ISSN 2041-4668 (Online)

The MACBETH Approach for Multicriteria Evaluation of Rural Development Projects in face of Cross-Cutting Issues

Ramiro Sanchez-Lopez [1,2] and Carlos A. Bana e Costa [2,3]

[1] KERMIT, Research Unit Knowledge-based Systems, Ghent University, Belgium
[2] Centre for Management Studies of Instituto Superior Técnico, Technical University of Lisbon
[3] Operational Research Group, Department of Management, London School of Economics
ramirosanchezl@gmail.com
c.bana@lse.ac.uk

Working Paper LSEOR 09.107

ISSN 2041-4668 (Online)

First published in Great Britain in 2009
by the Operational Research Group, Department of Management
London School of Economics and Political Science

Typeset, printed and bound by:
The London School of Economics and Political Science
Houghton Street
London WC2A 2AE

Working Paper No: LSEOR 09.107

ISSN 2041-4668 (Online)

The MACBETH approach for multicriteria evaluation of rural development projects in face of cross-cutting issues

Ramiro Sanchez-Lopez [1,2] and Carlos A. Bana e Costa [2,3]

[1] KERMIT, Research Unit Knowledge-based Systems, Ghent University, Belgium

[2] Centre for Management Studies of Instituto Superior Técnico, Technical University of Lisbon

[3] Operational Research Group, Department of Management, London School of Economics

ramirosanchezl@gmail.com

c.bana@lse.ac.uk

Abstract: The European Union effort towards poverty reduction is based on the principle of equitable and participatory Sustainable Human Development. Meeting the terms of this principle entails the consideration of the so called "cross-cutting issues" on development matters. Cross-cutting issues are aspects of general interest that ought to be considered at all levels of intervention. They comprise, among others, issues like Human Rights, gender equity, the environmental concern, democracy as a social value, and the empowerment of the beneficiaries of development. Consequently, any evaluation process of development projects needs to take into account these issues at an operational level. Such a need, however, challenges the adequacy of traditional project appraisal methods, for instance Cost Benefit Analysis, to deal with the qualitative and even intangible nature of these new dimensions. This article describes how the MACBETH multicriteria approach was implemented in Bolivia, in order to help an important Programme for rural development build a new project evaluation system, taking into account cross-cutting issues through a series of interviews with individuals, as well as decision conferences attended by specialists and the Programme staff.

Keywords*: Multicriteria decision analysis, MACBETH, Development cooperation, Project evaluation, Cross-cutting issues.*

Working Paper LSEOR 09.107

1. Introduction

According to the Organisation for Economic Cooperation and Development (Brolin 2007), the European Union stands as the biggest donor of Official Development Assistance (ODA), providing more than a half of the total global amount. European Union's efforts for poverty reduction are based on the principle of equitable and participatory Sustainable Human Development (European Commission 2001). Such a principle places people as the main object of development, promoting individual and collective capabilities so that people are at the same time actors and beneficiaries of the development processes.

Meeting the terms of this principle entails the consideration of the so called "cross-cutting issues" on development matters. Cross-cutting issues are aspects of general interest that ought to be considered at all levels of intervention. They comprise issues like Human Rights, gender equity, the environmental concern, democracy as a social value, and the empowerment of the beneficiaries of development. Figure 1 shows the relationship between cross-cutting issues and Sustainable Human Development. Nowadays – and formally since the European Union Council, the European Parliament, and the European Commission delivered the "European Consensus for Development" (European Commission 2006) – these issues constitute basic guidelines for the implementation of European Union's development cooperation initiatives.

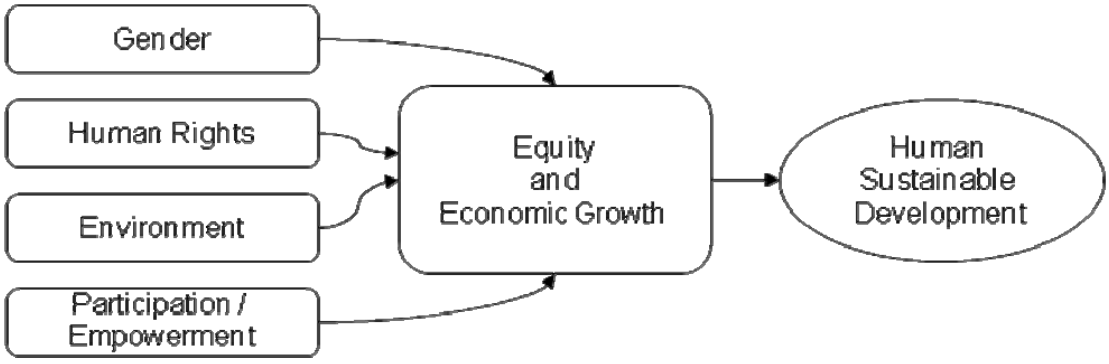

Figure 1: Relation between cross-cutting issues and the concept of Sustainable Human Development – adapted from (Gayraud and Quiroga 2005).

