Evaluating and Comparing Development Project Effectiveness

A concept to standardize and quantify results

by Jan Stiefel, IDEAS AidRating, aidrating@ideas-expert.ch IDEAS AidRating, P.O. Box 1992, CH-8401 Winterthur, Switzerland. www.aidrating.org

Introduction

A key problem in evaluation of development work is the difficulty to compare their results. The purpose of this paper is to demonstrate the extent to which new dimensions of analysis and learning can be achieved once this is solved by using a comprehensive standard for evaluation. Examples presented may show how far this can bring us if more or full transparency on development cooperation projects and their results becomes a reality, as would be the case if a standard like IATI gains widespread acceptance.

The IDEAS work group has developed a tool which is designed to appraise any kind of development cooperation project, and groups of projects, in a standardized way. This is done by

a) applying the same complete set of criteria (issues) in all projects: social/ intercultural/ livelihood/ environment/ management and assessing dimensions of each.

b) expressing the results both in a numeric rating and a descriptive part. The former of the two is essential to gain an impact profile which can be compared with other projects, the latter allows to include qualitative considerations and aspects not otherwise covered.

c) composing single ratings in a way appropriate to best express programs or complex projects with a number of differing objectives, or portfolios of entire organisations no matter what the size or number of projects is, and connecting them through a compiled rating.

Rating of single projects

The rating is proposed as a tool to standardize the way any intervention (project) is looked at. For each project, the rating produces an impact profile on all criteria mentioned under a) above. Profiles of different projects can be directly compared. Table 1 is an example of a condensed single project rating, with the basic rating result at bottom of the page.

Comparison of projects and cumulative ratings for entire portfolios

The standardization brings about an additional advantage not possible so far: The cumulative assessment of impact over an entire portfolio of interventions, e.g. for an entire agency, or for a subject matter, or for a geographic region, no matter how many interventions are to be summarized or how big or small they are. This is expected to be a major contribution to the enhancement of planning and performance of development cooperation. For this, graphs 1 to 4 provide examples.

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
www.aidrating.org

Keynote

The AidRating concept is designed to overcome problems commonly encountered with conventional approaches of "evaluation" in a wide sense: focus too limited or specific, differing formats, varying depths of enquiry, etc. and usually not comparable with other evaluations.

This is overcome by assimilating and including all issues linked to development projects and their setting in its widest sense[1]. At the same time, all issues are distilled into standardized but comprehensive groups (e.g. environment, livelihood, social issues)[2] and into a set of generally applicable parameters of dimensions and quantities.

To be generally usable, the concept implements a set of hypotheses which we deem to be "generally accepted" in the development community[3]:

> (1) "If a development activity impacts the single element "x" in a favourable way, then a positive development result can be assumed for this element. If the element is impacted in an unfavourable way, then we assume a negative development result for it."

If this can be accepted, we can take this assumption one step further and say:

> (2) "If a development activity impacts a specific set of sociocultural, environmental, and economic criteria in a favourable way, then a positive and sustainable development result can be expected from that activity, and vice versa."

From that reasoning comes the essential assumption that

> **(3) "Favourable and unfavourable impacts cannot be viewed each for itself alone. They always stand in relation to each other and to the whole[4]. Therefore, any intervention implies a ropewalk between desirable and undesirable effects. The outcome is always result of a tradeoff."**

Purpose, scope and modes of use

The Impact rating concept is designed to provide a universally applicable evaluation tool providing a decision basis for any kind of project intervention. It provides a profile for each intervention which allows decision makers, project leaders, and other stakeholders to identify strengths and weaknesses and take informed decisions.
The concept applies standards of desirability or not on which broadest acceptance can be expected.

---

[1] all actions called "project" designed to generate desired changes in a given sociocultural, environmental and economic setting, commonly termed "development setting".

[2] We call these issue groups „criteria": environmental, social, cultural, livelihood, management. See also table 1.

[3] Any rating needs a departure point. The departure point of this rating is that there are generally accepted principles of what "good development" is, and what not. This can be viewed, for example, expressed in the Millennium Development Goals and the declarations of the Rome, Paris, and Accra Summits.

[4] In other words: It is highly unlikely that any project (in fact any human activity) has only favourable or only unfavourable impact in all categories.

