THE EURO 2004 BEST POSTER AWARD

Choosing the Best Poster in a Scientific Conference

Raymond Bisdorff
University of Luxembourg – FSTC
Computer Science and Communication Research Unit
url: http://charles-sanders-peirce.uni.lu/bisdorff/

Abstract The chapter concerns the attribution of the EURO Best Poster Award at the 20$^{\text{th}}$ EURO Conference, held in Rhodes, July 2004. We present the historical decision making process leading to the selection of the winner, followed by a thorough discussion of the constructed outranking models and of the best choice recommendation.

Keywords: Best choice decision problem, Multiple evaluators, Multiple Criteria, Ordinal performances, Bipolar-valued outranking, Robust recommendation

Contents

Introduction

In this chapter, we report the elaboration of a recommendation for selecting the winner in a competition for the *EURO Best Poster Award* (EBPA[1]) at the EURO 2004 Conference in Rhodes (Greece). From an MCDA point of view, the real case study discussed here concerns a *unique best choice* decision problem based on multiple ordinal performance assessments given by the EBPA jury members, i.e. a multiple criteria group best choice decision problem.

The initiator of the EBPA, i.e. the Programme and Organisation Committees of the EURO 2004 Conference nominated five members in the award jury and fixed in advance four performance criteria: *Scientific Quality*, *Contribution to Theory and Practice of OR*, *Originality*, and *Presentation Quality*, to be taken into account in decreasing order of significance for selecting the EBPA winner. The call for participation in the EBPA resulted eventually in a pool of 13 poster submissions. Unfortunately, being quite busy at the conference, not all jury members had the possibility to inspect and evaluate all the competing posters. As a result the EBPA jury was left with an incomplete performance tableau showing some irreducibly missing values. With the help of an outranking based decision aid process, the EBPA jury could nevertheless agree on a unanimous final decision which was presented and scientifically argumented at the closing session of the EURO 2004 Conference.

The goal of this chapter is to present, comment and redo this decision aid process from an a posteriori – 2010 – perspective[2]. We therefore, first, report the historical case with its decision making process, – the involved actors, and – in particular, the actual decision aiding process with the historical unique best choice recommendation. We continue in a second section with discussing and analyzing more specifically the modelling of the EBPA jury's preferences. Finally, we propose a (re)building of the best choice recommendation with a particular focus on its robustness.

1. The historical case

In this first section we are going to present in detail the historical decision making process, followed by a thorough review of all the objects appearing in this decision making process. We close this section with a

[1] A glossary with abbreviations and symbols is provided at the end of the chapter.

[2] The seminal articles (Bisdorff et al., 2006, 2008) of the RUBIS decision aid methodology date from 2006 and 2008.

view on the decision aid process actually put into practice by the chair of the EBPA jury.

1.1 The decision making process

The decision making process we are going to describe here covers a period of approximately three months: from May to July 2004. We may grosso modo distinguish six steps.

Step 1: Defining and configuring the decision problem.
Apart from the traditional contributed and invited presentations, the Programme Committee (PC) of the 20th European Conference on Operational Research (EURO 2004) invited for special *discussion presentations* – a new kind of EURO K[3] conference participation consisting in a 30 minutes presentation in front of a poster in the style of natural sciences conferences. In order to promote this new type of poster presentations, the EURO 2004 conference organizers decided to attribute a special *EURO Best Poster Award* consisting of a diploma and a prize of 1000 € to the best poster. The actual selection procedure was dedicated to a special EBPA jury composed of three members of the Conference Programme Committee including the PC chair[4] and two members from the Organizing Committee. Furthermore, four selection criteria, in decreasing order of significance, were recommended: scientific quality, contribution to OR theory and practice, originality, and presentation quality.

Step 2: Collecting the competing posters.
Besides invited and contributed paper submissions, the EURO 2004 Programme Committee had called for discussion presentations based on posters to be held in parallel in an exhibition space in the main lobby of the Conference. Each presentation scheduled in such a session lasted 30 minutes, whereas the poster was informally exhibited during during a whole day (see Figure 1). At the end of the day all the posters were changed. Contributions that were suitable for a discussion presentation sessions: authors willing to present a poster, authors wishing to present more than one paper, contributions selected by the Programme Committee, and European Working Group promotional presentations.

[3] EURO K conferences, organised every three years, are the main dissemination instrument of EURO – the Federation of European Operational Research Societies, see http://www.euro-online.org.
[4] the author of this chapter.

Figure 1. The poster exhibition during the conference

Eventually 99 discussion presentations were scheduled at the EURO XX Conference in 8 sessions: MA (Monday 9:00–10:30), MC (Monday 14:00–15:30), MD (Monday 16:00–17:30), TA (Tuesday 9:00–10:30), TC (Tuesday 14:00–15:30), TD (Tuesday 16:00–17:30), WA (Wednesday 9:00–10:30) and WC (Wednesday 14:00–15:30) in separated time slots of 30 minutes parallel to the 15 or 16 regular organized and contributed session streams. They gave the authors the possibility to present and discuss their ongoing work with their poster illustration in the background. As illustrated in Figure 1, the posters attracted a large and interested audience.

