



IOB
Institute of Development Policy and Management
University of Antwerp

23/11/2018

Evaluability of Belgian Development Cooperation interventions: Insights from a multi-country evaluability study of 40 interventions

Nathalie Holvoet (IOB)

Based on Holvoet, N., Inberg, L., Peeters, B., Popelier, L., Van Esbroeck, D. and Verhofstadt, E. (2016) *To evaluate in a credible and meaningful way: between dream and reality. A study of the evaluability of (co)-financed interventions of the Belgian Cooperaton*. Brussels: Federal Public Service Foreign Affairs, Foreign Trade and Development Cooperation. The Special Evaluation Office of the Belgian Development Cooperation (available at http://diplomatie.belgium.be/en/policy/development_cooperation/how_we_work/special_evaluation_office)

OUTLINE

1. BACKGROUND
2. EVALUABILITY AND USEFULNESS
3. METHODOLOGY
4. SELECTED FINDINGS AND ANALYSIS

1. BACKGROUND

- study Belgian Office of Special Evaluator
- ↑importance evaluation ↔ quality, usability and effective use
 - ✓ DC: difficult settings, multiple stakeholders with diverging/competing interests, ...
- aim : “contribute to making all future actions evaluable in the not too long term” (initial), through the study process rephrased, broadened

2. 'EVALUABILITY' & USEFULNESS (1)

- no standardised definition

OECD/DAC definition (2002, p. 21)

“Evaluability is the extent to which an activity or a programme can be evaluated in a reliable and credible fashion”

- evaluability as a continuum
- evaluability ≠ inherent quality
- differ according to the different OECD/DAC criteria
- theoretical evaluability ≠ practical evaluability
- evaluability ≠ evaluability assessment

2. 'EVALUABILITY' & USEFULNESS (2)

- usefulness evaluability assessment
 - ↑ value added of evaluation
 - ↑ M&E system quality (~ formative evaluation)
 - depends on intervention phase
 - ✓ ex-ante: ↑ quality of intervention design
 - ✓ start implementation: input into set-up M&E system
 - ✓ during implementation & ex-post: assess whether evaluation is appropriate & desirable & redesign evaluation

3. METHODOLOGY (1)

Study framework (based on Davies, 2013)

- o three dimensions, subdivided into 9 components
- o differentiation according to 5 evaluation criteria

DAC Evaluation Criteria	Relevance	Effectiveness	Efficiency	Impact	Sustainability
Dimensions/Components	Number of items scored				
<i>1. Analysis of the intervention design</i>					
1.1 Underlying analysis (7)	(7)	(4)	(0)	(0)	(0)
1.2 The intervention logic and theory of change (8)	(1)	(6)	(5)	(5)	(3)
1.3 The proposed M&E system (9)	(5)	(7)	(9)	(7)	(7)
1.4 The consistency and adaptation of the intervention logic and theory of change (3)	(3)	(3)	(3)	(3)	(3)
<i>2. Practice with respect to implementation and management of the intervention and the context</i>					
2.1 Basic information concerning the intervention implementation (11)	(5)	(10)	(11)	(10)	(8)
2.2 The M&E system in practice (12)	(11)	(12)	(12)	(12)	(12)
<i>3. The evaluation context</i>					
3.1 Attitude of the key stakeholders(9)	(9)	(9)	(9)	(9)	(9)
3.2 The broader context (3)	(3)	(3)	(3)	(3)	(3)
3.3 Practical elements (2) (°)					
Aggregate score	(44)	(54)	(52)	(49)	(45)

→ 62 items

- o scoring based on 5 point-scale
 - ✓ at item and component level
 - ✓ detailed grid
- o calculation of indexes
 - ✓ ‘relative’ interpretation
- o comparative analysis
 - ✓ countries, **complex/less complex ToC**, types of actors

2.1 Methodology - Study framework

	Relevance	Effectiveness	Efficiency	Impact	Sustainability	
1.1 Underlying analysis	x	x				<p>1 = Analysis has not been performed/is lacking.</p> <p>2 = Analysis is vague and lacks crucial information regarding the situation (context, needs, problems) of the beneficiaries.</p> <p>3 = Analysis is clearly described but not comprehensive enough.</p> <p>4 = Analysis is clearly described and comprehensive but it is inadequately linked to the project objectives.</p> <p>5 = Analysis is clearly described and comprehensive and it supports the project objectives (neat matching between project objectives and analysis).</p>