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
www.aidrating.org

2

The novel quality about the IDEAS AidRating concept is three-fold:

1. for the first time all relevant aspects are viewed together, one beside the other and in a common overview, so as to obtain an overall insight into all possible effects or impacts caused by a given development activity.

2. the concept is not focused on selected project activities with the question whether they were "successful" or not, and why, as is usually done. The AidRating concept assumes that any activity always has impact on all criteria. The difference lies in the extent of this impact (it can be positive, neutral, or negative), and in the objectives and priorities of the intervention.

3. The "why" of success or failure is replaced by the assessment "how" this comes about. The way activities meet or miss their targets is expressed in several ways in this framework, numerically and also in comments. Side effects as part of the picture are also considered. This allows to make links visible between causes and effects, see their dimensions, and get directions as to what may need to be changed.

The rating will be most reliable when data cover all criteria and are complete. However, incomplete data do not mean that we cannot make an interpretation. Missing data or information can be provisionally substituted out of experience in similar cases. For such cases each detail rating is tagged which allows to verify its reliability at any time.
Essentially necessary are:

1. exact area or location of intervention, and its size ("geodata"[5])

2. population in the area, population size, and its social and ethnic background[6]

3. Details on project activities on the ground (hardware, travel, …)

4. Intended objectives

5. Information on organisational set-up including staff, and investments. These are mainly used to assess soundness of the managerial approach during project implementation. Later (after project completion), they are likely to lose their relevance.

Annex 1 holds a more detailed table explaining necessary project information. The better founded the information is, the more reliable will the resulting profile become. It is helpful if projects can be grouped along or close to classification systems used by large agencies, such as OECD (DAC CRS purpose codes).

---

[5] geodata are quite important, as they also allow to estimate related conditions: climate, hydrology, geology, etc.

[6] like geodata for ecological implications, and perhaps closely linked to these, information on the local population can allow to estimate cultural implications, positive as well as inhibitive: Taboos, local knowledge, solidarity systems, etc.

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
www.aidrating.org

3

Rating procedure and criteria

The single project rating may be the most obvious one and most frequent one to be applied. This section is intended to demonstrate the outcome of the rating procedure for a single project, and, combined with others, for a sample portfolio with a number of projects.

The detailed procedure includes a comprehensive set of steps from data collection to write-up and rating which is described in a separate paper. For more details, please contact jan.stiefel@ideas-expert.ch.
Only the essentials to understand the reasoning and procedure are given here. Annexes 1 to 3 summarize how we handle project information, data reliability, and dimensions. Some detail can be derived from the "SODIS" sample rating sheet given in table 1, and from the sample ratings of the projects presented.

Single Project Rating

The basic rating object is the single project. All projects are assessed along the key criteria Social impact, Intercultural Impact, Livelihood, Environment, and sound management. These criteria are subdivided in detail issues (4 or 5 each), which are then analyzed, rated on a scale between -2 (very unfavourable) to +2 (very favourable) and commented.
All issues and criteria are also assessed and rated for their dimension (number, size) and for reliability of information on them. Table 1 is an example for a test rating sheet showing the set-up and resulting graphs. The rating has been conducted on a section of the international SODIS drinking water project http://www.sodis.ch/index_EN and is based on available literature. SODIS works with solar disinfection of water in used PET drinking bottles.

The sheet in table 1 is a version conceived to show all key informations at a glance:

   a) at top: Name, general information and general statistical information

   b) in center section: Main criteria, their underlying issues[7] and their rating

   c) bottom left: a graphic representation of reliability of information used

   d) Bottom: The graph with the basic rating (but not corrected for dimensions), and a brief summary.

Overview of results

The principal profile is shown in the graph at the bottom of the page: It shows a fairly favourable rating in most criteria in this example. This would be the desirable case in project work. The environmental component resulted less favourable than clients expected. There were doubts about chemical softeners of the plastic in the bottles, and about the waste PET bottles still produce at the end of their useful life. Cultural issues could not be measured as no information was found on this aspect.

As can be seen in the round transparency rating graph, available information was scarce.

Size and format of tables and graph are examples for the purpose of making our point.

---

[7] Thus, the environmental criterium looks at the following issues: Natural resources (and their consumption), biodiversity, pollution, habitat quality, CO2 balance. These in turn may be built from differing elements, depending on location and situation.