Particularly in Bolivia, cross-cutting issues have been incorporated in the European Union's cooperation initiatives. For instance, for an important rural development programme called PRAEDAC (“Support Programme for the Alternative Development Strategy in Chapare”) financed mostly by the European Commission (Financial agreement BOL/B7-310/96/041), four specific cross-cutting issues started to be considered in 2003 (PRAEDAC 2006): (1) the degree of participation of the population in development processes; (2) gender equity; (3) the environmental concern; and (4) the degree of endorsement of Human Rights. They had to be considered in the formulation and evaluation of development cooperation projects, whatever is their nature, magnitude or objectives.

Within the European Union, Cost Benefit Analysis is the most widely used methodology for project evaluation (European Commission 1997). However, the intangible and qualitative nature of cross-cutting issues compromises the monetary commensurability of many impacts of development projects. Consequently, Cost-Benefit Analysis is not suitable to evaluate the benefits that a given project entails in terms of, for instance, gender equity, people's empowerment, Human Rights or the environment.

The Logic Framework is also a widely used technique for formulating and evaluating development programmes and projects (EVO 1997). One of its main features is the use of the so called "Objectively Verifiable Indicators", verifiable in terms of quality, quantity and time. In PRAEDAC context, the Logic Framework technique proved to be difficult to implement. Development cooperation agencies usually operate in areas with limited transport infrastructures and dispersed population, which makes it difficult to reach all locations. These areas frequently suffer from intense population mobility and limited documentary control over individuals. All these factors imply that the evaluation of rural development cooperation projects using statistically driven indicators is too complicated and excessively expensive. Such was the case for the Logic Framework created for PRAEDAC previous to the incorporation of cross-cutting issues in its operations, which became inadequate for the evaluation of projects in the light of cross-cutting issues, pushing forward to the creation of a specific system for the accomplishment of that task.

The challenge posed by the design of a new evaluation system was originated:

(I) First, on the difficulty to define informative performance measures regarding the pros and cons of projects, about topics of a qualitative and multi-dimensional nature, which are often intangible and incommensurable in monetary terms. Such a difficulty poses two questions to answer:

> (1) How can cross-cutting issues be transformed into evaluation criteria?;
>
> and
>
> (2) How can a given performance (or "impact") level of a project, with respect to a cross-cutting issue, be described qualitatively and quantitatively?

(II) Second, on the difficulty to differentiate between a project's performance and its corresponding value (attractiveness or utility). Such a difficulty in turn poses another two questions to answer:

> (3) How can a number be associated to every performance level of each criterion, so that it characterizes the value it has for a given rural community?;
>
> and
>
> (4) How can a given project be evaluated globally, taking into account criteria with different importance from the community's point of view?

Given the kind of problem we were facing, with four well defined and intuitively multi-dimensional cross-cutting issues, the use of a multicriteria approach was perfectly natural and convenient. In this direction, it is frequent to find, within official agencies for development cooperation, solutions to these difficulties based on the discretional assignment of weights to evaluation criteria and of scores to project performances, in order to subsequently calculate global scores by a simple process of weighed sum. Unfortunately, some weighing and scoring processes widely used in practice break the necessary theoretical conditions underlying a multicriteria additive value model (French 1988; Keeney and Raiffa 1993; Krantz et al. 1971).

Methodologically correct procedures for weighting and scoring are based on qualitative and quantitative value judgments, depending on the techniques in use. For instance, SMART (Edwards and Barron 1994) requires direct numerical estimations, while

MACBETH (Bana e Costa et al. 2003; Bana e Costa and Vansnick 1994) requires qualitative judgments about differences of attractiveness (i.e., value), (Bana e Costa and Chagas 2004). The present article describes how the MACBETH approach was implemented in order to help PRAEDAC's Evaluation Unit build a new multicriteria project evaluation model. The process comprised a series of single-person interviews and decision conferences (Phillips and Bana e Costa 2007) carried out in cooperation with four specialists, one on each cross-cutting issue, plus other staff members of PRAEDAC.

This article presents in Section 2 the answers to questions (1) and (2) relative to model structuring; Section 3 presents the answers to questions (3) and (4) relative to the construction of the quantitative model; and finally, Section 4 delivers some conclusions.

2. Model structuring

2.1. How can cross-cutting issues be transformed into evaluation criteria?

In order to carry out the identification of criteria, we made use of a "top-down" structuring process (von Winterfeldt and Edwards 1986; Watson and Buede 1987), which required asking each specialist to describe succinctly what it would mean, from the point of view of his area of expertise, a completely satisfactory project – referred to as "prototype" in (Sanchez-Lopez 2005; 2008). Subsequently, the specialist was asked to write an essay justifying, in theoretical grounds, the corresponding prototype just formulated. The purpose of such an essay was to subject the prototype and its corresponding theoretical explanation to the comments of the other specialists, in order to enrich it with additional concerns that the original formulation ignored or took too lightly. This recursive process was expected to develop a common understanding of the key-concerns, reflected in an improved version of the prototype. Let us take as example the cross-cutting issue "Participation" (Figure 2), for which the following prototype was formulated (Bazoberry Chali 2005):

> *The project addresses an 'inclusive we', includes democratic mechanisms for conflict resolution, and the actors hold developed capabilities and opportunities to participate in processes of local development.*

As a subsequent step, a semantic analysis of the prototype allowed to identify the embedded key-concerns that should be taken as criteria for project evaluation: “Inclusion”, “Conflict management”, “Capabilities enabling participation” and “Opportunity to participate”. The same was accomplished for every cross-cutting issue, yielding in total 17 evaluation criteria, as shown in the value tree in Figure 2.

