Chapter 13
Applying the EPISSURE Approach for the Evaluation of Business Sponsorship Performance

Stéphane André and Bernard Roy

Abstract This paper presents the application of an approach designed to evaluate non-financial performance in companies. Within a defined perimeter, the approach called EPISSURE produces an "evaluation of non-financial performance with a hierarchical set of synthesis indicators co-constructed during a process of framed dialogue." The paper discusses how the EPISSURE approach was tested and set up within several companies for the purpose of evaluating sponsorship projects and deciding on their follow-up. Test results seem to indicate that the EPISSURE approach is decidedly appropriate for evaluating non-financial performance.

13.1 Introduction

This paper examines the application of an approach designed to evaluate non-financial performance in companies. The application is applied to sponsorship projects. The approach was designed within a much broader vision [4]. The purpose was to meet the ever-growing need of business to take account of non-financial performance [7, 8, 14, 31, 40, 18], from the perspective of decision aiding. Synthesising non-financial performance with an operational perspective involved designing an approach that could:

- Adjust to special cultural, sociopolitical and environmental traits that, within the given company, delineate the perimeter of the relevant non-financial performance components

Stéphane André
LAMSADE, Université Paris-Dauphine, 120/122 rue Réaumur - 75002 Paris, France e-mail: stephane.andre@dauphine.fr

Bernard Roy
LAMSADE, Université Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, F-75 775 Paris Cedex 16, France e-mail: roy@lamsade.dauphine.fr

C. Zopounidis and P.M. Pardalos (eds.), *Handbook of Multicriteria Analysis*, Applied Optimization 103, DOI 10.1007/978-3-540-92828-7_13,

- Take account of the multiple factors involving heterogeneous aspects that can often only be grasped in qualitative terms, entailing some share of ill determination that should not be overlooked
- Produce a few synthesis evaluations that the various stakeholders will consider legitimate and appropriate to the set goal.

The approach (presented in Section 13.2) is designed to produce an "**evaluation of non-financial performance with a hierarchical set of synthesis indicators co-constructed within a process of framed dialogue**", within a defined perimeter. The approach is called EPISSURE (splice), which is a nautical term meaning a joint made by splicing, i.e., to join or unite (ropes or rope ends) by weaving together the end strands (adapted from Webster's New World Dictionary, 1976). The term reflects one of the concerns that helped guide the design of the approach, i.e., to interweave the separate components of performance. Section 13.3 examines how the EPISSURE approach was tested and set up in several companies for the evaluation sponsorship projects (with a view to deciding on their follow-up). Section 13.4 describes the main results at the end of the tests. Finally, the conclusion addresses the main hurdles that have to be overcome for the application of the EPISSURE approach.

13.2 The EPISSURE Approach

After outlining the approach, we describe the tools that were used to define the synthesis indicators, the dialogue process that is an integral part of the approach, and the set-up of the EPISSURE approach. This procedure clears the way for the users' proper appropriation of the tools and legitimises the resulting evaluations for the main stakeholders.

13.2.1 The Outline of the Approach

Two normative principles to ground the approach were laid down *ex-ante*.

- Principle no. 1: The suggested evaluation is hierarchical, i.e., classified into successive synthesis levels. These levels match a hierarchy of responsibilities into the organization. The importance of the hierarchical agglomerative clustering principle is advocated in numerous research papers on performance evaluation [1, 2, 9, 22, 27]. About evaluation in a context of decision aiding, authors as Keeney [23] and Saaty [39] highlighted the interest of hierarchical levels. Nevertheless they did not establish a link with the hierarchy of responsibilities into an organization.
- Principle no. 2: At each hierarchical level, the evaluations (except perhaps for some at the lowest levels) rely on purely ordinal verbal scales. The number of

degrees on the scales must be adjusted to its matching level; also, the number of degrees must be high enough to mirror evolutions and to be understandable by the stakeholders operating at the said level.

Four levels were selected. The purpose of the first set-up stage of EPISSURE was to detail how the levels were to be understood and tailored within the company where the approach was applied. The four levels are called, and characterised as follows:

- *Level 1 or elementary indicator level:* It involves the elementary indicators that are considered relevant for evaluating the performance of actions [33, 34](chapter 8).
- *Level 2 or criteria level:* It concerns the major points of view for defining the groups of elementary indicators: groups wherein each indicator is related to the same significance axis, which matches a criterion [36] (page 46).
- *Level 3 or level of the actions to evaluate:* We use this term to refer to the typical entity whose performance has to be evaluated. In each concrete case of application, the term must be replaced with the one that designates what - viz., a project, an investment, a procedure, an operational mode, etc. - must be evaluated in the company.
- *Level 4 or level of organisation:* It involves the structure in charge of the different actions.

The elementary indicators of level 1 are what may be considered as the primary data providing the basis for working out all the other evaluations. The elaboration involves different syntheses, each one taking into account the evaluations of the level immediately below. Three syntheses are required (see Figure 13.1):

- The **Operational** synthesis that clears the way for going from the evaluations on the elementary indicators to the evaluations on the criteria.
- The **Intermediate** synthesis that makes it possible to go from the evaluations on the criteria to the evaluations of the actions.
- The **Managerial** synthesis that involves going from the evaluations of the actions to the evaluation of the organisation.

Each synthesis uses tools that are described in the next subsection. Compliant with the second principle (see beginning of this subsection), the tools solely involve purely ordinal scales (save certain exceptions). For the following reasons, these types of scales [35] are deemed the most appropriate for non-financial performance evaluations:

1. The significance of each degree, defined by a verbal description, is easily understandable by the different stakeholders, corporate in-house and external players alike.
2. Any errors of interpretation due to a definition of degrees solely based on numeric data are avoided. These errors stem from individuals' tendency to attribute - automatically and without good reason - significance to the arithmetic relations

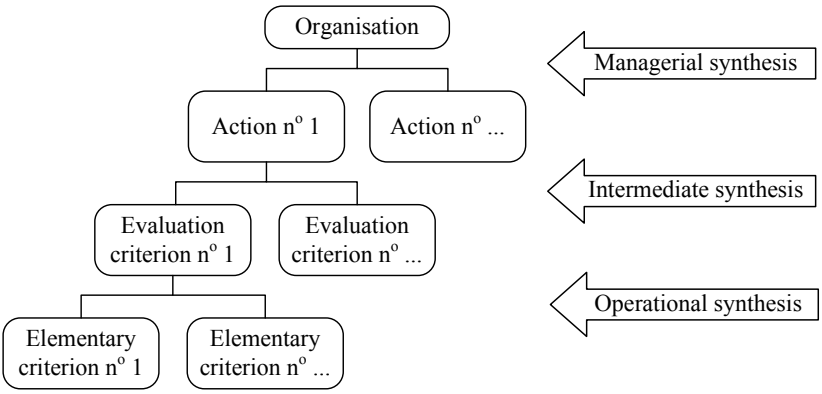

Fig. 13.1 Hierarchical Structure and Corresponding Syntheses

existing between numbers, i.e., a ratio equal to two translates into twice as much satisfaction or intensity of preference; equal differences reflect equal satisfaction or preference variations. This is only true if the numbers express a quantity or a variation of quantity; said quantity being understood as based on a well-defined unit. Non-financial performance cannot usually e grasped in these terms.

3. Stating that one degree rather than the next closest must be used to evaluate the relevant object (whatever the level) may seem arguable in many instances. This ill determination may be properly taken into account by indifference or preference thresholds linked to the scale.

13.2.2 The Tools for the Approach

They tools are dedicated to the given synthesis, and so are described below in succession for each synthesis. The grounds for the choices we have made can be found in André [4].

However, the three types of syntheses do share a common trait. Save certain exceptions, the evaluations they aggregate refer to very heterogeneous, not easily commensurable data. Accordingly, resorting to aggregation procedures (save for the operational synthesis in some cases) that strictly limit the possibility of systematically compensating for poor evaluations with good evaluations seemed appropriate. For non-financial performance, such compensations do not seem warranted especially since, with purely ordinal scales, grounding the equivalences implied by the said compensations is hard to do based on clearly understandable considerations. Actually, this is only feasible at the cost of recoding the scales that - more often than not - entails the loss of significance of the degrees.

13.2.2.1 The Operational Synthesis

At the lowest level (level 1), EPISSURE involves indicators, called elementary indicators whose purpose for each of the actions (sponsorship projects in this case) is to describe an aspect, characteristic, factor, expected consequence, etc. that ought to be taken into consideration in the evaluation process. The basic data, and often the indicators per se, exist within the organisation (the company or association in charge of the sponsorship project). In the operational synthesis, the point is to group them per point of view so as to define a coherent family of criteria [33]. In the case of sponsorship (see 13.3.2), each criterion c was defined as the synthesis of elementary indicators with a common scale (not necessarily the same from one criterion to the next). The evaluation of a project did not have any quantitative significance for any of the elementary indicators. When the evaluation was not naturally verbal and was in other words numeric, the numbers only served to reflect an order of preference. In these conditions, defining what we call the **partial evaluation of a project on one criterion**consisted in aggregating elementary verbal evaluations that were all located on the same purely ordinal scale (see 13.3.2).

We were able to use an unique tool in each business that is dealt with in this paper. This tool is the **rule-based weighted median** which we are going to describe below. This tool was especially well tailored to the aggregation of purely ordinal data with a view to restricting compensation possibilities (a goal that proved appropriate to the studied situations).

Let us denote $g_c(a)$ the partial evaluation of project a on any one of the criteria c that the operational synthesis has to build. Let $h_1,...,h_k,...,h_{H_c}$ be the set of elementary indicators that the synthesis must take into account. Evaluations $h_k(a)$ are on the same purely ordinal scale defined by the ordinate series $E_1,...,E_j,...,E_m$ of the set of degrees it comprises with E_1 designating the best and E_m the worst. For the application presented in this paper, the selected scale did not depend on the criteria c, it was defined (see 13.3.2.2) by 16 degrees reflecting the verbal description of a position according to a set goal. To define the median of the evaluations of a on the relevant H_c elementary indicators, the evaluations $h_k(a)$ had to be arranged in an order starting from the best to the worst. The degree located in the middle of the resulting ordinate series $s(a)$ is, by definition, the median of the evaluations.

