key players, in 3 interventions there is some interest of stakeholders in M&E findings, in 2 intervention there is no real interest (M&E is perceived as a threat or an necessary evil)	In 5 interventions there is real and strong interest of all stakeholders in M&E findings (related to sustainability) for purposes of strategic management and learning which is also supported by HQ of key players, in 3 interventions there is some interest of stakeholders in M&E findings, in 2 interventions there is no real interest (M&E is perceived as a threat or an necessary evil)

M&E results used for learning	In 1 intervention there is a strategy to use M&E (up to level of outcomes and related to relevance) systematically for learning purposes which targets all stakeholders, in 6 interventions M&E results are used for learning but not in a systematic way (rather ad-hoc), in 1 case M&E results are only implicitly used for learning, in 2 cases no efforts are made to use M&E for learning.	In 2 interventions there is a strategy to use M&E systematically for learning purposes (up to level of outputs) which targets all stakeholders, in 6 interventions M&E results are used for learning but not in a systematic way (rather ad-hoc), in 1 case M&E results are only implicitly used for learning, in 1 case no efforts are made to use M&E for learning.	In 1 intervention there is a strategy to use M&E (up to level of impact) systematically for learning purposes which targets all stakeholders, in 6 interventions M&E results are used for learning but not in a systematic way (rather ad-hoc), in 1 case M&E results are only implicitly used for learning, in 2 cases no efforts are made to use M&E for learning.	In 1 intervention there is a strategy to use M&E (related to sustainability) systematically for learning purposes which targets all stakeholders, in 6 interventions M&E results are used for learning but not in a systematic way (rather ad-hoc), in 1 case M&E results are only implicitly used for learning, in 2 cases no efforts are made to use M&E for learning.
M&E results used for accountability	In 1 out of 10 interventions there is a strategy to use M&E (up to level of outcomes and from a relevance perspective) systematically for accountability purposes which targets all stakeholders, in 2 cases a strategy has been elaborated but not all stakeholders are covered, in 6 cases results are used for accountability but not in a systematic way (ad-hoc), in 1 case no efforts are made to use results for accountability.	In 2 out of 10 interventions there is a strategy to use M&E (up to level of outputs) systematically for accountability purposes which targets all stakeholders, in 2 cases a strategy has been elaborated but not all stakeholders are covered, in 6 cases results are used for accountability but not in a systematic way (ad-hoc).	In 1 out of 10 interventions there is a strategy to use M&E (up to level of impact) systematically for accountability purposes which targets all stakeholders, in 1 case a strategy has been elaborated but not all stakeholders are covered, in 6 cases results are used for accountability but not in a systematic way (ad-hoc), in 1 case no efforts are made to use results for accountability.	In 1 out of 10 interventions there is a strategy to use M&E (related to sustainability) systematically for accountability purposes which targets all stakeholders, in 2 cases a strategy has been elaborated but not all stakeholders are covered, in 6 cases results are used for accountability but not in a systematic way (ad-hoc), in 1 case no efforts are made to use results for accountability.
Evaluations/studies have been conducted (where it can be expected) that are of good quality and provide useful information	In 1 of the 5 interventions where evaluations could be expected, no evaluations or studies have taken place (up to level of outcomes and from a relevance perspective). In 3 cases a good independent evaluation took place with some minor limitations, in 1 case a good independent evaluation took place that covered the entire intervention which produced useful information.	In 1 of the 5 interventions where evaluations could be expected, only a few activities with internal evaluative character have taken place (up to level of outputs). In 3 cases a good independent evaluation took place with some minor limitations, in 1 case a good independent evaluation took place that covered the entire intervention which produced useful information.	In 2 of the 5 interventions where evaluations could be expected, no evaluations or studies have taken place (up to level of impact). In 3 cases a good independent evaluation took place with some minor limitations.	In 1 of the 5 interventions where evaluations could be expected, no evaluations or studies have taken place (related to sustainability dimension). In 3 cases a good independent evaluation took place with some minor limitations, in 1 case a good independent evaluation took place that covered the entire intervention which produced useful information.

M&E system is regularly reviewed on its quality and adjusted accordingly	In 4 cases the M&E system is systematically reviewed and adjusted accordingly, in 2 cases the M&E system is reviewed but not in a systematic manner, in 4 cases no reviews of the quality of the M&E system is taking place.

3.4. Influence of contextual factors

As regards the influence of contextual factors, some difference was discernible among different channels of intervention. This is not very surprising as the degree of involvement of government actors and the interplay among the intervention and government actors is different in different intervention channels. Whereas efforts were done to limit social desirable answering (by asking specific questions and by triangulating among different interviewees' answers), it cannot entirely be excluded as in some of the cases we are probing into future behaviour *('what is likely to happen in future evaluations'*). Additionally it cannot be ruled out that the same contextual factors that influence monitoring and especially evaluations have also influenced this review and have thus led to some degree of social desirable answering.

3.4.1. Attitude of key players

From different items discussed in Table 9 it is clear that the most important key players have a positive attitude towards evaluation. In most (8 to 9) interventions (regardless of the intervention channel) the majority (if not all) of the key actors express a keen interest and demand for evaluation. The existing 'performance' culture in Rwanda, referred to by almost all interviewees, is certainly one of the factors that is of influence here (see also below). This is also obvious from the high degree of involvement of different stakeholders in evaluation processes. In almost all interventions at least some of the important stakeholders (not beneficiaries) have been involved in evaluation processes or will be involved in future evaluation processes and in more than half of the cases all the important stakeholders (including beneficiaries) are/will be involved. In 6 out of 10 cases it seems possible to contact all stakeholders without interference while in 4 cases it is clear that certain stakeholders (would) find it difficult not to interfere in the execution of an evaluation. As regards relationships among stakeholders, in 5 out of 10 cases the relationships can be considered constructive, in 4 cases there were no considerable tensions, whereas in 1 case there was an open conflict among the most important actors involved.