2.1 Methodology – Study framework

1	Effectiveness
2	Interventions
3	Country
4	Actor (channel)
5	Sector

	A	B	C	D	E	F	G	H	I	J	K	L
1	Effectiveness											
2	Interventions	RWA1	RWA2	RWA3	RWA4	RWA5	RWA6	RWA7	RWA8	RWA9	RWA10	
6	1											
7	1.1 (equal weights)	3,50	3,25	2,75	4,33	3,33	2,50	2,25	4,00	4,75	3,25	
8	1.1	5	3	5	5	5	3	3	5	5	4	
9	1.1.1	3	4	2	5	4	3	2	5	4	4	
10	1.1.2											
11	1.1.3	1	3	5	5	5	3	3	3	5	3	
12	1.1.4	5	5	3			3	3	3	5	3	
13	1.1.5	5	1	1	3	1	1	1	5	5	3	
14	1.1.6											
15	1.1.7											
16	1.2 (equal weights)	3,00	3,33	3,00	3,67	3,67	1,00	3,67	4,33	4,67	3,33	
17	1.2	2	3	2	3	3	1	2	4	4	4	
18	1.2.1	5	5	5	5	5	1	3	5	5	3	
19	1.2.2	3	3	3	3	3	1	3	3	5	5	
20	1.2.3	3	3	3	5	5	1	5	5	5	3	
21	1.2.4	1	1	1	1	1	1	5	3	3	3	
22	1.2.5											
23	1.2.6	5	5	5	5	5	1	3	5	5	3	
24	1.2.7	1	3	1	3	3	1	3	5	5	3	
25	1.2.8											
26	1.3 (equal weights)	3,14	1,57	2,33	3,50	3,00	1,00	2,00	3,43	3,71	3,29	
27	1.3	4	3	3	4	4	1	2	4	4	4	
28	1.3.1	5	3	3	4	3	1	3	5	5	4	
29	1.3.2	5	1			3		1	2	4	3	
30	1.3.3	5	1	3	5	5	1	3	5	5	5	
31	1.3.4	4	3	3	5	3	1	4	5	5	4	
32	1.3.5	1	1	1	1	1	1	1	3	3	1	
33	1.3.6	1	1	1	1	1	1	1	3	3	1	
34	1.3.7											
35	1.3.8											

Sheet for data entry per DAC criterion

2.1 Methodology – Study framework

Sheet for calculation of index per DAC criterion

Formula bar: $=((B24*\$B\$2)+(C24*\$C\$2)+(D24*\$D\$2)+(E24*\$E\$2)+(F24*\$F\$2))/(5*(10-COUNTBLANK(Effectiveness!B27:K27)))$

	A	B	C	D	E	F	G	J	K	L	M	N	O
1	Effectiveness scores												
2		1	2	3	4	5	Index (0-1)						
3	1												
4	1.1 (equal weights)						0,7						
5	1.1	0	0	3	1	6	0,9						
6	1.1.1	0	2	2	4	2	0,7						
7	1.1.2												
8	1.1.3	1	0	5	0	4	0,7						
9	1.1.4	0	0	5	0	3	0,8						
10	1.1.5	5	0	2	0	3	0,5						
11	1.1.6												
12	1.1.7												
13	1.2 (equal weights)						0,7						
14	1.2	1	3	3	3	0	0,6						
15	1.2.1	1	0	2	0	7	0,8						
16	1.2.2	1	0	7	0	2	0,6						
17	1.2.3	1	0	4	0	5	0,8						
18	1.2.4	6	0	3	0	1	0,4						
19	1.2.5												
20	1.2.6	1	0	2	0	7	0,8						
21	1.2.7	3	0	5	0	2	0,6						
22	1.2.8												
23	1.3 (equal weights)						0,5						
24	1.3	1	1	2	6	0	0,7						
25	1.3.1	1	0	4	2	3	0,7						
26	1.3.2	2	1	2	1	1	0,5						
27	1.3.3	2	0	2	0	6	0,8						
28	1.3.4	1	0	3	3	3	0,7						
29	1.3.5	8	0	2	0	0	0,3						
30	1.3.6	8	0	2	0	0	0,3						
31	1.3.7												
32	1.3.8												
33	1.3.9	6	0	1	0	3	0,5						
34	1.4 (equal weights)						0,4						
35	1.4	7	1	0	2	0	0,3						
36	1.4.1	9	0	1	0	0	0,2						
37	1.4.2	10	0	0	0	0	0,2						
38	1.4.3	10	0	0	0	0	0,2						

Navigation tabs: Relevance, Relevance-ind., Effectiveness, Effectiveness-ind., Efficiency, Efficiency-ind., Impact, Impact-ind., Sustainability, Sustainability-ind.