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
www.aidrating.org

4

Table 1: Example of a rating sheet: SODIS Latin America

Mar 2008	Date scored			a	**Project Scorecard**		
2	Step				Code:	**14030**	210.00	67
1	Depth	status:	**0**		Risk category:		

			Program/Project values		
Project Name:	**SODIS**		a	Client group size:	**4**		c
Country & Location:	**Various**	**Lat Am**	a	Client group coverage:	**1**		c
Lead implementer:	**EAWAG/SANDEC**		a	Sociocult. Sensitivity:	**1**		c
lead counterpart:	**SANDEC various**		a	Area size:	**5**		c
Lead Donor(s):	**various**		e	Area coverage:	**1**		c
contributors:			e	Ecological Sensitivity:	**1**		c
r funding € since start:			e	Human scope:	**6**	
t funding € since start:			e	Geographic scope:	**7**	
Field start date:	**1999**		a			
latest phase dates/ended:		e	Expected effect start:		Effect to last (y):		

Program/Project Specific Criteria:		Element Rating	info quality/	potential	percen-tages
Social impact:	on outside representation/equitability of stakeholder groups:	0	e		
	on relation to external/neighboring communities:	0	e		
	on in-group equitability (gender, age gr. etc):	1	e		
	on group self-determination:	0	e		
	Average:	0.25			12.50
Cultural adaptability:	knowledge of and adaptation to cultural context:	0	e		
	involvement/inclusion of local knowledge:	0	e		
	impact on cultural autonomy and diversity:	0	e		
	on cultural resilience:	0	e		
	Average:	0.00			0.00
Livelihood/ Economic impact:	general income opportunities and livelihood improvement:	1	e		
	equitability of income opportunities and livelihood improvement:	0	e		
	Emergency resilience:	1	a		
	Long term resilience:	0	e		
	stability and viability in view of phase-out:	0	e		
x	Average:	0.40			20.00
Environmental impact:	Impact on critical natural resources consumption:	0	c		
	Impact on biodiversity:	0	c		
	Pollution/contamination:	0	c		
	Habitat quality:	0	c		
	Greenhouse/ CO2 balance:	1	c		
	reserve:	0			
	Average:	0.17			8.33

Project management:				
Principal goals unequivocal (P):	1	c		
Link principal goals & means verified (P):	1	c		
lient consultation and empowerment (P/I):	0	e		
apacity building local partners/clients (P/I):	1	c		
functional project structures (P/I):	0	e		
tive learning and flexibility by project (P/I):	0	e		
Cost-effect ratio:	0	e		
Average:	0.43			21.43
Percent value per element:	3.85	U		
Basic score overall				12.45

Comments: idea with hygiene benefit at low cost
technical approach well described and founded
technology transfer may need specific answers to respond to
social and cultural restrictions. Availability bottles in long run?

100% unfavourable — 100% favourable

Basic score overall; 12.5
Project management:; 21.4
Environmental impact:; 8.3
Livelihood/ Economic impact:; 20.0
Cultural adaptability:; 0.0
Social impact:; 12.5

* = Principal objectives as per documentation

20.09.2011 23:02

Combined Rating of a portfolio of projects (Compound rating)

An entire portfolio of projects, no matter how many, can be rated once all its projects or a representative portion of them are documented and rated.

The compound rating composes the individual ratings of each project, and takes the respective size of each (population and area) into account. This allows to compile the project results while maintaining the proper proportion of each.

In this section, we would like to show what becomes possible when a portfolio of projects is rated this way.

For demonstration purpose, a portfolio of projects including the mentioned SODIS project, plus 10 others has been put together. This portfolio represents a realistic mix in a country, a region, or with a donor. Location is a hypothetical region in, say, Latin America. All projects are evaluated while they are still running. The projects with their data are listed in annex 4.

In brief, the ten are:

1. Three medium sized smallholder farmer projects: One with good adaptation to local needs and problems (A), one with medium adaptation (B), and one poorly adapted (C).