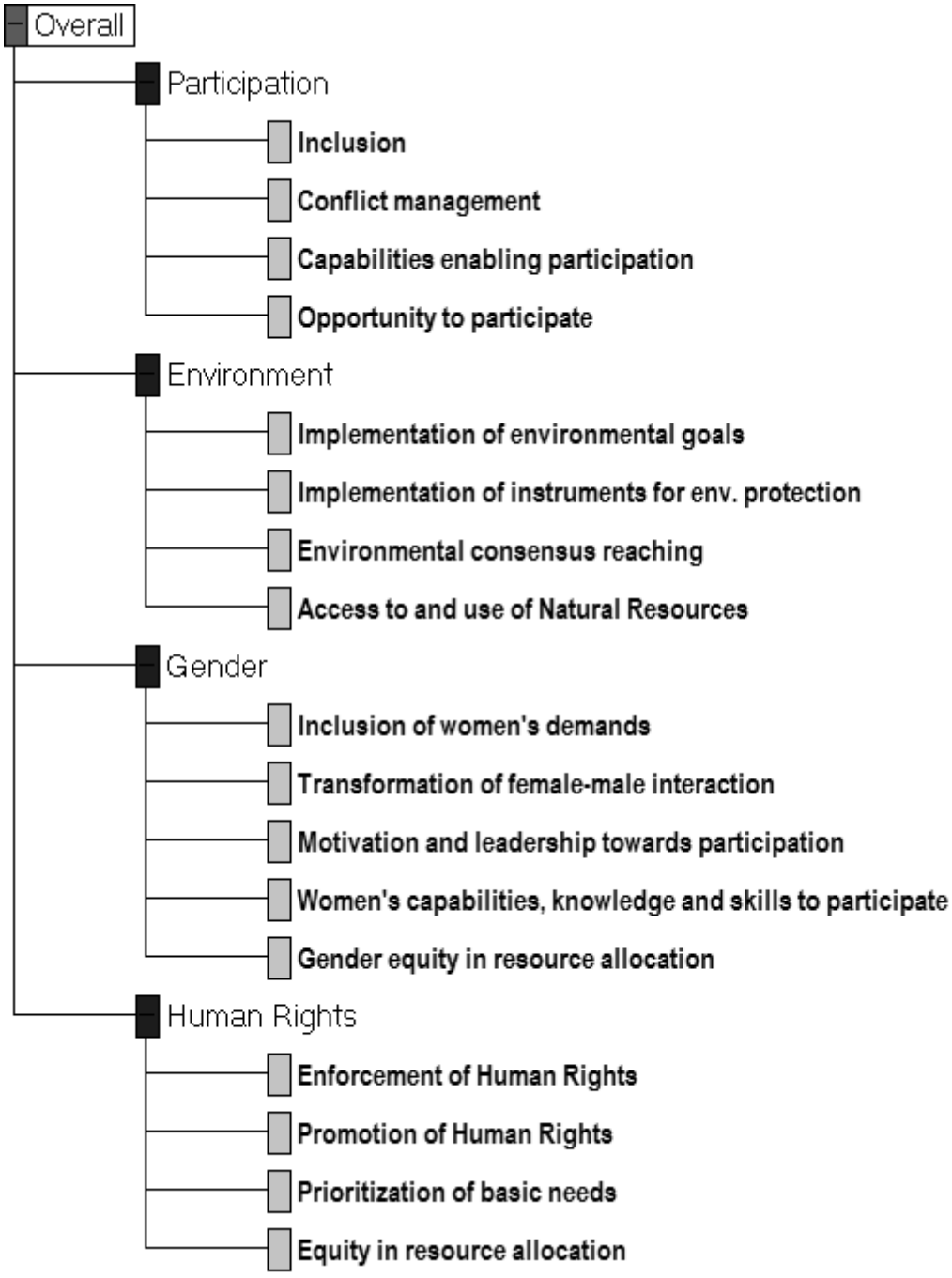

Figure 2: Cross-cutting issues and their corresponding evaluation criteria. Source: (PRAEDAC 2005)

Each specialist delivered a detailed description of the meaning of each of the criteria of his/her area of expertise (PRAEDAC 2005). These descriptions helped to conclude that there were no relevant problems of non-completeness, redundancy or preferential dependence of the criteria (Kirkwood, 1997).

2.2. How can performance levels be described qualitatively and quantitatively?

Given the intangible nature of concerns that derives from cross-cutting issues, it is not surprising that the corresponding criteria merge several interrelated elementary dimensions. Therefore, in order to adequately describe performance levels on each evaluation criterion, it was first necessary to identify its most relevant dimensions and subsequently describe sub-levels of performance for each dimension. This is the starting point of the application of the approach to construct a multidimensional performance scale proposed in (Bana e Costa and Beinat 2005) and implemented in real world contexts in (Bana e Costa et al. 2006, 2008). Next, plausible performance levels of the scale are constructed by combining (i.e. concatenating) sub-levels of performance of the different dimensions. The constructed scale is then defined after ranking the plausible performance levels, from the most attractive to the least (Table 1).