If one wants to differentiate the role that the elementary indicators should play in calculating the median, a set of weights characterising the relative importance to attribute to each indicator may be taken into account in the manner explained below. Let P_k be the weight attributed to indicator h_k Here, the weight must be defined by an integer (a high number if indicator h_k is to play a more preponderant role in the synthesis). Let $S(a)$ be the series of degrees deducted from $s(a)$ by having degree $h_k(a)$ come into play not once but P_k times. The weighted median is defined by the degree that is located in the middle of the ordinate series $S(a)$.

The partial evaluation of the project a on criterion c may be defined by median $M(a)$ of the ordinate series $S(a)$. As will be seen in 13.3.4.2, the evaluation of action a according to the point of view that criterion c should reflect may show some "defects", in some cases. The defects may be remedied by introducing rules that take

account of all the evaluations $h_k(a)$ on the elementary indicators and thus replace $M(a)$ with a somewhat different degree $\hat{M}(a)$. The rule base that should be introduced obviously depends on the defect to correct. The type of observed defects was as follows: $M(a)$ could lead to a partial evaluation on criterion c that was judged as too favourable considering that there were poor evaluations in $S(a)$. The said rules involve boundaries B (or milestones) defined as follows:

- If E_m is in $S(a)$, then $g_c(a)$ must not exceed B_m
- if E_{m-1} is in $S(a)$, then $g_c(a)$ must not exceed B_{m-1}
- and so forth
- if E_t is in $S(a)$, then $g_c(a)$ must not exceed B_t

For obvious reasons of coherence, the boundaries must be defined so that $B_m \leq B_{m-1} \leq ... \leq B_t$ The rule-based weighted median is written as follows:

$$g_c(a) = \min\{M(a), B_m(a), B_{m-1}(a), ..., B_t(a)\}$$

where

$$B_j(a) = \begin{cases} B_j \text{ if } E_j \text{ is in } S(a) \\ E_1 \text{ if } E_j \text{ is not in } S(a) \end{cases}$$

In these conditions, the partial evaluation $g_c(a)$ of project a on criterion c is necessarily one of the degrees of the ordinate series $E_1, ..., E_m$ defining the common scale for all the indicators h_k $(k = 1, ..., h_{H_c})$ that have be taken into consideration to build criterion c (see Figure 13.2 for example). As the scale is not necessarily the same from one criterion to the next, strictly speaking the degrees should be written $E_1^c, ..., E_m^c$. Up to now, exponent c was omitted for simpler notations. From now on, we will write E^c for the scale.

Evaluation scale of the criteria												
	E_1							$B_{..}$	$B_{..}$	$B_{..}$	B_2	B_1
	E_2						$B_{..}$					
	...					B_{j-1}						
	...											
	E_i				B_j							
	E_{i+1}			B_{j+1}								
	...		$B_{..}$									
	...	B_m										
	...											
	...											
	E_m											
		E_m	...	E_{j+1}	E_j	E_{j-1}	...	...	...	...	E_2	E_1
		Evaluation scale of the elementary indicators										

Fig. 13.2 Example of the Rule Base

13.2.2.2 The Tools for the Intermediate Synthesis

The purpose of this synthesis is to provide the means of appreciating the performance of each project a by associating what we call a **comprehensive evaluation** $g(a)$ with each project. Compliant with the second principle (see 13.2.1), $g(a)$ must be a verbally described degree on a necessarily purely ordinal scale $\boldsymbol{E}$. The comprehensive evaluation $g(a)$ must be regarded as a synthesis of partial evaluations $g_c(a), \forall c \in F$. To build the synthesis, EPISSURE implements a multi-criteria aggregation procedure. We felt that is was requisite for the aggregation procedure to take into account the following points.

i) The purely ordinal character of the scales E^c on which the partial evaluations $g_c(a)$ are located.
ii) The fact that two partial evaluations on different criteria may, even if their verbal description is the same, concern totally heterogeneous aspects of the actual project (in other words, non-commensurable aspects).
iii) The virtual impossibility of grounding rules that systematically compensate for poor partial evaluations with good partial evaluations, on clearly understandable bases.
iv) The fact that two projects with very similar partial evaluations on criterion c may not signify a preference in favour of the project with the best evaluation of the two.
v) The need to differentiate the importance of the role played by the different criteria in determining comprehensive project evaluation.

The ELECTRE methods [15] (chapter 5) and (see also chapter 3 of this book) take these requirements into account very well. As shown below, the ELECTRE TRI method is highly suitable (although it was not designed with this object in mind) for defining the comprehensive aggregation $g(a)$ based on partial aggregations, by taking into account the five above-mentioned requirements. That is why, we choose ELECTRE TRI method for the intermediate synthesis.

ELECTRE TRI [15] (chapter 3) and [36] was designed to assign actions (in this paper, we will still refer to them as projects) to pre-defined ordinate categories. In ELECTRE TRI, each project a is characterised by partial evaluations $g_c(a)$ on a family F of criteria, where the evaluations are on purely ordinal scales E^c. To each of these scales, indifference and preference thresholds could be associated. The purpose of these thresholds is to take into account the relative meaningfulness of the preference gap separating two different evaluations (see iv above). The criteria with their thresholds are usually called **pseudo-criteria**. For a precise definition of the thresholds and of the pseudo-criterion tool, see Roy and Bouyssou [36] and Roy and Vincke [37]. The importance of the role played by the criteria (see v above) may be differentiated to define the category to which a project must be assigned. To do so, a weight and a possible veto power is linked to each criterion. The greater the weight of a criterion, the greater the role that the criterion plays in defining the category of assignment. The veto makes it possible to block the assignment to certain categories due to the partial evaluation of the project on the relevant criterion.

Applying the ELECTRE TRI method, EPISSURE used the same scale **E** to assign a category to the relevant projects. Scale $\boldsymbol{E}$ (selected for the application in this paper) was defined (see 13.3.2.2) by the following degrees:

$\boldsymbol{E}_1$: projects largely exceeding set goals
$\boldsymbol{E}_2$: projects reaching or slightly exceeding set goals
$\boldsymbol{E}_3$: projects falling short of goals
$\boldsymbol{E}_4$: projects falling far short of goals

In ELECTRE TRI, categories are defined based on what is called their limit profiles. By definition, the limit profiles are projects that border on two consecutive categories. Such a project is defined by all the partial evaluations that are justifiably assigned to the project so that it effectively characterises the border. The project defined by the boundary between two categories is not only the lower limit profile of the best of these two categories but also the upper limit profile of the worse of these two categories (see Figure 13.3).

To assign a project to a category, ELECTRE TRI compares its partial evaluations with limit profile evaluations criterion by criterion. Based on the result of the comparisons, a project is assigned to one category rather than to another, i.e., this project is evaluated by one degree rather than another. We would like to point out that ELECTRE TRI has two possible assignment procedures. For EPISSURE, we have used the so-called pseudo conjunctive procedure. With the procedure, a project a can only be assigned to a category $\boldsymbol{E}_h$ if a sufficient majority of criteria (coalition of weights is high enough) evaluate project a at least as well as the lower limit profile of $\boldsymbol{E}_h$ and providing that none of the criteria (which do not evaluate the project as well) veto the assignment to the category (for a complete description and justification of this assignment procedure, see above-mentioned references).

13.2.2.3 Managerial Synthesis Tools

The purpose of the synthesis is to provide an overview of the performance of all the projects, to the general management of the organisation (level 4 of Figure 13.1). The main aim of the overview [13] (p. 44) and [1] (p. 314) is to enable the managers to:

- Identify the widest gaps compared to the set goals and to find the origin of the gaps.
- Communicate on the performance of the various projects and highlight its main traits, within the organisation.
- Identify the projects whose follow-up may pose a problem and help select new projects, if appropriate.

EPISSURE does not suggest any mathematical tool for the overview. The only suggested tools are graphics. The type of graphic components that could be selected may draw on tools such as Dashboards [1] (p. 314) or Balanced Scorecards [21, 22]. Obviously, the components depend on the context, and specifically on the reporting

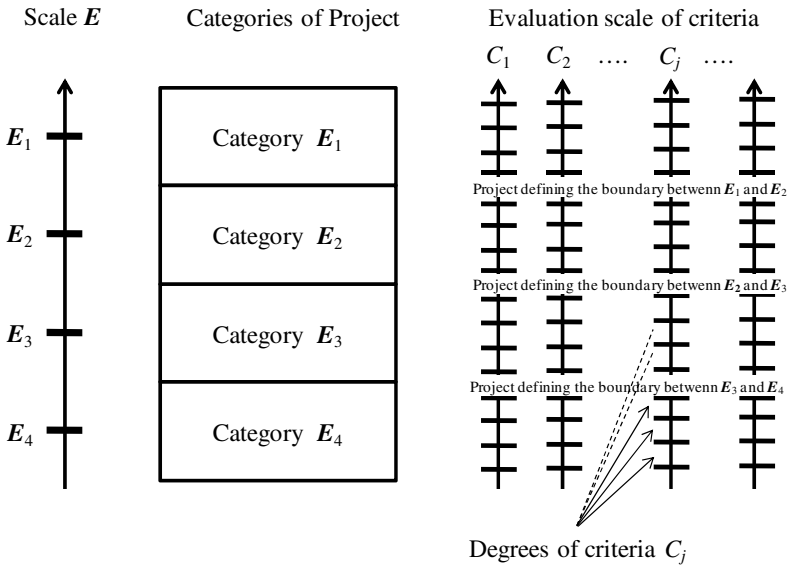

Fig. 13.3 Links between scale E, categories of projects and degrees of criteria

practices in the relevant organisation. Thanks to the hierarchical structure of the approach (principle no. 1), highlighting the characteristics and factors with the greatest impact on overall project performance is easily done. With a computer tool, it is easy to click on the graphics and scroll down from comprehensive performance to partial performance on any given criterion, and from the criterion to elementary evaluations on the various indicators taken into account.