In most interventions expectations of the most important stakeholders are realistic and also mutually compatible. However, the identification of the most important users and their specific expectations is less clear-cut. As there tends to be some trade-off among different objectives of evaluations, with different types of evaluations emphasizing different principles and methodologies (e.g. accountability-focused evaluations need external validity and representativity whereas learning-focused evaluations rather focus on some specific dimensions and benefit from a higher degree of internal validity), it may jeopardize an evaluation's quality when the principal users' needs are not clearly identified or when too many objectives are to be reached and too many OECD/DAC criteria are to be covered by one single evaluation. In some of the interventions we also noticed a slight difference in preferences for particular evaluation criteria with some stakeholders (often government actors) having a preference for effectiveness and efficiency criteria (focus on 'are we doing the things right?') whereas others (often beneficiaries) tend to be more interested in impact and relevance (focus on 'are we doing the rights things?').

Table 9: Attitude of key players: Overview of strengths and weaknesses

OECD/DAC criteria Items	Relevance	Effectiveness (check zou normal gelijk moeten zijn aan de rest)	Efficiency	Impact	Sustainability
Principal users of evaluation and expectations are clearly defined	In 4 out of 6 cases the principal users of the evaluation and their expectations are not defined, in 2 cases the principal users and their expectations are defined.				
Key actors express demand for evaluation	In 5 out of 10 cases, the key actors express demand for evaluation or are at least interested in evaluation (related to relevance, effectiveness and efficiency). In 4 cases the majority of the key actors express demand and in 1 case none of the key actors expresses demand.			In 5 out of 10 cases, the key actors express demand for evaluation or are at least interested in evaluation (up to impact). In 3 cases the majority of the key actors express demand and in 2 cases none of the key actors expresses demand.	In 5 out of 10 cases, the key actors express demand for evaluation or are at least interested in evaluation (related to sustainability). In 4 cases the majority of the key actors express demand and in 1 case none of the key actors expresses demand.
Expectations of key actors regarding evaluation are mutually compatible	In 6 out of 8 cases the expectations of the key actors regarding evaluation (regarding relevance, effectiveness and efficiency) are mutually compatible, in 1 case there are manageable differences regarding the process or the results, in 1 case there are large differences regarding process and results.				In 6 out of 8 cases the expectations of the key actors regarding evaluation (related to sustainability) are mutually compatible, in 2 cases there are manageable differences regarding the process or the results.
Expectations of key stakeholders are realistic	In 7 out of 9 cases, all key stakeholders have realistic expectations regarding evaluation (related to relevance and effectiveness), in 2 cases several key stakeholders have unrealistic expectations.		In 8 out of 9 cases, all key stakeholders have realistic expectations regarding evaluation (related to efficiency), in1 case several key stakeholders have unrealistic expectations.	In 7 out of 9 cases, all key stakeholders have realistic expectations regarding evaluation (up to level of impact), in 2 cases several key stakeholders have unrealistic expectations.	In 8 out of 9 cases, all key stakeholders have realistic expectations regarding evaluation (related to sustainability), in1 case several key stakeholders have unrealistic expectations.
Principal users were/will be involved in evaluation process	In 3 out of 4 cases, all principal users were/will be involved in the evaluation process, in one case a number of users were/will be involved.				
Most important stakeholders are/will be involved in	In 5 out of 9 cases, most important stakeholders were/will be involved in the evaluation process, in 4 cases some of the important stakeholders were/will be involved in evaluation process (but not the beneficiaries)				

evaluation process	
Relationships among stakeholders are healthy	In 5 out of the 10 cases the relations among the stakeholders can be considered as constructive, critical partnerships, in 4 cases there are no considerable tensions whereas in 1 case there are serious and open conflicts among stakeholders.
It is possible to contact all stakeholders without risking reciprocal influence	In 6 out of 10 cases it is possible to contact all stakeholders without interference, in 4 cases it is clear that certain stakeholders find/would find it difficult not to interfere in the execution of an evaluation.
Positive attitude towards independent evaluation	In 6 out of 10 interventions, all stakeholders have a positive attitude towards evaluation, in 3 interventions the number of those have a negative attitude is similar to the number having a positive attitude, in 1 case almost all stakeholders have a negative attitude towards independent evaluation

3.4.2. Broader context

An overview of the influence that the broader context might have on the evaluability of interventions in Rwanda is provided in Table 10. As mentioned above there is currently a strong 'performance' culture in Rwanda which has a number of positive consequences in terms of evaluability while it as well raises a number of challenges. More specifically, many actors have an interest and demand for evaluation which drives the supply of M&E and also influences the sustainability of the M&E (system) itself. Local expertise is generally and increasingly available, although sometimes considerable efforts are required to identify and engage specific local expertise.

Reversely, the performance culture and the strong focus on 'accountability within the system', and particularly the drive to arrive quickly at 'positive' results, may as well undermine the quality and reliability of data collection as individuals and organisations may be reluctant to disclose 'negative' results. It may as well negatively affect people's and organisation's scope for experimentation and put all the emphasis and efforts on those issues and items that are in 'performance contracts' and being monitored at the detriment of those issues that are outside such contracts. Another contextual element which influences evaluability (and development interventions in general) is the speed at which changes in national policies are taking place and are being implemented. At the least it is important for donors to acknowledge the need for flexibility.

The downsides of the performance culture obviously jeopardize evaluability and particularly the evaluation's learning function. While there is a keen interest in evaluation, not all interviewees considered it straightforward to conduct independent evaluations, or there were at the least different opinions among stakeholders within the same intervention. While collection of data at the level of beneficiaries was possible in 6 out of 10 interventions, some interviewees also highlighted that it was impossible to do this in an independent manner. Along the same lines, some interviewees also pointed out that collection of data on individuals that are not among the intervention's beneficiaries is even more difficult which obviously puts into perspective the possibility of using a 'counterfactual'.