Table 1: A procedure to develop a multidimensional descriptor.

Steps	Tasks
Step 1	Define a discrete set of performance levels in terms of each of the component dimensions.
Step 2	Establish all possible combinations of the levels of the various dimensions.
Step 3	Eliminate infeasible (implausible) combinations.
Step 4	Compare the desirability of the feasible combinations and group those that are judged to be indifferent in terms of the criterion; each group of profiles forms an equally plausible performance level of the scale (if convenient, give a label to each level). Rank the plausible levels by decreasing relative attractiveness in terms of the criterion.
Step 5	Make a textual description of each plausible performance level, as detailed as appropriate and as objective as possible.

For instance, for the criterion "Inclusion" from cross-cutting issue "Participation", the specialist identified two dimensions worth to be taken into account: (X) The extent to which different stakeholders are included in the project's decision making processes (termed "Who?"); and (Y) the extent to which the participation takes place in a truly participative way (termed "How?"). Subsequently, for each dimension, the specialist articulated three sub-levels of performance (X.1, X.2, X.3 and Y.1, Y.2, Y.3 respectively), ordered from the most attractive to the least. Every level was formulated avoiding as much as possible the use of ambiguous terms. This is a critical step, since the appeal of the evaluation model depends to a great extent on the careful selection of linguistic terms used to describe sub-levels of performance.

Example: Criterion "Inclusion" from cross-cutting issue "Participation"

Dimension X: "Who?"

The project manuscript explicitly incorporates in the decision making processes...

X.1 ...a wide variety of actors, including opposing forces and indirect actors (in other words, the project has been formulated following a truly inclusive rationale, paying attention to actors that are affected by the project but will seldom have a spontaneous voice in the process).

X.2 ...only actors as pointed out by the norm (example for municipal development: municipal authorities, technicians and the vigilance committee); or by the project formulation (example: industry representatives or farm associations; or the financing institution).

X.3 ...only the financing institution and/or the project-management staff (in other words, the project follows a political or bureaucratic rationale).

Dimension Y: "How?"

The project manuscript explicitly dictates actors to be included...

Y.1 ...assertively (in other words, as protagonists, actively).

Y.2 ...passively (in other words, actors are not persuaded enough to deliver propositions).

Y.3 ...instrumentally (in other words, careless, with the only purpose of fulfilling requirements or norms from the financing institution).

The combination of two sub-levels of performance, one for each dimension, yields a sentence that describes (qualitatively) the performance of a project with respect to a given criterion. For instance, the concatenation of sub-levels X.1 and Y.3 in our example yields the following sentence:

> *The project manuscript explicitly incorporates in the decision making processes a wide variety of actors, including opposing forces and indirect actors (in other words, the project has been formulated following a truly inclusive rationale, paying attention to actors that are affected by the project but will seldom have a spontaneous voice on the process). Moreover, the project manuscript explicitly dictates actors to be included instrumentally (in other words, careless, with the only purpose of fulfilling requirements or norms from the financing institution).*

The next step is to assemble all possible combinations of sub-levels of performance, identifying and discarding implausible combinations. Once all plausible combinations are defined, they have to be ranked from the most attractive to the least. In some cases, two or more combinations can be found to be equivalent in terms of attractiveness. In such a case they would all conform the same performance level, as in Levels D and E of Table 2. Finally, the scale levels are expressed in natural language. Table 3 shows the statements of all performance levels for criterion "Inclusion".

Table 2: Combination of sub-levels of performance for the creation of scale levels

Dimension: X	**Combinations**	**Scale levels**
Sub-level: X.1	(X.1, Y.1)	Level A: (X.1, Y.1)
Sub-level: X.2	(X.1, Y.2)	Level B: (X.2, Y.1)
Sub-level: X.3	(X.1, Y.3)	Level C: (X.1, Y.2)
	(X.2, Y.1)	Level D: (X.2, Y.2) or (X.1, Y.3)
	(X.2, Y.2)	Level E: (X.3, Y.1) or (X.3, Y.2) or (X.2, Y.3)
Dimension: Y	(X.2, Y.3)	Implausible: (X.3, Y.3)
Sub-level: Y.1	(X.3, Y.1)	
Sub-level: Y.2	(X.3, Y.2)	
Sub-level: Y.3	(X.3, Y.3)	