13.2.3 The Framed Dialogue[1] *Process*

The reasons that prompted us to introduce a dialogue process as an integral part of the EPISSURE approach are not discussed in this paper; please refer to Andre [4] and Andre and Roy [38]. As any other dialogue approach, the objective is that the different stakeholders involved in the evaluation reach a common vision. We deliberately chose framed dialogue because we considered that the framework of the shared vision had to be defined by the company. Actually for the business, performance evaluation is not a goal per se; it must serve the aim of creating value for all the stakeholders (shareholders, employees, environment and civil society). Consequently, performance evaluation must lead to an operational evaluation system for performance management. Framed dialogue is defined as a structured dialogue pro-

[1] The word dialogue is used as a translation of the French word "Concertation" for which there is no equivalent in English.

cess which take place within a delimited space called framework. This framework is built on:

- The two normative design principles (see 13.2.1).
- The synthesis tools selected for each of the three levels (see 13.2.2).
- The identified stakeholders involved in the evaluation (see 13.2.4.1).

In these conditions, dialogue must address the components that still need to be defined so that the synthesis tools become operational within this framework. The components involve what is called Characteristic to be Identified by Dialogue (CID). It should be pointed out that a group of experts could have identified all the components (without any dialogue). To provide readers immediately with a quick overview of what CID comprise, we have broken them down per major type; we list what the characteristics pertain to in each type (content or the value to attribute to certain parameters), for each of the relevant synthesis:

- The CID pertaining to performance definition:
  - Operational synthesis: The elementary indicators linked to the criteria.
  - Intermediate synthesis: The criteria.
  - Managerial synthesis: The graphic representation of the managerial synthesis.
- The CID contextualise the performance components and are instrumental in drawing up the different synthesis:
  - Operational synthesis: Relative importance of the elementary indicators and the rules involved in the aggregation.
  - Intermediate Synthesis: Relative importance of the criteria, information pertaining to the modelling of poor knowledge and veto thresholds.
  - Managerial synthesis: No generalisation is possible, but characteristics have to be determined according to the selected mode of graphic display.
- The CID determining performance goals.

We would like to point out that dialogue and framing may be perceived as conflicting. The designed framework was defined accurately enough so that the company could pursue its goals of value creation while leaving enough leeway for real discussions between the stakeholders. Consequently, the set-up of framed dialogue requires several precautions that we detail during the different set-up stages of the EPISSURE approach.

13.2.4 The Set-up Stages of the EPISSURE Approach

The set-up of the EPISSURE approach must be tailored to each evaluation context to become effectively operational. There are three set-up stages in the following order:

- Stage no. 1: Defining the components of the evaluation framework and tailoring the synthesis tools.
- Stage no. 2: Dialogue phase per se for identifying CID content or the value of their parameters, depending on the case.
- Stage no. 3: Dialogue phase per se for identifying CID content or the value of their parameters, depending on the case.

To ensure the coherence of the approach, it is advisable that the same individual, called a facilitator (consult Maystre and Bollinger [28] for additional information), conduct all the stages. His/her function is to implement the synthesis tools properly as well as facilitate dialogue:

- Present the approach so that it is understood by all.
- Facilitate discussions between the stakeholders while checking that each stakeholder takes the floor to express his/her point of view and goals, and that he/she is heard [36].

The neutrality of the facilitator is important to forestall any stakeholder from challenging the process [10, 24].

The detailed content of the stages depends on the evaluation context. This is discussed in Section 13.3. Nevertheless, the main components of stages 1 and 2 can now be detailed.

13.2.4.1 Stage no. 1

During the first step (stage no. 1.1), the purpose is to defined the framework for dialogue. Company appointed stakeholders (for instance, the organisational managers or performance control managers) with support from the facilitator construct the framework. These stakeholders form the steering group. After a phase where the approach is presented and clarified (specifically what the CID encompass and their *raison d'etre* (see 13.2.3), the group details the following items:

- Evaluation goal and the decisions the evaluation will aid.
- The four hierarchical levels of the evaluation.
- The ordinal verbal scales of the evaluations matching each evaluation level.
- The stakeholders that should be included in the dialogue; the latter are organised into what is called the working group (WG).
- The components that still need to be identified so that the synthesis tools are operational, i.e., the components related to the Characteristics to be Identified by Dialogue (CID).

During the second step (stage no. 1.2), the facilitator tailors the synthesis tools to the framework that was designed earlier (see 13.3.3 - stage no. 1.2). This includes:

- Selecting the synthesis tool that is the most appropriate for the operational synthesis.
- Tailoring ELECTRE TRI to the corporate evaluation context.
- Suggesting a graphics display for the managerial synthesis.

13.2.4.2 Stage no. 2

This is the dialogue phase per se. It should be organised within a working group whose content is decided during stage no. 1. The purpose of this stage is: one, to set-up the framework and two, foster a constructive discussion within the working group. This stage is incremental and iterative for each CID:

- Incremental: the target of each working group meeting (2 to 3 hours) is to reach an agreement on the content of an CID or on the value of its parameters. The definition of criteria *c* of *F* (see 13.2.2.1) is usually the target of the first meeting.
- Iterative: each meeting begins with the review and possible amendment of an CID that was on the agenda of the previous meeting. EPISSURE provides for the possibility of going over the work of the previous meeting. The iterative method may seem risky as it might question the progress made at the earlier WG meeting. However, we will see that this option fosters the construction of a common vision where each stakeholder feels that he/she is part of the process (see 13.4). This point corroborates what Joerin and Rondier [20] (p. 19) established about cognitive learning processes during which it is common to review an earlier stage to modify, polish, or complete results.

In conclusion to this Section, we would like to draw the reader's attention to two special features of EPISSURE:

- Dialogue only pertains to concrete proposals and not to general evaluation principles.
- The point is not to incorporate alternatives but to build a common vision for all the actors involved in performance evaluation. For instance, the working group members have to agree on one set of values to assign to the weights (see 13.2.2.1 and 13.2.2.2) and not on several sets according to the different value systems of the members [6]. We will see that this deliberate choice did not cause any particular problems.

Here we conclude the general presentation of the EPISSURE approach that proposes an innovative approach in an attempt to provide answers to the issue of broadening the value concept. In the next Section, we describe a series of operational implementations with a view to evaluating the sponsorship project.

13.3 Testing and Implementing EPISSURE in Companies, for Sponsorship Projects

Thanks to *IMS- Entreprendre pour la Cité* (an institute for sponsorship and solidarity)[2], we were able to contact about ten companies likely to be interested in

[2] Created in 1986, *IMS-Entreprendre pour la Cité* is the backbone for a network of 200 firms. Its purpose is to help the businesses integrate innovative societal commitment approaches into their

testing and possibly implementing EPISSURE to evaluate sponsorship project performance. After a first series of contacts, four companies were selected. The four companies suggested testing seven different sponsorship projects. In this Section, we first present the companies. Then, we describe the set-up stages of the EPISSURE approach.

13.3.1 Presentation of the Selected Companies

The four selected corporate foundations were:

- The Petzl Foundation[3] (Grenoble) with a project for a mountain school in Nepal
- The Kronenbourg Foundation[4] (Strasbourg) with two projects:
  - A project for a restaurant promoting social integration.
  - A project with the association *Jardin de la montagne verte* (green mountain garden), where fruit and vegetable crops are grown by adults with disabilities.
- The Olympique Lyonnais / Cegid Foundation[5] (Lyons) with two projects:
  - The Immersion project whose goal is to enable young people with disabilities to discover the corporate world.
  - The Doctor Clown project providing entertainment for sick children in hospital.
- The RATP Foundation[6] (Paris) with two projects:
  - The *T'as trouvé un job* (you've found a job) project for the professional integration of young people.
  - The *Rencontre sur Tatamis* (the tatami mat encounters) project teaching civic education through sports to young people.

An association (which was different for each project) ran each of the seven projects and the companies were project partners. Accordingly, the companies were seeking an evaluation approach tailored to their evaluation contexts:

- The evaluation required taking into account very heterogeneous data (see 13.2.2.2 ii).
- More often than not, the basic data was solely qualitative. Although the data underwent a quantitative evaluation, a great deal of the data was only meaningful for ordinal comparisons (see 13.2.2.2 i).

social responsibility policy, approaches that meet their development challenges as well as society's expectations.

[3] http://fr.petzl.com/petzl/frontoffice/static/fondation/fondation-petzl.htm

[4] http://www.brasseries-kronenbourg.com/_corporate/entreprise_engagee

[5] http://www.olweb.fr/index.php?lng=fr&pid=910101

[6] http://www.fondation-ratp.fr

- The data often also contained a share of arbitrariness, thus small gaps were devoid of meaning (see 13.2.2.2 iv).
- The favourable data for a good evaluation did not have to compensate systematically for the data that was unfavourable to a good evaluation (see 13.2.2.2 iii).

Building synthesis indicators based on dialogue was a first for the different stakeholders and fuelled huge expectations.

We would like to remark that the goals of the different evaluated projects covered very different fields, i.e., national vs. international, young people vs. adults, business vs. civil society. The companies sponsoring the projects were also of different sizes (ranging from small enterprises to major groups) and worked in different businesses. Last, it should be underscored that some associations had poor knowledge of evaluation practices and were worried that the topic might jeopardise their autonomy unlike others associations that already had a solid experience of evaluation practices.