Table 10: The broader context: overview of opportunities and threaths

OECD/DAC criteria Items	Relevance	Effectiveness	Efficiency	Impact	Sustainability
The broader political and institutional context is positive towards independent evaluation	In 3 out of 10 interventions, the principle of independent evaluation was accepted and embedded (for relevance), in 4 interventions there were different attitudes towards independent evaluations, in 3 interventions conducting independent evaluation was difficult at the level of the concerned institutions in the country.	In 3 out of 10 interventions, the principle of independent evaluation was accepted and embedded (for effectiveness and efficiency), in 5 interventions there were different attitudes towards independent evaluations, in 2 interventions conducting independent evaluation was difficult at the level of the concerned institutions in the country.		In 3 out of 10 interventions, the principle of independent evaluation was accepted and embedded (for impact), in 4 interventions there were different attitudes towards independent evaluations, in 3 interventions conducting independent evaluation was difficult at the level of the concerned institutions in the country.	In 3 out of 10 interventions, the principle of independent evaluation was accepted and embedded (for sustainability), in 5 interventions there were different attitudes towards independent evaluations, in 2 interventions conducting independent evaluation was difficult at the level of the concerned institutions in the country.
The socio-cultural context at the level of beneficiaries allows for adequate data collection	In 6 out of 10 interventions collection of data from and about beneficiaries can be achieved without great effort/challenge, in 3 interventions collection of data is feasible (but not without challenge or effort), in 1 intervention collection of data to and from the beneficiaries is very difficult.	In 6 out of 10 interventions collection of data from and about beneficiaries can be achieved without great effort/challenge, in 4 interventions collection of data is feasible (but not without challenge or effort).		In 6 out of 10 interventions collection of data from and about beneficiaries can be achieved without great effort/challenge, in 3 interventions collection of data is feasible (but not without challenge or effort), in 1 intervention collection of data to and from the beneficiaries is very difficult.	In 6 out of 10 interventions collection of data from and about beneficiaries can be achieved without great effort/challenge, in 4 interventions collection of data is feasible (but not without challenge or effort).
Local expertise with the required profile is available	In 4 interventions it was sufficient local expertise with the required profile, for 5 interventions it is possible to engage experts with the required profile but it requires considerable effort, in 1 intervention it is very difficult to engage experts with the required profile.			In 3 interventions it was sufficient local expertise with the required profile, for 6 interventions it is possible to engage experts with the required profile but it requires considerable effort, in 1 intervention it is very difficult to engage experts with the required profile.	In 4 interventions it was sufficient local expertise with the required profile, for 5 interventions it is possible to engage experts with the required profile but it requires considerable effort, in 1 intervention it is very difficult to engage experts with the required profile.

Annex 1. Review framework

	This review framework has been developed in the context of the assignment 'Practical Evaluability of development interventions' commissioned by the Office of the Special Evaluator of Belgian Development Cooperation. The framework has been elaborated on the basis of existing literature and insights of the different study team members (drawing upon practice and evaluation theory). So far, there does not exist a standardized framework or methodology to assess evaluability and also this framework is work in progress. The review itself will test the framework's practical applicability and feed into its further refinement. *The OECD/DAC definitions of the different criteria are used. For some items the scores on the different criteria can be the same, while for others they will differ.*	**relevance**	**effectiveness**	**efficiency**	**Impact**	**sustainability**
1.	**Analysis of the intervention design** ***(Part 1 is similar to what is labelled 'theoretical evaluability in the TOR)***					
1.1	*The underlying analysis*	X	X			
1.1.1	The beneficiaries (target groups) are clearly identified (demarcated) and described	X	X			
1.1.2	The rationale of the intervention and the (problem) situation of the beneficiaries is clearly described	X				
1.1.3	The role of the beneficiaries is clearly described	X	X			
1.1.4	The role of the most important actors (exclusive of beneficiaries) is clearly described	X	X			
1.1.5	Gender analysis is integrated in the analysis	X	X			
1.1.6	The link between analysis and intervention objectives is clearly described	X				
1.1.7	The linkage between the rationale of the intervention and the sector policy of the partner country is clearly described	X				
1.2	*The intervention logic and the theory of change (TOC)*		X	X	X	X
1.2.1	A clear and correct distinction is made between outputs, outcomes and impact		X		X	
1.2.2	The ToC from outputs to outcomes and impacts is clearly elaborated		X		X	
1.2.3	The ToC is logic and realistic		X	X	X	
1.2.4	Critical and crucial links/ingredients in the ToC are identified and can be tested		X	X	X	
1.2.5	The ToC includes concrete measures (input, activity and output level) which guarantee the sustainability of the intervention results					X
1.2.6	The internal risks are clearly identified and explored/estimated		X	X		X
1.2.7	The external assumptions are clearly identified and explored/estimated	X	X	X	X	X
1.2.8	The allocation of the resources foreseen (investments, personnel and operational costs, inputs from other stakeholders) to the outputs is clear			X		

1.3	*The proposed M&E system*	X	X	X	X	X
1.3.1	The most important results envisaged are made operational in an adequate way		X	X	X	X
1.3.2	Where necessary/relevant indicators are disaggregated by sex or other relevant characteristics		X	X	X	X
1.3.3	The proposed M&E system includes a consistent translation of the proposed intervention logic and ToC	X	X	X	X	X
1.3.4	The method for monitoring and evaluating the intervention results and their sustainability is clearly described	X	X	X	X	X
1.3.5	The method to monitor the assumptions is clearly described	X	X	X	X	X
1.3.6	The method to follow up the internal risks is clearly described	X	X	X	X	X
1.3.7	The human and financial resources involved in the M&E system are clearly described			X		
1.3.8	The proposed MIS allows an allocation of expenditures to specific outputs and intervention components			X		
1.3.9	The way in which the M&E system of the intervention is aligned/related to the national/local M&E system is clearly described	X	X	X	X	X
1.4	*Consistency and adaptation of the intervention logic and ToC*	X	X	X	X	X
1.4.1	Changes in the intervention logic and the ToC are clearly reported and justified	X	X	X	X	X
1.4.2	Information is available about the vision and the opinions of the most important stakeholders regarding changes in the intervention logic and ToC	X	X	X	X	X
1.4.3	Changes in the intervention logic and ToC are adequately integrated in the M&E system	X	X	X	X	X