Table 3: Constructed scale for criterion Inclusion

Level A	(X.1, Y.1) The project manuscript explicitly incorporates in the decision making processes a wide variety of actors, including opposing forces and indirect actors (in other words, the project has been formulated following a truly inclusive rationale, paying attention to actors that are affected by the project but will seldom have a spontaneous voice in the process). Moreover, the project manuscript explicitly dictates actors to be included assertively (in other words, as protagonists, actively).
Level B	(X.2, Y.1) The project manuscript explicitly incorporates in the decision making processes only actors as pointed out by the norm (example for municipal development: municipal authorities, technicians and the vigilance committee); or by the project formulation (example: industry representatives or farm associations; or the financing institution). Moreover, the project manuscript explicitly dictates actors to be included assertively (in other words, as protagonists, actively).
Level C	(X.1, Y.2) The project manuscript explicitly incorporates in the decision making processes a wide variety of actors, including opposing forces and indirect actors (in other words, the project has been formulated following a truly inclusive rationale, paying attention to actors that are affected by the project but will seldom have a spontaneous voice in the process). Moreover, the project manuscript explicitly dictates actors to be included passively (in other words, actors are not persuaded enough to deliver propositions).
Level D	(X.2, Y.2) The project manuscript explicitly incorporates in the decision making processes only actors as pointed out by the norm (example for municipal development: municipal authorities, technicians and the vigilance committee); or by the project formulation (example: industry representatives or farm associations; or the financing institution). Moreover, the project manuscript explicitly dictates actors to be included passively (in other words, actors are not persuaded enough to deliver propositions).
	(X.1, Y.3) The project manuscript explicitly incorporates in the decision making processes a wide variety of actors, including opposing forces and indirect actors (in other words, the project has been formulated following a truly inclusive rationale, paying attention to actors that are affected by the project but will seldom have a spontaneous voice in the process). Moreover, the project manuscript explicitly dictates actors to be included instrumentally (in other words, careless, with the only purpose of fulfilling requirements or norms from the financing institution).
Level E	(X.3, Y.1) The project manuscript explicitly incorporates in the decision making processes only the financing institution and/or the project-management staff (in other words, the project follows a political or bureaucratic rationale). Moreover, the project manuscript explicitly dictates actors to be included assertively (in other words, as protagonists, actively).
	(X.3, Y.2) The project manuscript explicitly incorporates in the decision making processes only the financing institution and/or the project-management staff (in other words, the project follows a political or bureaucratic rationale). Moreover, the project manuscript explicitly dictates actors to be included passively (in other words, actors are not persuaded enough to deliver propositions).
	(X.2, Y.3) The project manuscript explicitly incorporates in the decision making processes only actors as pointed out by the norm (example for municipal development: municipal authorities, technicians and the vigilance committee); or by the project formulation (example: industry representatives or farm associations; or the financing institution). Moreover, the project manuscript explicitly dictates actors to be included instrumentally (in other words, careless, with the only purpose of fulfilling requirements or norms from the financing institution).

3. Evaluation

3.1. How can a meaningful numerical value be associated to every performance level?

3.1.1. Numerical and nonnumerical techniques for the construction of value functions

A multidimensional constructed descriptor is very useful to produce a comprehensive qualitative description of performance, by associating one of the scale's performance levels to the project being evaluated. Yet, as we said before, one thing is the project's performance and quite another is the value (or attractiveness) that such a performance conveys. In order to measure the attractiveness of projects, it is required to construct a value function for every evaluation criterion in the model. Such a task can be accomplished through different numerical and nonnumerical techniques (Belton and Stewart 2002; Kirkwood 1997; von Winterfeldt and Edwards 1986). In cases where discrete scales of performance were constructed, direct rating (von Winterfeldt and Edwards 1986) is widely used. This technique requires three main tasks: (1) to select two reference levels for the rating scale, usually the highest and the lowest performance levels; (2) to assign numerical values to these reference levels, usually 100 and 0 respectively; and (3) to ask the evaluator (be it a single person or a group) to assign to the each of the remaining performance levels a score that denotes its attractiveness with respect to the two reference levels. The resulting rating scale should be constructed in such a way that the difference between the scores of two performance levels reflects their difference of attractiveness for the evaluator. In contrast, MACBETH is a technique that enables the construction of value functions derived from qualitative (i.e. nonnumerical) judgements about the difference of attractiveness between every two performance levels of the scale. Thus, in this way, MACBETH avoids the difficulty (von Winterfeldt and Edwards 1986) or cognitive uneasiness (Fasolo and Bana e Costa 2009) experienced by some evaluators when expressing their preference judgments numerically.

3.1.2. The MACBETH approach

MACBETH elicits preference information in the form of qualitative judgments about the difference of attractiveness between every two performance levels. The evaluator is required to judge qualitatively how big such a difference is, using one of the six MACBETH semantic categories: *very weak*, *weak*, *moderate*, *strong*, *very strong* and *extreme*. This characteristic is the origin of the acronym MACBETH: "**M**easuring **A**ttractiveness by a **C**ategory **B**ased **E**valuation **T**ec**h**nique". The M-MACBETH Decision Support System (Bana Consulting 2005), which implements the MACBETH approach, allows using not only one semantic category but several at the same time, making it possible for the evaluator to express judgemental hesitation. The answers so expressed are entered into the MACBETH judgment matrix, filling in the cells to the right of the matrix's main diagonal.

For illustrative purposes, let us use the criterion Inclusion. Figure 3 shows the specialist's qualitative judgments concerning the difference of attractiveness between the corresponding performance levels. Starting from the difference between the highest level (i.e. the most attractive) and the lowest one (i.e. the least attractive), judged in this case as "extreme" by the specialist, he continued to judge the difference of attractiveness between every two consecutive levels of the scale. This means that the main diagonal of the matrix was first completed. Later on, the specialist was asked to judge the difference of attractiveness between the first level and the third, the second level and the fourth and so on, completing in this way the second diagonal of the matrix. In this respect, (Bana e Costa and Chagas 2004) suggests an alternative way to perform this operation, which consist on starting to fill in, from top to bottom, the last column of the matrix, next to fill in from right to left the first row, and finally to complete the main diagonal of the matrix. In our case, we chose this second procedure for the weighting process of criteria (*cf.* Section 3.2.1). According to (Bana e Costa et al. 2008), both procedures are correct, and still others can be devised rightfully.