13.3.2 Framing the Dialogue Approach (Stage no. 1.1)

The stage was conducted by a group of stakeholders called the "steering group" (see 13.2.4.1). In the cases studied in this paper, the groups included the head of the foundation, an IMS representative and a facilitator (see 13.2.4). They had to define (see 13.2.4.1):

- The evaluation goals and the decisions they would aid.
- The four hierarchical levels of the evaluation.
- The ordinal verbal scales of evaluations matching each evaluation level.
- The stakeholders that should be included in the dialogue, i.e., the composition of the working group (WG).
- The components pertaining to Characteristics to be Identified by Dialogue (CID).

The framing stage was special to each company. However, the features common to each company can be presented.

13.3.2.1 Evaluation Goal

The evaluation had to enlighten the foundation's decision (usually by the Board on a yearly basis) on project follow-up: continue the investment in a project conditionally or unconditionally, or stop the investment. The evaluation did not only consist in judging whether a project was good or bad but in evaluating whether the project had reached the set goals. In the different cases, the companies had invested in the projects with a view to results: general interest for society and the company's own interests ("sponsorship is no longer an act of pure philanthropy, but a form of investment that business legitimately expects to produce a positive return"). Consequently, the set goals were impact goals and not resource-based goals. Although this

may seem a standard discourse in the corporate world, it seemed quite new within the context of the cases we studied. Actually, although the idea of a goal was in the air, the terms of the goals were not always clearly identified within the companies and associations that we met.

13.3.2.2 The Hierarchical Levels of the Evaluation

The following levels were used for all the corporate cases presented in this paper (see Figure 13.4):

- Level 4 corresponded to the foundation. At this level, an overview of the different projects funded by the company had to be feasible.
- Level 3 was selected as the level for each project. At this level, the approach had to situate a project vis-à-vis its goals.
- Level 2 was selected as the intermediate evaluation level of a project, i.e., what we called the success criteria. The criteria were defined as the ex-post requisites according to which the stakeholders involved in the project would consider it successful. The methodology for identifying the success criteria was detailed during stage no. 2 (see 13.3.4).
- Last, Level 1 was defined as the elementary indicator level.

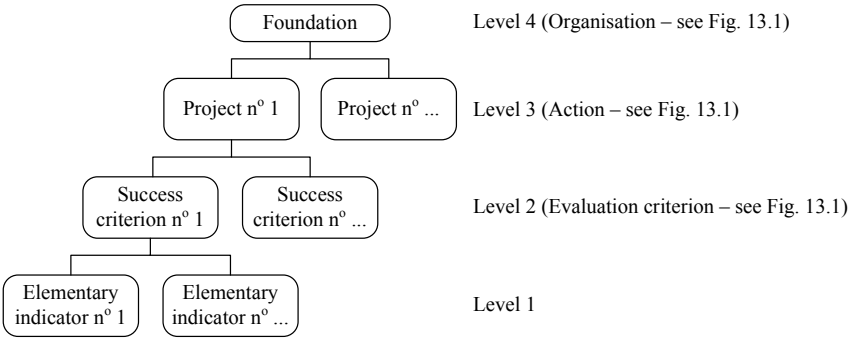

Fig. 13.4 The 4 hierarchical levels of evaluation

13.3.2.3 Ordinal Verbal Scales

The same scales were selected for all the corporate cases. For each scale, the principle consisted in taking a position compared to the goals (the above-mentioned evaluation goal).

For Level 1, a scale, which was the same for all the elementary indicators, was a 16 degree-scale arranged into 8 main degrees where each degree was broken down

into two ratings "high" and "low" (see Figure 13.5). This scale is also to be found automatically in level 2, because of the way the operational synthesis was defined (rule base weighted median). Consequently, level 2 is not mentioned below.

E_1	Goal very largely exceeded	High
E_2		Low
E_3	Goal largely exceeded	High
E_4		Low
E_5	Goal exceeded	High
E_6		Low
E_7	Goal reached	High
E_8		Low
E_9	Goal partially reached	High
E_{10}		Low
E_{11}	Goal very partially reached	High
E_{12}		Low
E_{13}	Goal not reached	High
E_{14}		Low
E_{15}	Goal that the project did not attempt to reach	High
E_{16}		Low

Fig. 13.5 Scale linked to levels 1 and 2

On the basis of the initial evaluations on elementary indicators, three reference situations will have to be defined:

- Satisfactory (level E_8) i.e., the value indicating that the stakeholders estimated that the project had reached its goals based on this one indicator (identifying goals was part of the dialogue phase, see 13.3.4).
- Ideal (level E_1) i.e., the value indicating that the stakeholders estimated that the project could not go any further, based on this one indicator.
- Floor value (level E_{16}) i.e., the value indicating that the stakeholders estimated that the project had not even attempted to reach its goal, based on this one indicator.

The detail of this correspondence can be found in 13.3.4.3.

Based on the same principle again, a levelled scale for level 3 was considered fine enough to evaluate the projects. The levels were defined as indicated in Figure 13.6.

At level 4, the selected graphic representation (see 13.3.3.3) took into account two ordinal scales jointly: the level 3 scale and another scale reflecting the size of the company's investment in the project. Investment size is not only financial as it

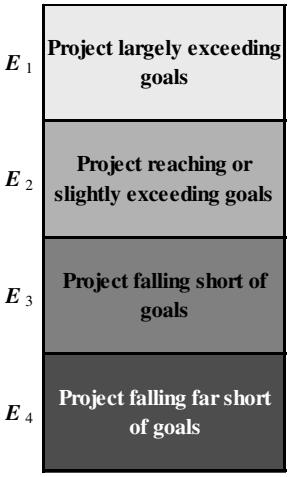

Fig. 13.6 Ordinal scale of evaluation for sponsorship projects

comprises other types of investment (for instance, number of employees involved in the project). The scale had two levels:

- Low Investment.
- High Investment.

13.3.2.4 Composition of the Working Group

Each working group had to count members which are legitimate and representative of the different goals of sponsorship projects. Therefore, the number of members of working group depends on each evaluation context and each of these members had to have leeway over their organization. The composition of the working groups was different in each company, ranging from 4 people (at Petzl) to 12 (at Kronenbourg). However, aside from the facilitator, each group included stakeholders outside the company (mainly from the association in charge of the project) and in-house players.

13.3.2.5 Elements Linked to the Characteristics to be Identified by Dialogue (CID)

So as to incorporate the different components of the evaluation context of a sponsorship action (emerging, not well structured issue, association and corporate stakeholders are not well acquainted, and so on, see 13.3.1), we suggested a mainly ascending process broken down into three sub-stages (see , and):

- Stage no. 2.1 - Identification, formalisation, and ranking of the success criteria (Level 2) corresponding to the identification of the parameters required for the Intermediate Synthesis. The CID at this stage were:
  - The success criteria.
  - The weights of the criteria.
  - The thresholds of discrimination (indifference, preference) and the veto possibilities.
- Stage no. 2.2 - Identification of the elementary indicators (Level 3) linked to the success criteria selected during stage no. 2.1. The CID at this stage were:
  - The elementary indicators.
  - The weights of the elementary indicators.
  - The rule base.
- Stage no. 2.3 - Identification of the goals linked to the elementary indicators. The CID at this stage were performance goals, related to the definition of the reference situations (see 13.3.2.3).
  - Satisfactory.
  - Ideal.
  - Floor value.

For each stage, the dialogue was focused on the CID (Characteristics to be Identified by Dialogue). For the presentation of the EPISSURE process, we would like to underscore the incremental and iterative approach (see 13.2.4.2). The incremental approach was exemplified by the fact that the CID were discussed synthesis operation per synthesis operation. The CID at stage no. 2.1 corresponded to the Intermediate Synthesis (IS) while the CID at stages no. 2.2 and no. 2.3 corresponded to the Operational Synthesis (OS). The iterative approach was exemplified by the fact that each working group meeting was an opportunity to rework points that had been examined earlier (see 13.3.4 for the conditions). Thus defined, the dialogue process required about 4 meetings with the working group. Figure 13.7 below shows how the dialogue principle unfolded.

Choosing the iterative approach was done deliberately to leave room for discussion, enable the issues to mature, and allow the stakeholders to appropriate the approach (evaluation was a new topic for numerous stakeholders - see 13.3.1). We describe the process in detail further down (see 13.3.4).

13.3.3 Tailoring the Synthesis Tools (Stage no. 1.2)

The purpose of this stage is to allow the facilitator to tailor the synthesis tool to each business.