		Relevance	**Effectiveness**	**Efficiency**	**Impact**	**Sustainability**
2.	**Practice regarding intervention implementation, intervention management and context** ***(Information is gathered through documents and field visits)***					
2.1	*Basic information regarding intervention implementation*	X	X	X	X	X
2.1.1	The basic documents (that could reasonably be expected) are available.	X	X	X	X	X
2.1.2	Baseline information (consistent with the intervention logic) relating to the target group (beneficiaries) is available.		X	X	X	X
2.1.3	Baseline information (consistent with the intervention logic) relating to the counterfactual is available.		X	X	X	
2.1.4	Disaggregated baseline information according to gender (or other relevant parameters) is available on relevant indicators.		X	X	X	
2.1.5	There is information (that could reasonably be expected) available regarding to the progress of the implementation of the intervention goals.		X	X	X	

2.1.6	There is information available regarding the participation of the initial target group (beneficiaries).	X	X	X	X	X
2.1.7	There is information available regarding the data collection process for the indicators.	X	X	X	X	X
2.1.8	The proposal regarding data collection allows in principle reliable data collection of the indicators	X	X	X	X	X
2.1.9	The information regarding the monitoring of internal risks is available and potential consequences for the intervention logic and intervention implementation are indicated.		X	X	X	X
2.1.10	The information regarding the monitoring of (external) assumptions is available and potential consequences for the intervention logic and intervention implementation are indicated.	X	X	X	X	X
2.1.11	Intervention expenditures are well documented/ recorded and linked to the outputs			X		
2.2	*The M&E system (in practice)*	X	X	X	X	X
2.2.1	The vison on and the role of monitoring and evaluation (both M&E and independent evaluation) are clear.		X	X	X	X
2.2.2.	The most important stakeholders agreed upon the proposed M&E system	X	X	X	X	X
2.2.3	There are sufficient resources (time, financial resources, human resources) provided to allow for the adequate functioning of the M&E system.	X	X	X	X	X
2.2.4	The responsibilities and procedures regarding the collection and analysis of M&E information are clearly defined.	X	X	X	X	X
2.2.5	The responsibilities and procedures regarding decision-making based on the analysis of M&E information are clearly defined.	X	X	X	X	X
2.2.6	The staff responsible for M&E is competent and independent	X	X	X	X	X
2.2.7	The M&E system of the intervention is aligned/related to the national/local M&E system	X	X	X	X	X
2.2.8	There is an internal drive for strategic management and learning	X	X	X	X	X
2.2.9	The M&E results are being used for learning	X	X	X	X	X
2.2.10	The M&E results are being used for accountability	X	X	X	X	X
2.2.11	Evaluations/studies have been conducted which are of good quality and these provide useful information (relative to what could have been expected)	X	X	X	X	X
2.2.12	The M&E system is regularly reviewed on its quality and adjusted accordingly	X	X	X	X	X

3.	**The evaluation-context** **The reference situation is independent evaluation** **Key actors: North & South partners, governments (Belgium and partner country), target groups** ***(This part is based on information from desk study, interviews and field visits)***	**Relevance**	**Effectiveness**	**Efficiency**	**Impact**	**Sustainability**
3.1	*Attitude of the key players*	X	X	X	X	X

3.1.1	The principal users of the evaluation and their expectations/interests with respect to the evaluation are clearly defined.	X	X	X	X	X
3.1.2	The key actors express demand for (or are at least interested in) the evaluation.	X	X	X	X	X
3.1.3	The expectations of the key actors with respect to the evaluation (process and results) are mutually compatible.	X	X	X	X	X
3.1.4	The expectations of the key stakeholders regarding the evaluation are realistic (in relation to the available resources).	X	X	X	X	X
3.1.5	The principal users were/will be involved in the evaluation process.	X	X	X	X	X
3.1.6	The most important stakeholders (inclusive of the beneficiaries) were involved/will be involved in the evaluation process	X	X	X	X	X
3.1.6	The relationships amongst stakeholders are "healthy"	X	X	X	X	X
3.1.7	It is possible to contact all stakeholders without risking reciprocal influence	X	X	X	X	X
3.1.8	There is a positive attitude towards independent evaluation amongst all stakeholders.	X	X	X	X	X
3.2	*The broader context*	**Relevantie**	**Effectiviteit**	**Efficiëntie**	**Impact**	**Duurzaamheid**
3.2.1	The broader institutional and political context is positive with respect to independent evaluation.	X	X	X	X	X
3.2.2	The socio-cultural context at the level of the beneficiaries allows for the adequate collection of information.	X	X	X	X	X
3.2.3	Local expertise with the required profile for evaluation is available.	X	X	X	X	X

4.	**Suggestions to increase the evaluability and the usefulness of evaluations in the future.**
	The aim is to provide a number of suggestions for each intervention which follow from the analysis conducted and which may increase the evaluability and the usefulness of evaluations in the future. The presented table will form the basis through which the suggestions will be formulated during the feedback moment.
5.	**Feedback from involved actors on the conducted analysis and provided suggestions.**
	Feedback from the involved actors on the conducted analysis and the provided suggestions will be registered and will be incorporated accordingly in the synthesis report.