	A	B	C	D	E	Current scale
A	no	moderate	moderate	positive	extreme	100
B		no	weak	moderate	positive	73
C			no	positive	strong	55
D				no	moderate	36
E					no	0

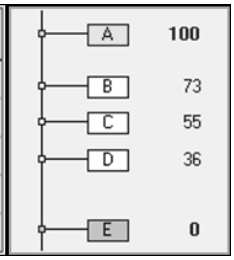

Figure 3: MACBETH matrix and its value function for criterion Inclusion

Let m be the number of performance levels. It is not compulsory to ask for all the $m(m-1)/2$ judgments. For instance, the MACBETH matrix presented in Figure 3 shows two cells containing a "*positive*" difference of attractiveness. It means that those judgments are merely ordinal. Despite that m-1 is the minimum acceptable number of judgments, which corresponds either to the last column or the first row or the main diagonal of the matrix, good practice recommends to perform some additional judgments in order to cross-check consistency (Bana e Costa et al. 2008).

Despite that two evaluators may interpret MACBETH semantic categories differently (for instance, "strong" can correspond to a different sensation of difference in attractiveness depending on the evaluator), the key characteristic of the process is that if the difference of attractiveness assigned to two given performance levels is stronger than the difference of attractiveness assigned to other two performance levels, the value function should be such that v_1-v_2 > v_3-v_4, where v_1, v_2, v_3, v_4 represent the corresponding scores of the four performance levels.

The judgments are entered into the MACBETH matrix as the evaluator formulates them. As each judgement is entered into the matrix, the M-MACBETH software uses a mathematical programming algorithm (Bana e Costa et al. 2005) to check its consistency with the judgments already inserted into the matrix and if an inconsistency is detected, suggestions to overcome it presented to the evaluator. Once a consistent MACBETH matrix is available, the software uses a linear program to suggest a preliminary value function (the so called MACBETH scale). The proposed scores are displayed graphically, together with the interval within which the evaluator can vary each score and still remain consistent with his qualitative judgements. Figure 3 shows a validated value function for criterion "Inclusion".

A similar process was accomplished in order to construct the corresponding value functions of the other criteria of PRAEDAC. The entire process was developed in a single session for each cross-cutting issue. During the session, the specialist of the corresponding cross-cutting issue constructed the value functions for all its criteria, guided by the first author of this article with the support of M-MACBETH (PRAEDAC 2005).

3.2. How can a project be evaluated globally taking into account different weights of criteria?

3.2.1. Criteria weighting

In order to explain the process of criteria weighting, let us consider two hypothetical projects P_1 and P_2, equally attractive in all criteria but two. For illustrative purposes, let us choose criterion [Exercise] and criterion [Promotion] both pertaining to the cross-cutting issue Human Rights (*cf.* Figure 2). Suppose that the performance of P_1 in criterion [Exercise] corresponds to the highest performance level; while its performance in criterion [Promotion] corresponds to the lowest performance level. In contrast, the performance of P_2 in criterion [Exercise] corresponds to the lowest performance level; while its performance in criterion [Promotion] corresponds to the highest performance level. Which of the two projects, P_1 or P_2, would contribute the most to (i.e. would be more important for) the cross-cutting issue Human Rights?

The answer to this question mirrors the relationship between the weights, $w_{[Exercise]}$ and $w_{[Promotion]}$, corresponding to these two criteria. Thus, if the answer is favourable to P_1, the weight of [Exercise] should be higher than the weight of [Promotion], and vice versa. If the answer suggests the same contribution to Human Rights from P_1 or P_2, the weights of [Exercise] and [Promotion] should be equal. Let us now suppose that the answer is favourable to P_1, that is, P_1 is globally more attractive than P_2. Then, $w_{[Exercise]} > w_{[Promotion]}$. It is not difficult to realize that, if the highest performance level on criterion [Exercise] is, in a different context, much lower than originally, the most favourable project now for Human Rights would be P_2. This would imply a rank reversal of weights, i.e., $w_{[Promotion]} > w_{[Exercise]}$. Such is the reason why weights of criteria shouldn't be determined without taking into account the range of their corresponding

performance scales, as when criteria weights are determined directly in terms of relative importance – Keeney (1992) called this the "most common critical mistake".

Consequently, in our case, criteria weights were determined separately for each cross-cutting issue, in accordance with the MACBETH weighting process. For illustrative purposes, let us consider the cross-cutting issue Human Rights and its four criteria: [Exercise], [Promotion], [Prioritization] and [Equity]. In order to weight them, the specialist was asked to consider the following five hypothetical projects: A hypothetical project named [All lower], whose performance profile in Human Rights shows the lowest performance levels in the four criteria; plus other four hypothetical projects named [Exercise], [Promotion], [Prioritization] and [Equity], each of them constructed from [All lower] by switching, within the corresponding criterion, from the lowest performance level to the highest performance level.