2. Three rural livestock improvement projects: One with good adaptation to local needs and problems (A), one with medium adaptation (B), and one poorly adapted (C)

3. A medium sized (CHF 500'000 p.a.) rural alphabetisation project with good track record

4. A medium sized rural health centre (CHF 300'000 p.a.)

5. A technology oriented livestock project (1'500'000 CHF p.a.), poorly adapted to local needs

6. A technology-driven project for production and promotion of fuel crops (CHF 2'000'000 p.a.), poorly adapted to local communities

The compound rating sums up the single project ratings corrected by their respective dimension, and adds up their budgets. For visualization, logarithmic scaling allows to view very large projects alongside very small ones. **Proportions are consistently maintained and expressed!**

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
www.aidrating.org

6

Portfolio Impact

In the following cumulative graph, The impact and total cost of all projects together is shown in a logarithmic graph[8].

Graph 1: Cumulative effects and cost of 11 hypothetic rural development projects

Overview of Cumulative Impact of Entire Portfolio

Values are logarithmic
totals of all projects in portfolio

negative	positive

Category	Value
cost since start	7.573451822
average	-4.026328939
Environmental impact	4.01494035
Livelihood impact	4.390051496
Cultural impact	-4.687528961
Social impact	-4.457881897

The graph shows that some positive environmental and livelihood improvement is achieved, while cultural and social impact of the portfolio is unfavourable. Given equal importance to all criteria, the unfavourable effects outweigh Total cost was CHF 37.45 Million *(inv log 7.573451822)*

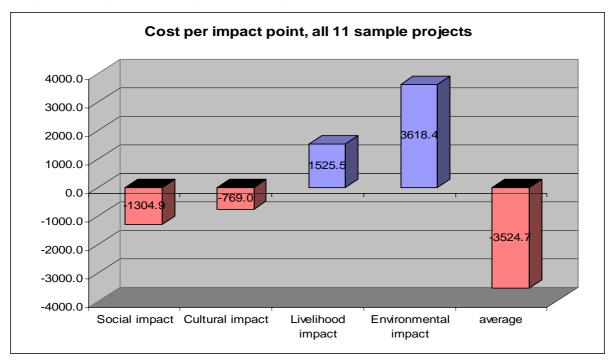[8] The numeric values can be deducted by de-logarithmizing the values given.

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
www.aidrating.org

7

Cost effectiveness of portfolio

The concept also allows to assess the cost effectiveness of a portfolio as follows:

Graph 2: Cost per "impact point" of the portfolio

Cost per impact point, all 11 sample projects

The standardization allows to track impact versus cost relations to their source because all are known: The dimensions of impact, and total cost since inception.
If we take into account that

> (4) one impact point (human) is one point on the scale for one standard family of 5 persons (social, cultural, livelihood impact)
>
> and
> one impact point (geographic) is one point on the scale for one sq km (km^2) of land in the project area (environmental impact)

then we find on the positive side that:

- The cost of the whole program to improve livelihood for one family by one point was CHF 1525.50

- The cost of the whole program to improve the ecological condition of one km^2 land by one point was CHF 3618.40

This means that the program was more cost efficient to achieve livelihood than in achieving environmental improvement.

If the concept finds acceptance, planners and researchers might find such comparison tools attractive. They may further like to assess whether these absolute values are "high" or "low" in comparison to other portfolios. Such things would become possible once the database and number of rated projects and portfolios would be large enough to compare projects and programs in terms of their outcome, no matter what issue they are dealing with.

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
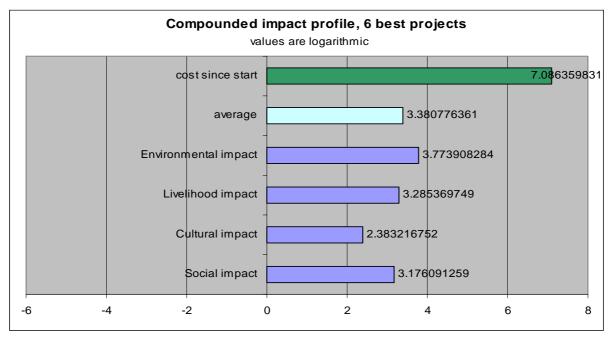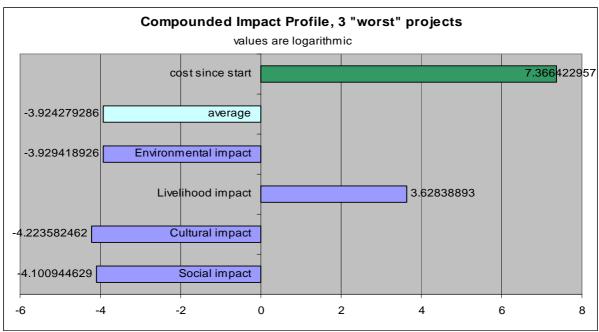www.aidrating.org

8

The graph shows also how easy it was to generate undesired effects: This is shown by the negative (red) bars: The shorter these are, the more undesired effects the portfolio as a whole has produced.