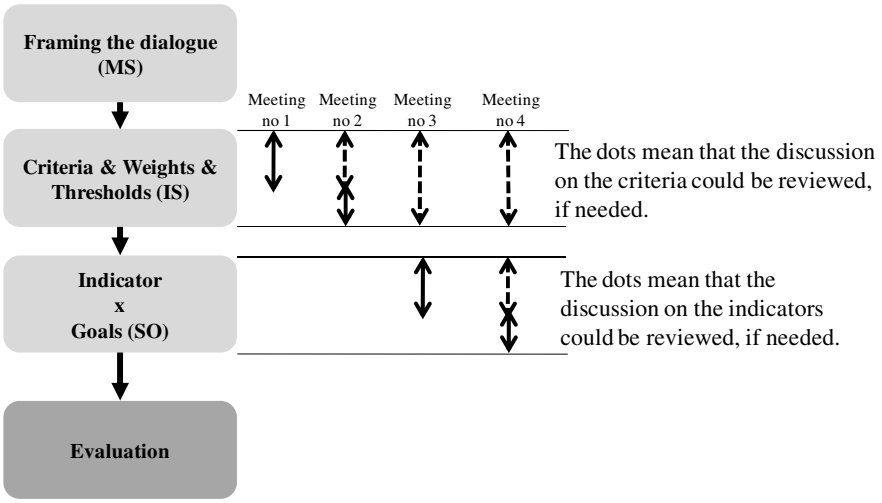

Fig. 13.7 Principle of the dialogue approach

13.3.3.1 Operational Synthesis

The necessary adjustments concern both the way the ordinal 16 levelled scale was used and the rule base adjustment.

a) Positioning Principle for the Evaluation of the Elementary Indicators on the Ordinal Scale

We have to keep in mind that the elementary indicators already existed. Adaptation consisted to link the way the project was previously evaluated to its position on the 16 degrees verbal scale. For the elementary numeric indicators, each degree was demarcated by a high limit and a low limit. For an indicator of growing satisfaction, the value of the low limit corresponded to the value that had to be determined during the dialogue process (see 13.3.4.3; stage no. 2.3). By definition, the high limit corresponded to the low limit of the degree above. In conclusion, project *a* was positioned on a degree if its evaluation ranged within the two limits demarcating the said degree. For elementary indicators on a verbal scale, positioning was done by hand during the evaluation phase.

b) Adjustment of the Rule Base

Based on the principles defined for the rule base (see 13.2.2), the adjustment provided a default rule base (see Figure 13.8), which comprised two rules that fully identify the base:

- If, on all the elementary indicators linked to the criterion, the weakest evaluation of project a ranged between E_1 and E_4, then the value of project a evaluation on the success criterion was not marked with a boundary.
- In the other cases, if the weakest evaluation was equal to degree E, on all elementary indicators linked to the success criterion, then the value of the criterion evaluation was at best equal to degree E_{i-3}.

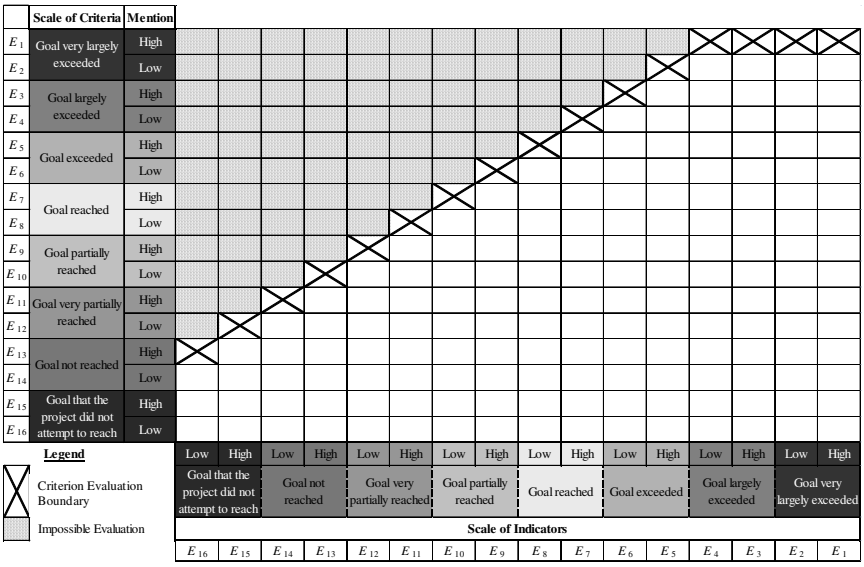

Fig. 13.8 Default Rule Base

This leads to adopt $B_j = E_{j-3}$ for $j > 4$ (see 13.2.2.1).

13.3.3.2 Intermediate Synthesis: ELECTRE TRI Implementation

In this paragraph, we present the outline of how ELECTRE TRI was adjusted. The data taken into account by the synthesis tool were the evaluation degrees linked to each success criterion during a given period (written j; usually the year) for the project (a_j, *which will be simplify as a*). Based on this data, the tool was used to identify the most appropriate degree to which each project was properly assigned.

Four limit profiles were defined for the synthesis of the success criteria. The profiles set the boundary limits for each category:

- "Project largely exceeding its goals" ($\boldsymbol{E}_1$)
- "Project reaching or slightly exceeding goals" ($\boldsymbol{E}_2$)
- "Project falling short of goals" ($\boldsymbol{E}_3$)
- "Project falling far short of goals" ($\boldsymbol{E}_4$)

The identification of the profiles was based on a simple logic. Each category included 4 degrees on the scale used for the success criteria (see Figure 13.9).

Limit Profile of category no. 1	**Project largely exceeding goals**	E_1
		E_2
		E_3
		E_4
Limit Profile of category no. 2	**Project reaching or slightly exceeding goals**	E_5
		E_6
		E_7
		E_8
Limit Profile of category no. 3	**Project falling short of goals**	E_9
		E_{10}
		E_{11}
		E_{12}
Limit Profile of category no. 4	**Project falling far short of goals**	E_{13}
		E_{14}
		E_{15}
		E_{16}

Fig. 13.9 Profile Limits

The different discrimination thresholds were set as follows:

- The value of the indifference threshold set at one degree
- The value of the weak preference threshold set at three degrees

In specific cases and for some criteria, these two thresholds could be choosen as equal to zero. The veto power (when chosen for a criteria) was set to prevent the evaluation on the degrees:

- "Project largely exceeding goals" ($\boldsymbol{E}_1$), if one criterion was evaluated between degrees E_{13} and E_{16}
- "Project reaching or slightly exceeding goals" ($\boldsymbol{E}_2$), if one criterion was evaluated at degrees E_{15} or E_{16}
- No veto threshold prevented the evaluation on the last two degrees: "Project falling short of goals" ($\boldsymbol{E}_3$) and "Project falling far short of goals" ($\boldsymbol{E}_4$).

13.3.3.3 Managerial Synthesis

The adjustment consisted in designing a graphic display providing an overview of the performance of the projects funded by a company (or corporate foundation). For the display, particular care had to be paid to the graphic components of the synthesis overview: shape, type, size, and colour of the graphics. The graphic choices

depended on context, and specifically on the reporting practices of the relevant organisations. The care paid to the graphics was very similar to the one involved in Dashboard designs [1] (p. 314).

The figure consisted in crossing the result of the Intermediate Synthesis,- with the size of the investment in each project (see 13.3.2.2), thus making it possible to identify the number of projects crossing both axes. The company assessed the evaluation of the size of the investment. Figure 13.10 below provides an example of the display.

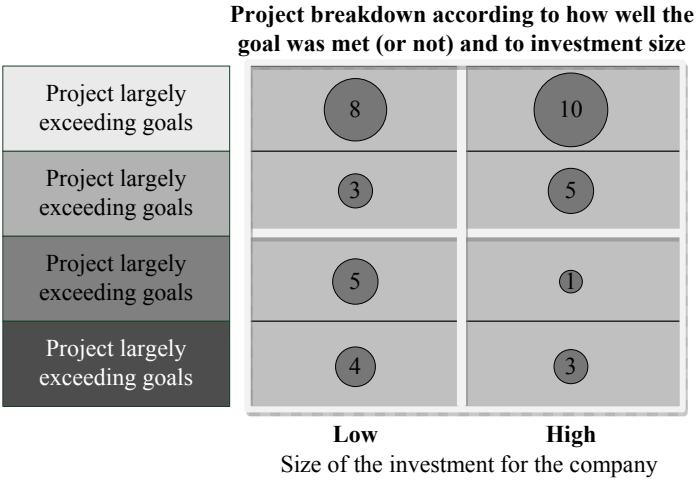

Fig. 13.10 Example of a display of managerial synthesis results

The purpose of the display was to enlighten corporate or foundation authorities' decision by allowing them to identify the projects that needed reviewing (for instance, projects for which the company had made substantial investments and that clearly did not reach their goals). Once the projects were identified, the criteria basing the evaluation and then the related elementary indicators could be detailed thanks to hierarchical interleaving. Consequently, the stakeholders had all the facts for managing the performance of the sponsorship projects funded by the company.

13.3.4 Dialogue Sequence (Stage no. 2)

13.3.4.1 Identification, Formalisation and Ranking of the Success Criteria (Stage no. 2.1)

The CID (Characteristic to be Identified by Dialogue) at this stage cleared the way for implementing the Intermediate Synthesis. The CID corresponded to the parameters required for ELECTRE TRI implementation (see 13.2.3 and 13.3.2.3):

- IS (Intermediate Synthesis) - CID no. 1: the success criteria
- IS - CID no. 2: the weights of the criteria
- IS - CID no. 3: the veto powers

The discussions were usually held during the first two meetings of the working group (WG).

a) Success criteria (IS - CID no.1)

The identification of the success criteria (Level 2) was on the agenda at the first WG meeting. After a general presentation of the EPISSURE approach, the facilitator organised the dialogue. He/she first asked all the members of the working group the following question, "From your standpoint, and only from your standpoint, what are your success criteria for the project?" In other words, "From your standpoint, what would prompt you to say, at project completion, that it was a success?" The purpose of taking a position solely from the standpoint of one's role was to enable each stakeholder to express his/her point of view as well as understand the other stakeholders' points of view.

Each participant was given time to think about the question. During this time, the participants had to jot down their ideas on post-its (a legibly written idea on a post-it), which were in different colours according to origin (foundation, company, association in charge of the project, and others).

Once they had thought it over, each participant presented his/her success criteria one by one. The facilitator grouped the criteria into broad categories. The only questions that were allowed during the presentation phase were questions to clarify understanding.

Once every idea was expressed, they were regrouped or even reformulated. At this stage of the discussion, each participant could express a broader point of view than his/her own position. The recommended reformulation was to use an action verb to express each success criterion. For instance, the project was a success if it managed to "strengthen young people's ties with the corporate world", or "improve the behaviours of young beneficiaries." This type of formulation might seem somewhat restrictive but it allowed each stakeholder to take a position within an impact evaluation logic and not within a resource evaluation logic, which is linked to the evaluation goal (see 13.3.2.1).