Annex 2. Schedule field mission and list of interviewees

Participants Briefing 22nd April 2015

Name	Organisation
Astrid de Laminne de Bex	Belgian Embassy
Anne Pierre Mingelbier	BTC
Jef Peeraer	VVOB
Angélique Ekirabo	VVOB
Richard Niwenshuti	SACB
Anna-Maria Schreven	BTC-SACB
Gratien Gasaba	BTC-SACB
Vincent Tihon	BTC-Minisanté
Dr Achour Ait Mohand	BTC-Minisanté
Raf Somers	BTC-SPAT II
Sylvia Salama Gata	SPAT II
Johan Nieuwenhuis	BTC-PAREF II
Jean Damascene	PAREF II
Philippe Daout	Protos
Jotham Mujyalibu	Rwandan Mountain Tea
Félix Byamungu	Caritas
Prosper Sebagenzi	Caritas
Pierre-Claver Ndahayo	Light for the World
Emmanuel Gasana	Independent consultant
Liesbeth Inberg	IOB

Interviews

Belgian Embassy			
Date	**Name**	**Function**	**Organisation**
Wednesday April, 22nd	Astrid de Laminne de Bex	Attaché	Belgian Embassy
BTC			
Wednesday April, 22nd	Anne Pierre Mingelbier	Director of Programs	Belgian Technical Cooperation
VVOB			
Thursday April, 23rd	Jef Peeraer	Programme Advisor/ Manager	VVOB
	Angélique Ekirabo	Programme Advisor Change Management (incl. M&E)	VVOB
	Leon Mugenzi	Programme Advisor Capacity Development	VVOB
	Christian Karasira	Programme Advisor for IR1, ICT and Knowledge Management	VVOB
	Alex Mukizwa Mahe	Programme Advisor for IR 1 and Gender Focal person	VVOB

Friday April, 24th	Alfred Otara	Focal person cooperation VVOB-URCE	University of Rwanda – College of Education (partner VVOB)
	Alphonse Uworwabayeho	Focal person cooperation VVOB-URCE	University of Rwanda – College of Education (partner VVOB)
	Genevieve Ayinkamiye	M&E Officer	Rwanda Education Board (partner VVOB)
	Feedback VVOB		
Minisanté			
Friday April, 24th	Dr. Daniel Ngamije	Director of Intervention	Ministry of Health
Monday April, 27th	Bob Mugisha	Representative in the steering committee	Ministry of Finance
	Dr. Vincent Tihon	Co-manager/ DELCO	BTC
	Dr Achour Ait Mohand	Technical Assistant Mental Health	BTC
	Nancy Misago	Focal point Mental Health	Ministry of Health
	Gervais Baziga	Focal point planning M&E and decentralisation	Ministry of Health
	Didier Nishimwe	Biomedical Technician RBC/MTI	Ministry of Health
SPATII			
Tuesday April, 28th	Raf Somers	Co-manager	BTC
	Joseph Higiro	National FFS coordinator	Rwanda Agriculture Board
	Sylvia Salama Gata	Director of Intervention	Ministry of Agriculture
	Consolata Nakure	M&E specialist	Ministry of Agriculture
	Feedback SPATII		
Thursday April, 30th	Christian Nbazigarire	FFS Master trainer	Rwanda Agriculture Board
SACB			
Wednesday April, 29th	Anna-Maria Schreven	Change Co-manager	BTC
	Gratien Gasaba	Project Co-manager	BTC
	Richard Niwenshuti	Project Coordinator	NCBS
	Judith Kayitesi	Head of SOQA Division	NCBS
PAREF Be II			
Thursday April, 30th	Gad Sibomana	M&E person	RNRA/MINERENA
	Jean-Claude Sebahire	Forest Inventory and Mapping Officer	RNRA
	Paul Hangamimana	Forestry Technical Assistant - Gakenke	PAREF Be II
	Claude Niyigena	Chef de Mission. Forest Service Provider – PAREF Be II	OPDSA Company– Gakenke District
	Jean Claude Nsengiyumva	Forest Service Provider – PAREF Be II	OPDSA Company– Gakenke District
Monday May, 4th	Johan Nieuwenhuis	Programme Co-manager	BTC
	Jean Damascene	Intervention Director	RNRA/MINERENA
	Gad Sibomana	M&E person	RNRA/MINERENA
	Jacque Peeters	International Technician	BTC

	Feedback PAREF II		
Caritas			
Monday May, 4th	Félix Byamungu	Coordinator PASAB	Caritas
	Prosper Sebagenzi	M&E person, programme manager	Caritas
	Abbé Donatien Twizeyumuremyi	Director	Caritas
	Abbé Anaclet Mwumvaneza	SG	Caritas
	Louis Rwagaju	Mayor Bugasera, member of CoPI PASAB	PASAB
	Eprhem Sebarundi	SE Ngeruki, member of CoPI PASB	PASAB
	Berthilde Mukantwari	SE Ntarama, member of CoPI PASAB	PASAB
	Aimable Gasinzingwa	President of Ikozamubu, member of Covapanya cooperation	Covapanya cooperation
	Drocela Umupfasoni	Accountant Nyamata Parish, member of Covapanya cooperation	Covapanya cooperation
	Felix Habincuti	President of Covapanya cooperation	Covapanya cooperation
Tuesday May, 5th	Tharcissie Mucunguyinka	Development Animator	Caritas
	Feedback Caritas		
ARES			
Friday May, 8th	Magnus Kirori	Programme Manager	ARES
	Laetitia Nyinawamwiza	Coordinator	ARES
Sunday May, 10th	Vivien Munyaburanga	Team Manager	ARES
PROTOS			
Wednesday May, 6th	Philippe Daout	Regional Coordinator	Protos
	Carinie Masumbuko	M&E person	Protos
	Joseph Uwizeye	Programme Manager	Protos
	Philbert Hakizimana	Executive Director	Coforwa (partner Protos)
	Claver Hagumintwali	Administrator	Coforwa (partner Protos)
	Callixte Habyarimana	Project engineer	Coforwa (partner Protos)
	Célestin Kanyamugenge	Field technician	Coforwa (partner Protos)
	Joseph Sibomana	Focal point MYP	Coforwa (partner Protos)
	Chantal Umulisa	Field animator	Coforwa (partner Protos)
	Francois Uhagaze	Vice Mayor Economic Affairs	Muhanga District Office
	Védaste	Environment responsible	Muhanga District