Comparing parts of portfolios (sub-portfolios)

Planners and evaluators may want to find out which part of a portfolio will produce the best results at the least cost, and which part might better be discontinued. The following example is to show how this can be done through a comparison between two "sub-portfolios":

Graphs 3 and 4: Profiles of „best" and „worst" group

Compounded impact profile, 6 best projects
values are logarithmic

Category	Value
cost since start	7.086359831
average	3.380776361
Environmental impact	3.773908284
Livelihood impact	3.285369749
Cultural impact	2.383216752
Social impact	3.176091259

(x-axis: -6, -4, -2, 0, 2, 4, 6, 8)

Compounded Impact Profile, 3 "worst" projects
values are logarithmic

Category	Value
cost since start	7.366422957
average	-3.924279286
Environmental impact	-3.929418926
Livelihood impact	3.62838893
Cultural impact	-4.223582462
Social impact	-4.100944629

(x-axis: -6, -4, -2, 0, 2, 4, 6, 8)

In the sample portfolio, some projects provide good results in all aspects (not only regarding the project objective; graph 3). Others provide some income, but do damage culturally, socially, and environmentally. This is exemplified in the second graph below (graph 4).

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
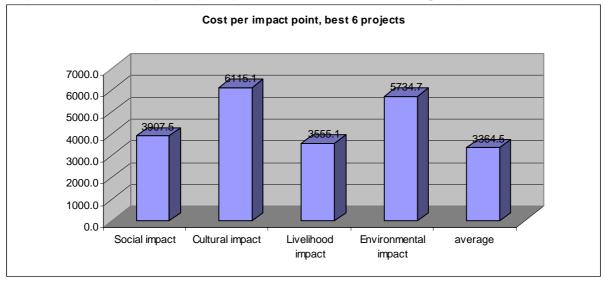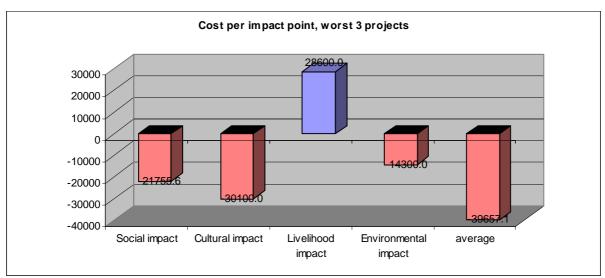www.aidrating.org

9

The six best projects cost a total of 12.2 million, while the three worst cost 23.25 Million Swiss Francs. The former, along with the impact profile of graph 3, is what the wide public will want, just like planners will, unless other priorities (political, commercial, etc) influence decision-making otherwise.

If this tool becomes widely used, the possibility to demonstrate desirability of choices this way may provide a strong argument for those who want effectivity in development work rather than rhetoric or vested interest.

The argument can be emphasized if the cost-impact ratio in terms of "impact points" (see keynote 4) is compared. This can be done as follows:

Graphs 5 and 6: Cost per impact point of „best" and „worst" group:

Cost per impact point, best 6 projects

Cost per impact point, worst 3 projects

These graphs show in proper proportion and in detail that the cost of the three "worst" projects to generate positive livelihood impact is not only eight times higher than for the same impact value of the "best" six (graph 5), but also that this remains the only positive impact of these projects.

We do not know of any other concept that will provide such findings in a similarly transparent and straightforward way. It works with facts relevant out in the "field". Let us remember that transparency on "field facts" counts as much as transparency on the flows of funds.

Original version presented at the Openaid Conference, Heinrich Böll Stiftung, Berlin, 29th September 2011

For questions please contact Mr Jan Stiefel, Project head, at jan.stiefel@ideas-expert.ch or at the address below!