During this phase, participants could express their opinions freely. The facilitator made sure that everyone spoke. At the end of the phase, each working group in the different corporate cases managed to set down the terms of the first version of the success criteria. The version changed in the course of the discussions.

b) Weights of the Criteria (IS - CID no. 2)

Before the discussion got started, the facilitator reminded the participants of the significance of the weights in terms of veto power. The purpose of the first question

was to find out if the stakeholders conferred the same importance to all the success criteria.

If the answer was no, a discussion was held to identify the weights, using J. Simos's revised card method [16]. To get the discussion going, the facilitator provided one white card per criterion. The facilitator then asked the stakeholders to rank the criteria cards by order of importance, explaining that criteria could be tied, and that one or several white cards could be inserted in the ranking order. He/she went on to explain that the purpose of the white cards was to reflect the size of the gap separating the relative importance assigned to two consecutive criteria in the ranking order.

By default, all the success criteria were put on the same level. Once an agreement had been reached about ranking, the facilitator asked the different stakeholders the following question so that they would identify a range to indicate the ratio between the most important and the least important criterion, "Indicate an interval for the ratio between the weight of criterion (name of the most important criterion) and the weight of criterion (name of the least important criterion)."

During this phase, the only condition laid down by the facilitator was to reach an agreement on the sets of weights. Actually keeping several sets of weights was not envisaged. This choice can be explained by the evaluation goal (see 13.3.2.1) that is to enlighten the company's decision on project follow-up. The set of weights was defined using the SRF software [16].

c) Discrimination threshold and veto power linked to Each Criterion (IS - CID no. 3)

In the first two cases, the facilitator suggested that the working group identify: one, the discrimination thresholds (so as to take into account poor knowledge) and two, the veto thresholds (so as to prevent a good evaluation of the project if one of the criteria received a bad evaluation). Both discussions (on discrimination and veto thresholds) were quite long, because the stakeholders had trouble understanding the different underlying these two notions. Grouping in a same CID of the discrimination threshold and veto power (which do not have the same function) may have facilitated comprehension. Once this was understood, the actual discussion dealt with giving a veto power to different success criteria, or not. In the third case, default values were set for the discrimination thresholds (see 13.3.3.2) and the discussion only dealt with whether or not to give a veto to the different criteria.

During the discussions, the stakeholders often questioned about the criteria of success (definition and/or formulation) and of the weights.

13.3.4.2 Link between Elementary Indicators and Success Criteria (Stage no. 2.2)

The CID at this stage enabled the implementation of the operational synthesis via the rule-based weighted median. The Characteristics to be Identified by Dialogue corresponded to the choice of the elementary indicators that had to be considered for each criterion, for project evaluation. They also corresponded to the values to attribute to the parameters required for method implementation:

- OS (Operational synthesis) - CID no. 1: the elementary indicators
- OS - CID no. 2: the weights of the elementary indicators
- OS - CID no. 3: the rule base

These discussions were often held during the third WG meeting.

a) Elementary Indicators (OS - CID no. 1)

In the course of the first discussions, the associations shared their fear of increased reporting assignments. Although they had understood the advantage of the suggested evaluation approach, they did not want to spend too much time on reporting. From the second corporate case, the facilitator noticed that, although the discussions on the elementary indicators would begin based solely on what existed, needs (save the odd exception) were covered. The approach also provided several advantages. First, it cleared the way for addressing the evaluation faster, because defining and setting up new indicators was not required. Two, it provided a base for the discussion on the goals (see 13.3.4.3 - stage no. 2.3).

The facilitator organised the dialogue in the following manner, "On the criterion (name of the criterion), what extant indicators could be used to evaluate whether goals were reached on this success criterion?" The discussion on the elementary indicators often prompted new questions and discussions on the success criteria. The discussions lead to reformulation or often to the merge of two criteria that were felt to be quite similar after analysing the related elementary indicators.

b) Weights of the Elementary Indicators (OS - CID no. 2)

The purpose of the first question was to find out whether all the stakeholders conferred the same importance to all the indicators. In the different cases and unlike the weights of the success criteria (see 13.3.4.1), the stakeholders agreed on equitable weighting. If this hadn't been the case, the facilitator would have used the same method (J. Simos's cards) and the same dialogue process as for the weights of the success criteria (see 13.3.4.1).

If the stakeholders preferred equal weighting, it was to simplify things. They found it complicated to have different types of weights, something which puzzled most of the stakeholders.

c) Rule Base (OS - CID no. 3)

Before starting the discussion, the facilitator used a numeric example to present the principle of the rule-based median. He/she specifically explained that the rules served to limit the evaluation on a criterion in the case where an indicator had a very bad evaluation whereas all the other evaluations were good. At this stage of the discussion, the facilitator asked the WG members to agree on the need to set up the rule base. In every corporate case the answer was yes. The facilitator suggested a default rule base (see 13.3.3.1). In every corporate case, the default base was accepted.

13.3.4.3 Identification of the Goals Linked to the Elementary Indicators (Stage no. 2.3)

The CID at this stage cleared the way for contractualising the performance goals on each elementary indicator. For each elementary indicator, the CID concerned the matching between the way they were first used to evaluate a project and the scale E_1 to E_16 on which they were eventually located. This matching (see 13.3.2.1) was linked to three reference situations (Satisfactory, Ideal and Floor) that set the meaning attributed to degrees E_8, E_1 and E_{16}, linked to the goals.

These discussions were often held during the fourth and last WG meeting. In one corporate case, a fifth meeting was needed. The dialogue process depended on the type of elementary indicator (described further down). The discussions on identifying the goals often prompted new questions and discussions on the elementary indicators or even on the associated criteria. The discussions led to modifying elementary indicators (one indicator added, merge of two indicators, removal of an indicator) or even changing the associated success criterion (reformulation of the name, merge of two criteria that were, in the end, judged to be quite similar).

a) Dialogue Process for an Elementary Indicator (in numeric form, at first)

The facilitator asked the WG members for at least three reference numeric values, by asking the three following questions in the order below:

- According to you, what is the target value to reach within a year for you to consider that, from the sole standpoint of this indicator, the project could be considered a success? (identification of the value setting the meaning of degree E_8).
- According to you, which value would you consider indicates that the project, from the sole standpoint of this indicator, is a success beyond all expectation? (identification of the value setting the meaning of degree E_1).
- According to you, below which value would you consider that the project, from the sole standpoint of this indicator, is a total failure? (identification of the value setting the meaning of degree E_{16}.

For each of these questions, the stakeholders who originally suggested the elementary indicator answered first. Then, an open discussion began until an agreement was reached. On all the elementary indicators in the seven corporate cases, a consensus on these performance goals was rapidly reached. According to the stakeholders, the main reason for this was that the discussion on the performance goals took place at the end of the process. The stakeholders knew each other well and could have constructive discussions. For instance, Figure 13.11 below illustrates the result for a indicator monitoring the number of young people to train.

			No. of graduates benefiting from arrangement	
			Goal	**Eval.**
E_1	Goal very largely exceeded	High	**1300**	
E_2		Low		
E_3	Goal largely exceeded	High		
E_4		Low		
E_5	Goal exceeded	High		
E_6		Low		
E_7	Goal reached	High		
E_8		Low	**800**	
E_9	Goal partially reached	High		
E_{10}		Low		
E_{11}	Goal very partially reached	High		
E_{12}		Low		
E_{13}	Goal not reached	High		
E_{14}		Low		
E_{15}	Goal that the project did not attempt to reach	High		
E_{16}		Low	**600**	

Fig. 13.11 Identification of the performance goals for the elementary indicators

Once an agreement was reached on the three reference values, linear interpolation was used to calculate the values corresponding to the other degrees. The facilitator presented the result and asked the WG members if they agreed with the values calculated for degrees $(E_2, ..., E_7, E_9, E_{10}, ...E_{15})$. Going back to the same example, Figure 13.12 illustrates what the facilitator presented.

In cases where the proposal was not validated by the working group, the stakeholders could choose to set other reference values for other degrees (supplemental to degrees E_1, E_8 and E_{16}). Another linear interpolation calculation was then carried out to set the value of the degrees between the reference values. Based on the same example as before, Figure 13.13 illustrates the case where the value matching degree E_4 was selected as the reference.

			No. of graduates benefiting from arrangement	
			Goal	**Eval.**
E_1	Goal very largely exceeded	High	**1300**	
E_2		Low	1229	
E_3	Goal largely exceeded	High	1157	
E_4		Low	1086	
E_5	Goal exceeded	High	1014	
E_6		Low	943	
E_7	Goal reached	High	871	
E_8		Low	**800**	
E_9	Goal partially reached	High	775	
E_{10}		Low	750	
E_{11}	Goal very partially reached	High	725	
E_{12}		Low	700	
E_{13}	Goal not reached	High	675	
E_{14}		Low	650	
E_{15}	Goal that the project did not attempt to reach	High	625	
E_{16}		Low	**600**	

Fig. 13.12 Calculation of the value of the performance goals based on the three values identified by the stakeholders of the evaluation process

b) Dialogue Process for an Ordinal Verbal Elementary Indicator

The same process was applied except that the different degrees of the scale were described with precise definitions. First, the facilitator asked the WG members to describe precisely the degree matching the relevant reference situations (satisfactory, ideal and floor). Then, the facilitator asked for the description of at least two additional degrees (see illustrations in Figures 13.13 and 13.14).

For instance, figure 13.14 below illustrates an indicator monitoring the set-up of a local organisation.

For each situation, the stakeholders that came up with the elementary verbal indicator were the first to submit a formulation. An open discussion then began until an agreement was reached. For every of the seven corporate cases, a consensus was rapidly reached on the verbal description of the reference situations which contractualised the performance goals.