	Mpagaritswenimana	person	Office
	Jean Marie Abizeye	Responsible person for development of the Gasagara cell, Rongi sector	Muhanga District Office
	Etienne Nzajyibwami	Responsible person for development of the Rukeri cell, Kiyumba sector	Muhanga District Office
Thursday May, 7th	Jean Damascene Ndahimana	Coordinator	UGAMA (partner Protos)
	Marie Claire Providence Abizeye	Administrator	UGAMA (partner Protos)
	Callixte Twagirayezu	Focal point MYP	UGAMA (partner Protos)
	Malachie Habanabashaka	Field technician	UGAMA (partner Protos)
	Liliose Umurerwa	Field Animator	UGAMA (partner Protos)
	Feedback PROTOS		
BIO			
Saturday May, 9th	Jean Damascene Gasarabwe	Tea Plant Manager Nyabihu Tea Plantation	Rwanda Mountain Tea
	Innocent Ndayishimiye	Hydro Power Plantation Manager – Giciye Hydro Plant	Rwanda Mountain Tea
	Augustine Okea	Factory Engineer	Rwanda Mountain Tea –Nyabihu Factory
	Japhet Kwizera	Internal Auditor	Rwanda Mountain Tea –Nyabihu Factory
Wednesday, May 13th	Jotham Majyalibu	Director General Rwandan Mountain Tea LTD	Rwandan Mountain Tea - Kigali
	Patrick Tuyisenge	Human Resources & Administration Officer	Rwandan Mountain Tea - Kigali
Light for the World			
Monday May, 11th	Pierre-Claver Ndahayo	Service Administrator	Service d'Ophtalmologie de Kabgayi Light for the World
	Jean Muhayimana	Techncien de maintenance de equipments medicaux	Service d'Ophtalmologie de Kabgayi Light for the World
	Patrick Shyaka	Paramedical chief Ophtalmologie service	Service d'Ophtalmologie de Kabgayi Light for the World
	Jean Baptiste Karibuhunde	Statistian/Data Manager	Service d'Ophtalmologie de Kabgayi
	Dr. Piet Noe	Médecin Ophtalmologue Belge & Chef du Service d'Ophtalmologie de Kabgayi	Service d'Ophtalmologie de Kabgayi Light for the World

	Dr. Leon Kagabo Hakizimana	Médicin/Directeur	Nyanza District Hospital
	Dominique Nsabiyaremye	Technician Superier en Ophtalmologie	Nyanza District Hospital
Tuesday May, 12th	Dr Mumena Legis	Head of Ophthalmology department	Kigali Health Institute
	Anita Ahayo	Head of Department	Department Injuries and Disabilities, Ministry of Health
	Feedback Light of the World		

Participant debriefing 13th May

Name	**Organisation**
Astrid de Laminne de Bex	Belgian Embassy
Léa Ingabire	BTC
Jef Peeraer	VVOB
Anna-Maria Schreven	BTC-SACB
Gratien Gasaba	BTC-SACB
Vincent Tihon	BTC-Minisanté
Raf Somers	BTC-SPAT II
Sylvia Salama Gata	SPAT II
Johan Nieuwenhuis	BTC-PAREF II
Philippe Daout	Protos
Joseph Uwizeye	Protos
Félix Byamungu	Caritas
Prosper Sebagenzi	Caritas
Laetitia Nyinawamwiza	ARES
Pierre-Claver Ndahayo	Light for the World
Ivo Hooghe	Evaluation Desk Belgian Ministry of Foreign Affairs
Emmanuel Gasana	Independent consultant
Nathalie Holvoet	IOB
Lisa Popelier	IOB

Annex 3. Ppt of debriefing session

Revue de l'évaluabilité: Session de restitution

South Research
Institut de Politique et de Gestion du Développement (IOB)

13/05/2015

IOB
L'Institut de Politique et de Gestion du Développement
Université d'Anvers

Aperçu

1. Introduction de la revue
2. Résultats préliminaires
   i. Analyse de l'analyse sous-jacente, la logique d'intervention et la théorie du changement
   ii. Analyse du système de S&E proposé
   iii. Analyse de l'évaluabilité pratique
   iv. Analyse des facteurs contextuels qui (co-)déterminent l'évaluabilité
3. Feedback et discussion

1. Introduction – Général

- Revue de l'évaluabilité (pratique) de l'intervention
  ←→ Ce n'est **pas** une évaluation d'intervention

- Sélection de 40 interventions dans 4 countries (Belgique, RDC, Rwanda, Bénin)

- Session de restitution: observations générales basées sur l'ensemble des interventions visités au Rwanda

1. Introduction – Echantillon

Titre	Organisation
SPAT II - Market oriented advisory services and quality seeds	BTC
PAREF II - Projet d'appui à la reforestation dans les provinces de l'Est et du Nord	BTC
MINISANTE IV- Appui Institutionnel au Ministère de la Santé	BTC
SACB - Strategic Approach to Capacity Building	BTC
LOPE - Learning Outcomes in Primary Education (2014-2016)	VVOB
Rwandan Mountain Tea	BIO
Program d'Appui Institutionnelle	ARES (CUD)
PASAB II - Projet d'appui à la sécurité alimentaire au Bugesera	Caritas
Accès à et gestion de l'eau et de l'assainissement au Rwanda	Protos
Rendre la vue au Rwanda: prévention et traitements des problèmes de vues	Lumière pour le monde

1. Introduction – Processus de revue

	Activité	Période
1	Compilation de l'échantillon provisoire	☑
2	Élaboration du cadre de la revue	☑
3	Collection des documents des interventions retenues	☑
4	Etude documentaire	☑
5	Focus group discussion	☑
6	**Étude de terrain**	***En cours***
7	Analyse et synthèse	juin–août
8	Version provisoire du rapport final	août
9	Rapport final	septembre
10	Dissémination des résultats (séminaire)	sept. – oct.