2.1 Methodology – Study framework

Summary of indexes per DAC criterion for each component of the study framework

	A	B	C	D	E	F
		Relevance	Effectiveness	Efficiency	Impact	Sustainability
1						
2	1					
3	1.1 (equal weights)	0,77	0,68			
4	1.1	0,86	0,86			
5	1.1.1	0,72	0,72			
6	1.1.2	0,88				
7	1.1.3	0,72	0,72			
8	1.1.4	0,75	0,75			
9	1.1.5	0,52	0,52			
10	1.1.6	0,92				
11	1.1.7	0,88				
12	1.2 (equal weights)	0,56	0,67	0,71	0,40	0,76
13	1.2	0,44	0,56	0,70	0,38	0,70
14	1.2.1		0,84		0,72	
15	1.2.2		0,64		0,36	
16	1.2.3		0,76	0,84	0,56	
17	1.2.4		0,40	0,48	0,24	
18	1.2.5					0,80
19	1.2.6		0,84	0,88	0,28	0,84
20	1.2.7	0,56	0,56	0,68	0,32	0,64
21	1.2.8			0,68		
22	1.3 (equal weights)	0,44	0,54	0,58	0,38	0,53
23	1.3	0,58	0,66	0,66	0,42	0,66
24	1.3.1		0,72	0,78	0,44	0,76
25	1.3.2		0,54	0,53	0,31	0,47
26	1.3.3	0,64	0,76	0,80	0,60	0,76
27	1.3.4	0,60	0,74	0,72	0,44	0,72
28	1.3.5	0,28	0,28	0,32	0,24	0,28
29	1.3.6	0,28	0,28	0,32	0,24	0,28
30	1.3.7			0,58		
31	1.3.8			0,68		
32	1.3.9	0,40	0,48	0,48	0,40	0,44
33	1.4 (equal weights)	0,35	0,35	0,36	0,35	0,35
34	1.4	0,34	0,34	0,34	0,34	0,34
35	1.4.1	0,24	0,24	0,24	0,24	0,24



3. METHODOLOGY (2)

- **sample**

- 40 interventions co-financed by Belgian DC
- 4 countries (Benin, DRC, Rwanda, Belgium)
- different intervention channels
 - ✓ 25% bilateral (country to country)
 - ✓ 50% indirect (NGOs, trade unions)
 - ✓ 25% other stakeholders (universities, BIO, VVOB/APEFE, ...)
- complex (60%) versus less complex (40%) theory of change
 - ✓ initially sectoral division but did not hold, more refined classification based on document review

3. METHODOLOGY (3)

- **data collection and analysis**

- secondary & primary data collection (4 days/intervention)
- field study
 - ✓ focus group discussions
 - ✓ interviews key stakeholders (20 to 25 persons)
 - ✓ feedback at intervention level
- scoring by different team members + calibration

4. SELECTED FINDINGS AND ANALYSIS (1)

Item	Relevance	Effectiveness	Efficiency	Impact	Sustainability	Evaluability index
Dimension 1 (intervention plan) (°)	3.11	3.22	3.29	2.52	3.15	3.16
1.1 The underlying analysis (7)	3.65	3.60				3.63
1.2 The intervention logic and theory of change (8)		3.00	3.50	2.10	3.25	2.96
1.3 The proposed M&E system (9)	2.45	2.83	2.85	1.95	2.46	2.51
1.4 Consistency and adaptation of the intervention logic and theory of change	3.55	3.44	3.52	3.50	3.75	3.54
Dimension 2 (implementation practice)	2.99	3.25	3.51	2.33	2.73	2.96
2.1 Basic information as regards the intervention implementation (11)	2.95	3.20	3.53	2.13	2.62	2.88
2.2 M&E system in practice (12)	3.03	3.30	3.50	2.53	2.85	3.04
Dimension 3 (context)	3.89	3.98	3.98	3.84	3.89	3.91
3.1 Attitude of the key stakeholders (9)	3.68	3.70	3.70	3.58	3.63	3.66
3.2 Broader context (3)	4.10	4.25	4.25	4.10	4.15	4.17
General score for evaluability (°°)	3.26	3.38	3.52	2.70	3.13	3.23

Index 4,01 – 5,00

Index 3,01 – 4,00

Index 2,01 – 3,00

Index 1,01 – 2,00

4. SELECTED FINDINGS AND ANALYSIS (2)

- general findings

- maximum score around 3.5, many in yellow zone → room for improvement
- dimensions: context > intervention design > implementation practice
- components:
 - ✓ strongest: in context dimension (but needs to be nuanced!)
 - ✓ weakest: proposed M&E system + basic information regarding intervention implementation → room for improvement in areas most under control
- evaluation criteria:
 - ✓ strongest: effectiveness & efficiency ← good M&E practice at implementation/operational level (incentives)
 - ✓ weakest : impact & sustainability ← methodological challenges, limited attention (incentives)