			No. of graduates benefiting from arrangement	
			Goal	**Eval.**
E_1	Goal very largely exceeded	High	**1300**	
E_2		Low	1233	
E_3	Goal largely exceeded	High	1167	
E_4		Low	1100	
E_5	Goal exceeded	High	1014	
E_6		Low	943	
E_7	Goal reached	High	871	
E_8		Low	**800**	
E_9	Goal partially reached	High	775	
E_{10}		Low	750	
E_{11}	Goal very partially reached	High	725	
E_{12}		Low	700	
E_{13}	Goal not reached	High	675	
E_{14}		Low	650	
E_{15}	Goal that the project did not attempt to reach	High	625	
E_{16}		Low	**600**	

Fig. 13.13 Calculation of the value of the performance goals based on the four values identified by stakeholders of the evaluation process

13.3.5 Validation and Implementation of the EPISSURE approach (Stage no. 3)

The evaluation approach that was built then had to be formally validated by the companies. The foundation authority usually validated the approach. Notably, once EPISSURE was set up, it could be used in a contract situation between the different stakeholders. Although dialogue was critical during the set-up phase, it then had to make way for a conformation tool whose goal was to prompt the stakeholders to adopt the desired type of behaviour [29].

13.4 Observed Results

Out of seven corporate cases EPISSURE approach was stopped for three of them (two Kronenbourg projects ant Petzl project) mainly for lack of follow-up. However, EPISSURE approach is still in place in four of them (RATP and Olympique Lyonnais projects). Several people were trained to act as facilitators within the IMS association and several corporate foundations. Accordingly, implementation has now

			Set-up of a durable local organisation	
			Goal	**Eval.**
E_1	Goal very largely exceeded	High	Financial autonomy of the system	
E_2		Low		
E_3	Goal largely exceeded	High		
E_4		Low		
E_5	Goal exceeded	High	The entire team is in place and operates autonomously	
E_6		Low		
E_7	Goal reached	High		
E_8		Low	The entire team is in place but is not autonomous vis-à-vis the foundation	
E_9	Goal partially reached	High		
E_{10}		Low		
E_{11}	Goal very partially reached	High	A mixes team Népal / Europe is in place	
E_{12}		Low		
E_{13}	Goal not reached	High	A technical manager is appointed	
E_{14}		Low		
E_{15}	Goal that the project did not attempt to reach	High		
E_{16}		Low	No contact, or at the study stage	

Fig. 13.14 Illustration of the evaluation degrees for a verbal indicator, at first

been broadened to include the evaluation of other sponsorship projects, specifically within the RATP foundation. The facts lead us to believe that the advantage of the EPISSURE approach has been validated for the evaluation of sponsorship action performance. Below, we discuss a series of results that were established within the different corporate cases.

a) Contribution of the Approach to Sponsorship Action Evaluation

In sponsorship action evaluation, there is a very strong demand for tailored, relevant tools enabling effective evaluation and providing an answer to companies' standard question, "were we right to invest in this project?" Although it is not expressed as such, the greatest expectation is to have synthesis providing a better understanding of the available information. Actually, the task is tricky not because of any lack of information but because of an overabundance of data and of the multi-criterion aspect of the data (see 13.3.1). The strong expectation is reflected by the fact that the WG members were readily available. They all wanted to reach concrete satisfactory results. We organised several dozen working groups and met diverse stakeholders (a total of fifty different stakeholders; companies, foundations, and associations). In every case, attendance was close to 100% at every single meeting.

Although the working groups were held independently, they led to the identification of success criteria (Level 2) that belong to three families of the same type:

- Social criteria evaluating the social impact (or general interest) of the project (for instance, introducing the corporate world to young people with disabilities).
- Association related criteria, specifically the association's ability to become self-reliant (for instance, reach financial autonomy).
- Corporate related criteria (for instance, strengthen employee motivation, clear the way for recruitment, provide ideas for new products).

The name of the criteria and the related elementary indicators were obviously particular to each company. Interestingly, compared to the literature on sponsorship action evaluation [12, 30, 32, 41], association related criteria are innovations. This undoubtedly bears witness to the fact that the construction of sponsorship action evaluation is still underway.

b) Ability of the Framed Dialogue Process to Foster Genuine Dialogue

The actual problem consisted in finding a process that was serious and reliable enough to reassure evaluation "regulars", but not too "sophisticated" to scare the "novices" (see 13.3.1). The process also had to prompt each stakeholder to become a responsive listener. As a result, virtually all the stakeholders involved in the dialogue process expressed their satisfaction. The only actual problem was addressing the case of an association representative who was unable to attend the first two WG meetings. He was uncooperative when work began on the elementary indicators (see 13.3.4.2 - stage no. 2.2). He was afraid that the company would use the data to lower financial aid. The positive contribution of the framed dialogue process is clear, provided that the process is experienced from beginning to end.

The seven corporate cases highlight that the process fostered individual and collective learning mechanisms contributing to the design of shared solutions. Consequently, we can talk about a cognitive side of the EPISSURE approach. Actually, the agreements on the different CID (Characteristic to be Identified by Dialogue)

were not reached based on pre-existing components, but on components that the WG members constructed together. An instance of the cognitive aspect is the list of success criteria (IS - CID no. 1 - see 13.3.4.1). The first meeting was an opportunity for each participant to list his/her success criteria (solely from his/her point of view; see 13.3.4.1). The list of criteria was then reworked during the entire process (iterative aspect of the approach; see 13.3.4.2 and 13.3.4.3). At the end of the process, in each corporate case, the working groups were able to draw up substantially revised lists of success criteria. The lists were shorter, better defined, and above all were common to all, unlike the lists that had been put together at the start of the first meeting. Several factors facilitated the collective learning process. First, there was the dialogue framework, which was never experienced as being overly restrictive. On the contrary, it fostered discussions within a common space. Because the framework was established during the framing phase of the dialogue process (meaning it was tailored to the special features of the context; see 13.3.2.3), it was well accepted. The iterative and incremental aspects (see 13.3.1) of the EPISSURE approach also fostered ties and trust within the working group. The WG members actually had time to get to know one another. They felt much freer to discuss as the process allowed them to review the result of their discussions at each working group meeting. Last, there was the facilitator; one of his/her assignments was to stimulate a common thought process.

c) Ability of the Synthesis Tools to Take into Account the Special Features of Sponsorship Action Performance

Identifying synthesis indicators was not a foregone undertaking, see Alfsen and Moe [3] (p. 10), Henderson [19], Anielski [5], and Faucheux & Nicolaï [14] to find numerous reserves about synthesis indicators. Nevertheless, the synthesis indicators, which were designed thanks to the selected synthesis tools (see 13.2.2) were very productive for evaluating non-financial performance. These good observed results were possible here thanks to the deliberate choices of the EPISSURE approach (see 13.2.1):

- The hierarchical structure of the synthesis indicators enabled stakeholders to have a good grasp on the components making up the synthesis.
- The use of ordinal scales meant that only the ordinal significance of the evaluation data was taken into account.
- The use of synthesis tools limited compensation and incorporated poor knowledge.

The positive feedback was also due to the fact that the approach was easily tailored to the different situations thanks to:

- The existence of different tools for the operational and managerial synthesis, the choice was quite free, meaning that a tool close to the field and to the standard practices of the relevant stakeholders could be selected.

- The choice of evaluation compared to goals - this fairly standard idea made it possible to set goals tailored to each situation. As a result, the synthesis addressed the gaps with the goals listed on an ordinal scale, which had been built based on simple ideas, viz., goal has not been reached, has been reached, has been exceeded. This choice meant that indicators whose evaluations were on a quantitative or qualitative basis could be incorporated.

d) Advantage of the EPISSURE Approach for Decision-Aid

The EPISSURE approach seems to have a real advantage within the context of the decision-aid process by promoting collective action twofold. One, it contributes to improving stakeholders' representation of the relevant issue (cognitive aspect of the approach mainly experienced during stage no. 2). Two, it prompts stakeholders to take action to reach the goals they are assigned (compliance aspect of the approach mainly experienced during use -see stage no. 3). Here we find the two sides of management tools [29] (p. 42.): the cognitive side and the conformation side. During stage no. 2, the cognitive aspect prevails. We have highlighted this through the observed learning processes (see 13.4b). The cognitive components provide an answer to the complexity of non-financial performance evaluation. During this phase, it should be pointed out that the conformation aspect is not missing. The fact that stakeholders agree to conduct dialogue within a given framework is a point in case. The stakeholders' representation of the context can evolve thanks to the alliance of both aspects. These findings are similar to the findings of the research conducted by Lebreton et al. [26]. Actually, in the course of participating and designing a system of indicators, a stakeholder becomes aware of what he/she perceives of the workings of the actual systems and learns from the other stakeholders' different points of view. The evolution involves a sizeable cognitive load, relying essentially on the acquisition of structured information [20]. Faced with new information, stakeholders find themselves in one of two possible situations [17]. Either the information is "coherent" with their representation of the system and enriches them, or the information creates "incoherence" and they have to deconstruct/reconstruct their representation of the system. The dialogue process then enables stakeholders to learn about the situation they are facing, better understand it, and gradually feel they are in a position to act. Here, we clearly find the foundation of our framed incremental and iterative dialogue process (see 13.2.4). Once this phase is over, the conformation side prevails to the extent where the stakeholders are "urged" to act to reach the goals they have been assigned. As the goals regularly come up for discussion (usually every year), conformation will again make way for dialogue. Here too, we find results that have been highlighted at the Canada Research Chair on territorial decision aid. Actually, once the indicator system is implemented, it provides a formalised representation of the dynamics of the evaluated actions, thus enabling the different stakeholders to communicate and exchange on the complexity of the evaluation.