2. Résultats préliminaires

- Analyse des **forces et faiblesses** et **suggestions/recommendations** relatives aux quatre composants:
  1. Analyse de l'analyse sous-jacente, la logique d'intervention et la théorie du changement (1.1, 1.2, 1.4, 1.5)
  2. Analyse du système de S&E proposé (inclusive la clarté sur la position des différentes parties prenantes) (1.3, 1.4)
  3. Analyse de l'évaluabilité pratique (2)
     - Disponibilité d'informations (2.1)
     - Le système de S&E en pratique (2.2)
  4. Analyse des facteurs contextuels qui (co)déterminent l'évaluabilité (3)

2.1 Analyse de l'analyse sous-jacente, la logique d'intervention et la théorie du changement

Forces	Faiblesses
L'analyse sous-jacente est en général la partie la plus développée	
• Le bien fondé de l'intervention et de la situation (problématique) du groupe cible sont clairement décrits • Le lien entre l'analyse sous-jacente et les objectifs de l'intervention est clairement décrit • Le lien entre le bien fondé de l'intervention et la politique sectorielle du pays partenaire est clairement décrit	• L'analyse genre ne fait pas partie intégrante de l'analyse sous-jacente

2.1 Analyse de l'analyse sous-jacente, la logique d'intervention et la théorie du changement

Forces	Faiblesses
La logique de l'intervention et la théorie de changement	
• Il existe une distinction claire et correcte entre les outputs, outcomes et impact • La théorie de changement contient des mesures concrètes à tous niveaux qui doivent assurer la durabilité des bénéfices de l'intervention	• La théorie de changement n'est parfois pas élaborée jusqu'au niveau de l'impact • Des maillons critiques et cruciaux de la chaîne ne sont parfois pas identifiés jusqu'au niveau de l'impact et sont difficile à tester • Pas de distinction claire entre risques et hypothèses • Les changements éventuels dans la logique de l'intervention et la théorie de changement ne sont pas toujours indiqués ni argumentés

2.1 Analyse de l'analyse sous-jacente, la logique d'intervention et la théorie du changement

Suggestions/Recommendations
• Expliciter les maillons/liens entre les differents niveaux des résultats (éventuellement representer la théorie de changement d'une manière plus schématique) • Élaborer la théorie de changement jusqu'au niveau de l'impact • Intégrer une analyse genre dans l'analyse sous-jacente • Identifier et analyser les risques (probabilitités des risques, effets potentiels, action à prendre) (aussi au niveau outcome et impact) • Actualiser le cadre logique et la théorie de changement conformément aux changements réels (certainement dans un contexte qui change rapidement, l'importance de flexibilité)

2.2 Analyse du système de S&E proposé

Forces	Faiblesses
• Les résultats sont bien opérationnalisés (certainement jusqu'au niveau d'outcomes) • Le système S&E proposé contient une opérationnalisation consistante de la logique de l'intervention • L'approche de suivi de la réalisation des résultats de l'intervention est clairement décrite • Les parties prenantes sont souvent d'accord avec le système S&E proposé • Les dépenses sont souvent liées aux outputs (opportunité pour des analyses d'efficience)	• Les indicateurs ne sont parfois pas spécifiés selon le genre ou un autre paramètre pertinent • L'approche de l'évaluation (surtout impact) est parfois moins clairement décrite • Parfois le rôle des groupes cibles dans le système S&E n'est pas clairement décrit et leurs contributions ne sont pas incorporées dans les "ressources" • Il manque souvent un vrai système MIS • Il manque souvent une description de la manière et degré d'articulation du système S&E aux systèmes nationaux/locaux de S&E

2.2 Analyse du système de S&E proposé

Suggestions/Recommendations
• Utiliser l'étude de base/référence pour décider selon quels paramètres ça serait utile de collecter des données désagrégées • Décrire mieux l'approche d'évaluation (analyse) • Décrire mieux le rôles des groupes cibles dans le système S&E (ressource additionnelle et parfois activité additionnelle avec des résultats spécifiques) • Décrire la façon d'articulation au système de S&E national/local

2.3 Analyse de l'évaluabilité pratique – Disponibilité d'informations

Forces	Faiblesses
• Les documents de base sont disponibles • L'information quant au progrès vers la réalisation des objectifs de l'intervention est disponible • L'information quant à la collecte des données liées aux indicateurs est disponible (sauf pour l'impact)	• L'information de base concernant le groupe cible n'est pas toujours disponible et presque jamais de l'information de base concernant le counterfactual • L'information quant à la participation du groupe cible initial n'est pas disponible • L'information quant au suivi de risques (et hypothèses) et les conséquences éventuelles pour la logique de l'intervention et la mise en œuvre du projet n'est souvent pas disponible • Le plan de collecte de données ne permet pas toujours une appréciation de la fiabilité de la collecte de données

2.3 Analyse de l'évaluabilité pratique – Disponibilité d'informations

Suggestions/Recommendations
• Utiliser l'étude de référence pour collecter les valeurs de base des indicateurs définis dans le cadre logique • Collecter l'information quant à la participation du groupe cible initial et comparer le groupe cible effectif avec le groupe cible initial (analyse des drop outs) • Identifier un counterfactual adéquat (non-intervention group, other intervention group, generic controls, more complex types of before-after such as time series, panel etc) • Inclure l'information quant au suivi des risques et hypothèses et les conséquences éventuelles d'une manière plus systématique • Intégrer dans le plan de collecte de données de l'information sur la fréquence, l'échantillon, etc.