Next, the specialist was required to rank them, in terms of Human Rights, from the most attractive to the least attractive. Once the hypothetical projects were ranked, the MACBETH weighting process was followed to pairwise compare hypothetical projects using the MACBETH semantic categories. The corresponding weighting scale proposed by M-MACBETH was then discussed, in order to define the final weights of the criteria (Figure 4). A similar process was accomplished for the other cross-cutting issues.

	[Exercise]	[Prioritization]	[Equity]	[Promotion]	[All lower]	Current scale
[Exercise]	no	very weak	weak	moderate	extreme	30
[Prioritization]		no	weak	weak	extreme	27
[Equity]			no	very weak	v. strong	23
[Promotion]				no	v. strong	20
[All lower]					no	0

Figure 4: MACBETH weighting for criteria under cross-cutting issue Human Rights

To complete the evaluation model, weights should be defined for the higher-level criteria, i.e. at the level of cross-cutting issues (Figure 2). These weights were of a strategic nature, different from the technical nature of criteria weights inside each cross-cutting issue. This is beyond the specialists' scope of intervention, requiring also and above all the intervention of PRAEDAC top managers. For this reason, no precise weights were defined for the cross-cutting issues (since it could have been considered illegitimate in social terms) letting their discussion aside for an ulterior process.

4. Conclusions

The evaluation model so created was used to produce a spreadsheet computer program (Sanchez Lopez 2005) that enabled the implementation of the model in PRAEDAC's day to day operations. Due to the time savings its use implied for the evaluators, the programme was named "Rapid Project Evaluation System regarding cross-cutting issues".

The Monitoring Unit of PRAEDAC proposed the following 5-step working procedure for the fieldwork evaluation of projects (Gayraud 2005):

Step 1 Project formulation and its corresponding delivery to the Monitoring Unit;

Step 2 Project evaluation by means of the spreadsheet programme, accomplished by the Monitoring Unit;

Step 3 Participative diagnosis meeting, attended by the Monitoring Unit and the staff in charge of the project. Such a meeting is intended to produce adjustments in the project formulation, based on a joint analysis of the individual evaluation results and conclusions about the project performance on cross-cutting issues.

Step 4 Project reengineering according to the necessary adjustments defined in Step 3.

Step 5 In itinere project evaluation and monitoring, intended to verify the correct implementation of cross-cutting issues during the project execution phase.

In this way, the spreadsheet programme was put into service at the level of projects and activities, making it possible for the Monitoring Unit to provide support and assessment to the technical staff in charge of projects, regarding the better inclusion of cross-cutting issues on their day to day actions.

According to PRAEDAC Monitoring Unit staff (Gayraud 2005), the descriptive nature of the model's performances scales provided the necessary guidelines for the inclusion of cross-cutting issues, diminishing significantly the need to hire external consulting services for this purpose. Moreover, an additional asset was the self-evident commitment and awareness the model generated on technicians, who may have been less sensitive to cross-cutting issues otherwise.

In addition to project evaluation, and despite the model was not specifically conceived for that purpose, the PRAEDAC Monitoring Unit made use of it in the PRAEDAC's

Annual Operative Plan and Global Operative Plan. By means of the model's value tree was used as a check-list in order to identify specific actions regarding cross-cutting issues. These actions were included in the matrix of activities of the different units (also called "components") of PRAEDAC.

Acknowledgments

The authors acknowledge the invaluable contribution of the Bolivian experts in cross-cutting issues Guillermo Bazoberry Chali, Carlos Crespo Flores, Javier García Soruco and Maria Lourdes Zabala Canedo. The case-study described in this paper was developed with the financial support of VLIR (Flemish Interuniversity Council) and PREADAC.

References

Bana Consulting. (2005). "M-MACBETH Version 1.1: User Manual." Lisbon.

Bana e Costa, C. A., and Beinat, E. (2005). "Model-structuring in public decision-aiding." *LSEOR 05.79*, The London School of Economics and Political Science.

Bana e Costa, C. A., and Chagas, M. P. (2004). "A career choice problem: An example of how to use MACBETH to build a quantitative value model based on qualitative value judgments." *European Journal of Operational Research*, 153(2), 323-331.

Bana e Costa, C. A., De Corte, J. M., and Vansnick, J. C. (2003). "MACBETH." *LSEOR 03.56*, The London School of Economics and Political Science.

Bana e Costa, C. A., De Corte, J. M., and Vansnick, J. C. (2005). "On the mathematical foundations of MACBETH." Multiple Criteria Decision Analysis: State of the Art Surveys, J. Figueira, S. Greco, and M. Ehrgott, eds., Springer, New York, 409-442.

Bana e Costa, C. A., Lourenco, J. C., Chagas, M. P., and Bana e Costa, J. C. (2008). "Development of Reusable Bid Evaluation Models for the Portuguese Electric Transmission Company." *Decision Analysis*, 5(1), 22-42.

Bana e Costa, C. A., and Vansnick, J. C. (1994). "MACBETH – An interactive path towards the construction of cardinal value functions." *International Transactions in Operational Research*, 1(4), 489-500.