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
www.aidrating.org

10

Annexes:

Annex 1: Ten Key Questions to Analyze Projects

Information on the following ten points is sought in order to complete an impact rating. If information is not available, then plausible assumption may be made. This and information availability on each point will bear on the transparency rating (round circle at bottom left of table 1 in text). The latter is a measure of reliability of rating, but also of quality of information gathered by the project planners:

Information	What information is needed
1. Project Environment	Geographic, dempgraphic, infrastructural **baseline.** Exact location and *size* of project intervention area.
2. Target Groups/ Beneficiaries	Knowledge about **target groups,** their *number*, and their characteristics, including problems that should be solved with (for) them.
3. Objectives of Project or Program	The „**Objective**" represents the concrete results a project tries to achieve. Example: "better health/ better income for…."
4. Who conducts project operations	**Who** is responsible for planning and executing the project? How is the project organized?
5. Mode of operation	**What** are the **operations** conducted in the project, with what equipment, what technology; what hardware and know-how is provided?
6. Risks and side effects	What risks or (negative) side effects are project planners expecting, and how are these countered (risk management)?
7. Start and duration of project	**How long** has the project lasted since its start, what phases; what is the planned end date?
8. Output, outcome, impact	Data, estimates, and/or indicators of results, namely related to project objective(s) but also others. **Impact Assessment.**
9. Sustainability	How is **sutainability** of results ensured after project departure?
10. Total cost	**Total Costs** for planning, materials, equipment, operations, personnel, advisories and follow-up from beginning

Note that these ten questions contain all the basics of any project or program. These questions are also one of two core subjects for the transparency rating (TCR) which AidRating conducts among large agencies in order to promote better information to the public. The other is coverage (of all projects/activities).

Annex 2: Dimensions

Dimensions are essential in order to compare any projects or project groups. They are part of the questions 1 and 2 in annex 1 (*in italic*). For human dimensions, the unit is number of persons/families, for environment related issues, dimension is km^2. By logarithmizing the dimensions, any size between "a few" (10^0) and "billions" (10^9) can be accommodated.

Annex 3: Time Frame

The impact rating takes care of the point in time a project is evaluated. An exact time vector can be applied. Simply put, 3 time frame groups are of interest:

	Typically relevant	Typically less relevant
1. During project activity	Direct project output	Long term results
2. briefly after project departure (0-3 years)	Direct and mid-term results	-
3. More than 3 years after project departure	Long term sustainable results	Direct project output

AidRating, c/o IDEAS Independent Development Experts Association, Postfach, CH-8401 Winterthur
Tel +41 52 203 52 50 Fax +41 52 203 52 55; e-mail: aidrating@ideas-expert.ch
www.aidrating.org

11

Annex 4: Rating database of 11 sample projects "Rural Region in The South"

Project titles	(A) Social impact	(B) Cultural impact	(C) Liveli-hood impact	(D) Environ-mental impact	Manage-ment	A-C Human Dimension (families)	Area Dimension (km^2) D	Average Cost CHF/y	Dura-tion years
Biofuel plantations in traditional farmer setting	-0.75	-1.00	0.25	-0.25	0.50	50'000	100'000	2'000'000	5
Sodis solar water sanitation	0.25	0.00	0.40	0.17	0.43	20'000	200'000	500'000	11
Smallholder crops A	0.50	0.20	0.50	0.70	0.50	500	1'000	200'000	5
Smallholder crops B	0.30	0.10	0.35	0.30	0.40	500	1'000	200'000	5
Smallholder crops C	0.00	0.00	0.10	-0.10	0.30	500	1'000	200'000	5
Large high yield livestock farm	-0.30	-0.20	0.20	-0.40	0.10	1'000	1'000	1'500'000	8
Livestock smallholder A	0.40	0.30	0.50	0.60	0.50	500	1'000	250'000	5
Livestock smallholder B	0.30	0.20	0.35	0.25	0.40	500	1'000	250'000	5
Livestock smallholder C	-0.10	0.00	0.10	-0.10	0.20	500	1'000	250'000	5
Rural alphabetization	0.60	0.70	0.50	0.10	0.60	1'000	1'000	500'000	4
Rural health	0.70	0.10	0.60	0.00	0.50	4'000	1'000	300'000	4