2.4 Analyse des facteurs contextuels qui (co)déterminent l'évaluabilité

Forces *(opportunités)*	Faiblesses *(menaces)*
• Les acteurs clés sont partie requérante d'une évaluation et la plupart ont une attitude positive par rapport à l'évaluation indépendante • Les attentes des acteurs clés par rapport à l'évaluation sont réalistes • Les parties concernées les plus importantes sont associées au processus d'évaluation	• Les utilisateurs les plus importants de l'évaluation ainsi que leurs attentes ne sont pas toujours clairement définis • Les attentes des acteurs clés par rapport à l'évaluation pourraient être différentes (efficacité, efficience versus pertinence et impact; evaluation scope)

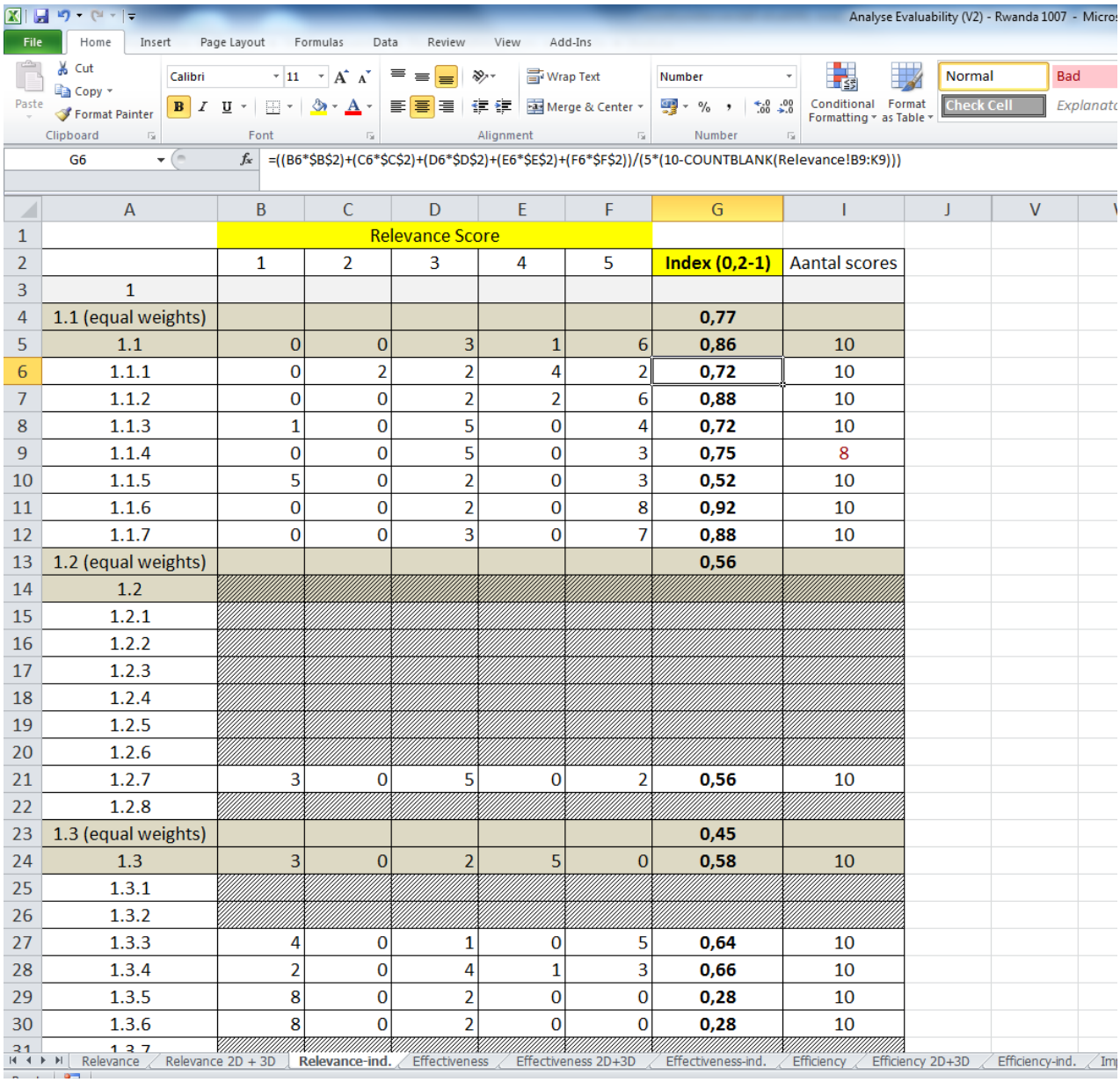
Remarques/Recommendations
• Culture de performance (contexte pays): en principe positif, mais peut aussi avoir des effets négatifs pour l'évaluabilité • Les changements rapides ('reformitis') peuvent rendre difficile le suivi et l'évaluation (flexibilité) • Clarifier les utilisateurs et leurs attentes (atteindre différents objectifs avec une seule évaluation risque de n'atteindre aucun des objectifs)

3. Feedback et discussion

Merci de votre coopération

Au nom de toute l'équipe

Annex 4. Printscreen (index calculation)



G6 =((B6*B2)+(C6*C2)+(D6*D2)+(E6*E2)+(F6*F2))/(5*(10-COUNTBLANK(Relevance!B9:K9)))

	A	B	C	D	E	F	G	I
1		Relevance Score						
2		1	2	3	4	5	**Index (0,2-1)**	Aantal scores
3	1							
4	1.1 (equal weights)						**0,77**	
5	1.1	0	0	3	1	6	**0,86**	10
6	1.1.1	0	2	2	4	2	**0,72**	10
7	1.1.2	0	0	2	2	6	**0,88**	10
8	1.1.3	1	0	5	0	4	**0,72**	10
9	1.1.4	0	0	5	0	3	**0,75**	8
10	1.1.5	5	0	2	0	3	**0,52**	10
11	1.1.6	0	0	2	0	8	**0,92**	10
12	1.1.7	0	0	3	0	7	**0,88**	10
13	1.2 (equal weights)						**0,56**	
14	1.2							
15	1.2.1							
16	1.2.2							
17	1.2.3							
18	1.2.4							
19	1.2.5							
20	1.2.6							
21	1.2.7	3	0	5	0	2	**0,56**	10
22	1.2.8							
23	1.3 (equal weights)						**0,45**	
24	1.3	3	0	2	5	0	**0,58**	10
25	1.3.1							
26	1.3.2							
27	1.3.3	4	0	1	0	5	**0,64**	10
28	1.3.4	2	0	4	1	3	**0,66**	10
29	1.3.5	8	0	2	0	0	**0,28**	10
30	1.3.6	8	0	2	0	0	**0,28**	10