The last point is the fact that the EPISSURE approach seems to help legitimise decisions. Actually, managing increasingly means legitimising, i.e., producing argu-

ments likely to make a corporate decision acceptable by its stakeholders [25] (p. 81). Yet everything happens as if the said legitimacy is now the result of management systems able to build both objects and collective agreements [11] (p. 4). Thanks to the framed dialogue process, the EPISSURE approach makes it possible to build the objects and collective agreements via the identification of the CID (Characteristic to be Identified by Dialog), among others. The evaluation stemming from the EPISSURE approach is the end-result of collaborative work and does not seem to be some sort of hat trick. From this standpoint, the EPISSURE approach does not legitimise the decision per se but the components underpinning the decision.

13.5 Conclusion

For the successful set-up of EPISSURE, for sponsorship or, more broadly, for non-financial performance evaluation, some obstacles must be overcome.

The first obstacle is to find a facilitator who is sufficiently knowledgeable about the different multi-criteria synthesis tools to select the most appropriate (operational and managerial synthesis tools) and to tailor the use of ELECTRE TRI to the evaluation context.

The second obstacle involves the framed dialogue process. The EPISSURE approach requires constructive behaviour during the work sessions. This supposes that one, the stakeholders regularly attend meetings to share their opinions and two, they agree to let their guard down enough to foster a discussion on the "real" issues. For the purpose of facilitating the stakeholders' constructive behaviour, the role of the facilitator is again decisive. To do so, he/she must have teaching skills to explain the approach, be a proficient group organiser, and have a knack for mediation to foster consensus on tricky points.

The third obstacle is undoubtedly the most complex. It concerns incorporating the tools within the organisation. Actually, the EPISSURE approach can only enable the management of non-financial performance if it is implemented over time. The first stage is the successful set-up of the EPISSURE approach. This supposes one, receiving strong support from management and two, tailoring the approach to the customs and practices of the organisation (for instance, adjusting the technical vocabulary of the synthesis tools to the corporate language). The next step is to ensure the durability of the approach within the organisation. This assignment is trickier and, for the time being, we do not have enough perspective to observe durability in the different cases. However, we can assume that durability supposes that the EPISSURE approach must:

- Enrich extant practices without adding overly strict constraints, the goal is to prompt organisation stakeholders to appropriate the approach gradually.
- Remain coherent with the corporate strategy; any policy moving toward a solely financial strategy would void the EPISSURE approach.
- Be incorporated into the reporting processes that already exist in the company.

Although we believe we have validated the advantage of the approach, the different companies still have to be monitored to analyse the durable integration of EPISSURE within an organisation. The analysis of coherence compared to other management tools and the analysis of the capacity to change behaviours are the main points requiring special attention.

References

1. F. Adam and J.-C. Pomerol. Critical factors in the development of executive systems - leveraging the dashboard approach. In M. Mora, G. Forgionne, and J. Gupta, editors, *Decision Making Support Systems: Achievements and Challenges for the New Decade*, pages 305–330. IGI Publishing, Hershey, PA, 2002.
2. F. Adam and J.-C. Pomerol. Developing practical support tools using dashboards of information. In F. Burstein and C.W. Holsapple, editors, *Handbook on Decision Support Systems 2*, pages 151–173. Springer-Verlag, 2008.
3. K. Alfsen and T. Moe. An international framework for constructing national indicators for policies to enhance sustainable development. In *Expert Group Meeting on Indicators of Sustainable Development*, New York, 13-15 December 2005. UN Department of Economic and Social Affairs, Division for Sustainable Development.
4. S. André. Evaluation de la performance non financière des entreprises: Apport des méthodes multicritère d'aide à la décision. Thèse d'Informatique / Aide à la Décision, Université Paris Dauphine, Paris, France, 2005.
5. M. Anielski. Alberta sustainability trends 2000: The genuine progress indicators report 1961-1999. Pembina Institute, Drayton Valley, 2001.
6. V. Belton and J. Pictet. A framework for group decision using a MCDA model: Sharing, aggregating or comparing individual information. *Revue des Systèmes de Décision*, 6(3):283–303, 1997.
7. CEE. Direction générale de l'emploi et des affaires sociales: Promouvoir un cadre européen pour la responsabilité sociale des entreprises. Commission Européenne, 2001.
8. CERES. Dix principes de protection de l'environnement et de transparence. Coalition for Environmentally Responsible Economies, 1989.
9. R. Chenhall and K. Langfield. Multiple perspective of performance measures. *European Management Journal*, 25(4):266–282, 2007.
10. S. Damart and B. Roy. The uses of cost-benefit analysis in public transportation decision-making in France. *Transport Policy*, 16(4):200–212, 2009.
11. A. David, A. Hatchuel, and R. Laufer. Sciences de gestion et sicences sociales: Un déficité d'identité. In A. David, A. Hatchuel, and R. Laufer, editors, *Les Nouvelles Fondations des Sciences de Gestion*, pages 1–12. Vuibert, Paris, 2001.
12. F. Debiesse. *Le Mécénat*. Presses Universitaires de France, Paris, 2007.
13. C. Dover. How dashboards can change your culture. *Strategic Finance*, 86(4):43–48, 2004.
14. S. Faucheux and I. Nicolaï. Quels indicateurs pour la responsabilité sociales des entreprises? Une application à l'industrie de l'aluminium. *Revue Gérer et Comprendre*, 76:42–54, 2004.
15. J.R. Figueira, B. Roy, and V. Mousseau. ELECTRE methods. In J.R. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 133–162. Springer, New York, 2005.
16. J. Figuiera and B. Roy. Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. *European Journal of Operational Research*, 139:317–326, 2002.
17. F. Flourentzos. Constructivisme piagétien dans l'aide à la décision - Contribution au développement durable en architecture. Thèse no 2418, Ecole Polytechnique Fédérale de Lausanne, 2001.

18. World Bank Group. Pollution prevention and abatement handbook, toward cleaner production. The World Bank Group in collaboration with the United Nations Environment Programme and the United Nations Industrial Development Organization, Washington, D.C, 1999.
19. H. Henderson. What's next in the great debate about measuring wealth and progress? *Challenge*, 39:50–56, 1996.
20. F. Joerin and P. Rondier. Indicateurs et décision territoriale: Pourquoi? quand? comment? In G. Sénécal, editor, *Les Indicateurs Socio-territoriaux et les Métropoles*, pages 9–36. Presses del'Université Laval, Québec, 2007.
21. R. Kaplan and D. Norton. The balanced scorecard: Measures that drive performance. *Harvard Business Review*, 70(1):71–79, 1992.
22. R. Kaplan and D. Norton. Using the balanced scorecard as a strategic management system. *Harvard Business Review*, 74(1):75–85, 1996.
23. R. Keeney. *Value-Focused Thinking: A Path to Creative Decisionmaking*. Harvard University Press, Cambridge, MA, 1996.
24. S. Labbouz, B. Roy, Y. Diab, and M. Christen. Implementing a public transport line: Multicriteria decision-making methods that facilitate concertation. *Operational Research: An International Journal*, 8(1):5–31, 2008.
25. R. Laufer. Les institutions du management: Légitimité, organisation et nouvelle rétorique. In A. David, A. Hatchuel, and R. Laufer, editors, *Les Nouvelles Fondations des Sciences de Gestion*, pages 45–81. Vuibert, Paris, 2001.
26. M. Lebreton, G. Desthieux, and P. Rondier. Chaire de recherche du Canada en aide à la décision territoriale. Université de Laval, available at http://www.adt.chaire.ulaval.ca/2_recherche/programme_recherche.php#systemes_indicateurs, 2009.
27. H. L¨ning, V. Malleret, J. Meric, Y. Pesqueux, E. Chiapello, D. Michel, and A. Solé. *Le Contrôle de Gestion: Organisation et Mise en Oeuvre*. Edition Dunod, Paris, 3rd edition, 2008.
28. L. Maystre and D. Bollinger. *Aide à la Négociation Multicritère: Pratique et Conseils - Collection Gérer l'Environnement*. Presses Polytechniques et Eniversitaires Romandes, Lausanne, 1999.
29. J.C. Moisdon. *Du Mode d'Existence des Outils de Gestion*. Edition Seli Arslam, Paris, 1997.
30. K. Nielsen. *Le Mécénat Mode d'Emploi*. Economica, Paris, 2007.
31. ONU. Action 21: Document publié par les Nations Unies à la suite du sommet de Rio. Available at www.un.org/french/ga/special/sids/agenda21/index.html, 1992.
32. M.E. Porter and M.R. Kramer. The competitive advantage of corporate philantropy. *Harvard Business Review*, 80(12):56–67, 2002.
33. B. Roy. *Méthodologie Multicritère d'Aide à la Décision*. Economica, Paris, 1985.
34. B. Roy. *Multicriteria Methodology for Decision Aiding*. Kluwer Academic Publishers, Dordrecht, 1996.
35. B. Roy. L'aide à la decision ajourd'hui: Que devrait-on en attendre? In A. David, A. Hatchuel, and R. Laufer, editors, *Les Nouvelles Fondations des Sciences de Gestion*. Vuibert, Paris, 2000.
36. B. Roy and D. Bouyssou. *Aide Multicritère à la Décision: Méthodes et Cas*. Economica, Paris, 1993.
37. B. Roy and P. Vincke. Pseudo-orders: Definition, properties and numerical representation. *Mathematical Social Sciences*, 14:263–274, 1987.
38. André S. and Roy B. Conception et mise en place d'un outil d'évaluation de la performance environnementale - Le cas des raffineries de la société Total. *Journal of Decision Systems*, 16(3):335–367, 2007.
39. T.L. Saaty. *Decision Making for Leaders: The Analytical Hierarchy Process for Decisions in a Complex World*. Wadsworth, Belmont, MA, 1982.
40. S. Schmidheiny. *Changing Course: A Global Business Perspective on Development and the Environment*. MIT Press, Cambridge, MA, 1992.
41. V. Seghers. *Ce qui Motive les Enterprises Mécènes*. Autrement, Paris, 2007.