Bazoberry Chali, G. (2005). "La Participación Ciudadana en el Desarrollo Local." Responsabilidad Social y Desarrollo: Análisis Multicriterio para la Evaluación de Proyectos, PRAEDAC, ed., Delegación de la Comisión Europea en Bolivia y Viceministerio de Desarrollo Alternativo del Gobierno de Bolivia, Cochabamba, 115-130.

Belton, V., and Stewart, T. J. (2002). *Multiple Criteria Decision Analysis: An Integrated Approach*, Kluwer Academic Publishers, Boston.

Brolin, T. (2007). "The EU and its policies on Development Cooperation: A learning material." *2007:1*, Swedish Agency for Development Evaluation, Karlstad.

Edwards, W., and Barron, F. H. (1994). "SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement." *Organizational Behavior and Human Decision Processes*, 60(3), 306-325.

European Commission. (1997). *Manual: Financial and economic analysis of development projects*, Office for Official Publications of the European Communities, Luxembourg.

European Commission. (2001). "Communication from the Commission to the Council and the European Parliament of 8 May 2001: The European Union's role in promoting human rights and democratisation in third countries." European Commission.

European Commission. (2006). "The European Consensus on Development: The development challenge - Joint statement by the Council and the representatives of the governments of Member States meeting within the council, the European Parliament and the Commission on European Union Development Policy: "The European Consensus"." *Official Journal of the European Union*, 2006/C 46/01.

EVO. (1997). *Evaluación: Una herramienta de gestión para mejorar el desempeño de los proyectos*, Banco Interamericano de Desarrollo (BID). Oficina de Evalulación (EVO), Washington.

Fasolo, B., and Bana e Costa, C. A. (2009). "Tailoring value elicitation to decision makers' numeracy and fluency: Expressing value judgments in numbers or words?" London School of Economics.

French, S. (1988). *Decision Theory: An Introduction to the Mathematics of Rationality* Ellis Horwood, Chichester.

Gayraud, E. (2005). "Implementación del Sistema de Evaluación Rápida en el PRAEDAC." Responsabilidad Social y Desarrollo: Análisis Multicriterio para la Evaluación de Proyectos, PRAEDAC, ed., Delegación de la Comisión Europea en Bolivia y Viceministerio de Desarrollo Alternativo del Gobierno de Bolivia, Cochabamba, 149-154.

Gayraud, E., and Quiroga, E. (2005). "Promoviendo el Desarrollo con Responsabilidad Social." PRAEDAC, ed., Delegación de la Comisión Europea en Bolivia y Viceministerio de Desarrollo Alternativo del Gobierno de Bolivia, 11-21.

Keeney, R. L. (1992). *Value Focused Thinking: A Path to Creative Decision Making*, Harvard University Press, Cambridge.

Keeney, R. L., and Raiffa, H. (1993). *Decision with Multiple Objectives: Preferences and Value Tradeoffs*, Cambridge University Press, Cambridge.

Kirkwood, C. W. (1997). *Strategic Decision Making: Multiobjective Decision Analysis With Spreadsheets*, Duxbury Press, Belmont.

Krantz, D. H., Luce, R. D., Suppes, P., and Tverksky, A. (1971). *Foundations of Measurement. Volume I: Additive and Polynomial Representations*, Academic Press, New York.

Phillips, L. D., and Bana e Costa, C. A. (2007). "Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing." *Annals of Operations Research*, 154(1), 51-68.

PRAEDAC. (2005). "Responsabilidad Social y Desarrollo: Análisis Multicriterio para la Evaluación de Proyectos." Delegación de la Comisión Europea en Bolivia y Viceministerio de Desarrollo Alternativo del Gobierno de Bolivia, Cochabamba, 177.

PRAEDAC. (2006). *8 años en el Desarrollo - Trópico de Cochabamba*, Viceministerio de Desarrollo Alternativo del Gobierno de Bolivia y Delegación de la Comisión Europea en Bolivia, Cochabamba.

Sanchez-Lopez, R. (2005). "El Sistema de Evaluación Rápida de Proyectos respecto a Temas Transversales." Responsabilidad Social y Desarrollo: Análisis Multicriterio para la Evaluación de Proyectos, PRAEDAC, ed., Delegación de la Comisión Europea en Bolivia y Viceministerio de Desarrollo Alternativo del Gobierno de Bolivia, Cochabamba, 55-74.

Sanchez-Lopez, R. (2008). "Evaluating development projects based on multiple intangible criteria: Theoretical framework and applications to coca producing regions of Bolivia," Ghent University, Ghent.

~~Sanchez-Lopez, R. (2005). "El Sistema de Evaluación Rápida de Proyectos respecto a Temas Transversales." Responsabilidad Social y Desarrollo: Análisis Multicriterio para la Evaluación de Proyectos, PRAEDAC, ed., Delegación de la Comisión Europea en Bolivia y Viceministerio de Desarrollo Alternativo del Gobierno de Bolivia, Cochabamba, 55-74.~~

von Winterfeldt, D., and Edwards, W. (1986). *Decision Analysis and Behavioural Research*, Cambridge University Press, New York.

Watson, S. R., and Buede, D. M. (1987). *Decision Synthesis: the principles and practice of decision analysis*, Cambridge University Press, Cambridge.