4. SELECTED FINDINGS AND ANALYSIS (4)

DIMENSION 1: INTERVENTION DESIGN

	STRONGER	WEAKER
Underlying analysis (7)	-generally ok ← continuity in interventions	-gender analysis & social differentiation → implications for later phases
Intervention logic and theory of change (8)	-implementation theory → ↑ effectiveness & efficiency evaluability	-impact theory → ↓ impact evaluability → difficult to correct for later ← lack of incentives (internal & external)
Proposed M&E system (9)		-weakest component in general ← lack of incentives ← similar weaknesses as in analysis → implications for M&E in practice, but can sometimes be corrected for later
Changes in intervention logic & ToC (4)	-highlighted in documents	-not transferred to M&E system

4. SELECTED FINDINGS AND ANALYSIS (5)

DIMENSION 2: IMPLEMENTATION PRACTICE		
	STRONGER	WEAKER
Basic information related to progress of intervention implementation (11)	-implementation level (inward looking focus) → ↑ effectiveness, efficiency evaluability	-gender & social differentiation -higher levels in ToC -external hypotheses -counterfactual -data collection process
M&E system in practice (12)	-M&E practice > M&E on paper -implementation level → reflection & learning at organisational level	-higher levels in ToC -monitoring > evaluation (time, staff, institutional arrangements) -inward looking .limited role of stakeholders .articulation with local & national M&E systems .focus on donor accountability

4. SELECTED FINDINGS AND ANALYSIS (6)

DIMENSION 3: THE EVALUATION CONTEXT

	STRONGER	WEAKER
Attitude of key stakeholders (9)	-key stakeholders positive & open	-limited number of stakeholders involved -interests and expectations of stakeholders not included → inward looking
Broader context (3)	-few restrictions (at first sight), challenges easily solved ↔ ○ technical level > political level ← inward looking ○ social desirable answering ? (our study similarly affected as evaluations)	

Table 13: Overview of the evaluability scores for interventions with a complex and less complex TOC

Item	Less complex interventions	Complex interventions	Evaluability index (°)
Dimension 1 (intervention plan)	2,87	3,30	3,16
<i>1.1 The underlying analysis</i>	3,31	3,83	3,65
<i>1.2 The intervention logic and theory of change</i>	3,03	2,92	2,96
<i>1.3 The proposed M&E system</i>	2,11	2,78	2,51
<i>1.4 Consistency and adaptation of the intervention logic and theory of change</i>	3,03	3,67	3,54
Dimension 2 (implementation practice)	2,82	3,07	2,96
<i>2.1 Basic information as regards the intervention implementation</i>	2,68	3,03	2,88
<i>2.2 M&E system in practice</i>	2,95	3,10	3,04
Dimension 3 (context)	3,82	3,98	3,91
<i>3.1 Attitude of the key stakeholders</i>	3,44	3,80	3,66
<i>3.2 Broader context</i>	4,20	4,15	4,17
General score for evaluability (°°)	3,04	3,34	3,23

(°)The score for the general evaluability index is calculated on the basis of indexes (not averages). The scores calculated for complex and less complex interventions are *averages* (not indexes). Consequently, it is possible that the general evaluability score deviates slightly from the weighted average of the scores of less complex (40%) and complex (60%) interventions. The differences are very tiny and do not change the contents of this analysis.

(°°) For the general evaluability score a weighted average was calculated in which dimensions 1 and 2 are weighed equally and dimension 3 counts for half of each of the other dimensions

4. SELECTED FINDINGS AND ANALYSIS (7)

- no significant differences of evaluability scores between intervention with less or more complex ToC
 - slightly higher for more complex ToC ← ? more investment of staff in proving results and value of interventions

5. SELECTED RECOMMENDATIONS

- integration of evaluability assessment by Belgian actors in different phases of interventions
 - usefulness broader than evaluations as such
- invest in proper intervention design → ↑evaluability directly and indirectly
 - e.g. gender/social differentiation in analysis
 - impact level in ToC
- increased attention for impact (theory) level
 - external (donor) incentives needed
 - establishment of pooled fund for impact evaluations
 - ↓ inward looking focus (more involvement of stakeholders, link to local M&E systems)



Thank you

Nathalie Holvoet (nathalie.holvoet@uantwerpen.be)



IOB

Institute of Development Policy and Management
University of Antwerp