Step 3: Gathering the performance assessments.
The actual evaluation of the competing posters by the five jury members was done at the occasion of the discussion presentations in front of the poster. The grading on the four selection criteria, recommended by the EBPA organizer, was guided by an evaluation sheet template provided by the jury chair (see Figure 2[5]).

Step 4: Aggregating all the preferential information.
Considering the ordinal nature of the recommended grading scale, the chair of the EBPA jury decided to follow an ordinal aggregation method

[5]Here the evaluation sheet template has been made anonymous. The real instance showed the actual coordinates of the poster authors instead of the abstract identifier shown here.

template

EURO Best Poster Award 2004

Evaluation sheet

Evaluator:

Evaluation marks should be given from 0 (very weak) – 10 (excellent)

authID	Scientific quality	Contribution to theory or practice of OR	Originality	Presentation quality
P_01				
P_02				
P_03				
P_04				
P_05				
P_06				
P_07				
P_08				
P_09				
P_10				
P_11				
P_12				
P_13				

Figure 2. The evaluation sheet used by the jury members

and to aggregate the jury members' evaluations into a global valued outranking relation, representing the bipolar valued characterisation of pairwise "at least as good" preference situations on the set of competing posters. All jury members were considered equi-significant whereas decreasing integer significance weights (from 4 to 1) were allocated, in accordance with the EBPA organizers regulations, to the four recommended selection criteria.

Step 5: Selecting the best poster.
The EBPA jury was asked to select the best – in the sense of the selection criteria retained by the Programme Committee – out of the 13 competing posters on the basis of their grades as gathered in the individual evaluation sheets. In July 2004, the jury unanimously accepted the best choice recommendation elaborated by their chair on the basis of the proposed global pairwise outranking relation and consequently attributed the EURO Best Poster Award 2004 to poster p_{10}, a poster with title: *Political Districting via Weighted Voronoï Regions* and authored by Federica RICCA, Bruno SIMEONE and Isabella LARI from the University of Rome "La Sapienza".

Step 6: Auditing the result.
The report on the selection procedure of the EBPA 2004 was eventually presented by the Programme Committee Chair at the closing session of

the EURO 2004 conference. The audience positively acknowledged the winner and the arguments which led the jury to particularly select this poster.

Before looking more precisely on the actual decision aiding process which guided the EBPA jury in the selection procedure, we review, first, the formal MCDA data appearing in this case study.

1.2 The formal data of the decision problem

We may distinguish a list of general MCDA data that can be identified in the decision making process above.

The actors and stakeholders.

1 The EBPA responsible organizer: in fact the joint EURO 2004 Programme and Organising Committees;

2 The EBPA jury: The Programme Committee nominated a jury of five members, three members of the Programme Committee (J. Blazewicz, R. Bisdorff (chair), G. Wäscher) and two members of the Organising Committee (N. Matsatsinis, C. Zopounidis), to evaluate the submitted posters on the basis of the proposed selection criteria and eventually attribute the award to the best submission. The chair of the award jury acted as decision aid analyst.

3 The authors having submitted their poster to the EBPA;

4 The conference participants: witness of the eventual winner and potentially the actual auditor of the overall decision making process.

The potential decision alternatives.
The Conference organizers offered the EURO Best Poster Award EBPA 2004 with the goal to encourage discussion presentations based on posters. This award, granted during the Closing session, consisted in a diploma and a prize of 1000€. All accepted discussion presentations authors were invited to compete for the EBPA. In order to participate the authors had to submit a reduced electronic PDF version of their poster before June 15th 2004. 13 candidates actually submitted an abstract and an image of their poster in due time. These were the potential decision alternatives for the best choice decision problem under review, denoted A in the sequel.

The subjects of the competing posters concern: – a variety of traditional OR topics like inventory planning and project management tools; – discrete mathematics problems with set covering and dice games; – applications in software development, in data and information systems, in the wood industry, in higher education, in the banking industry, and in political districting.

The selection criteria.
To evaluate the submitted poster images, the Programme Committee retained the officially recommended selection criteria: *Scientific Quality* (sq), *Contribution to OR Theory and/or Practice* (tp), *Originality* (or) and *Presentation Quality* (pq) in decreasing order of importance.

The performance tableau.
The EBPA jury members were invited to listen to the discussion presentations and evaluate the corresponding poster. In Table 1 are shown, for instance, the evaluation marks given by three jury members.

We may notice that j_1 expressed himself moderately by using only a reduced set of ordinal values: from 4 (lowest) to 9 (highest). Jury member j_2 used nearly the whole range of the given ordinal performance scale, from 1 (lowest) to 10 (highest), whereas j_3 used almost only the upper part (from 5 to 10) of the performance scale. Beside this apparent in-

Table 1. Evaluation marks given by three jury members j_1, j_2 and j_3

Poster ID	Scientific quality			Theory or practice of OR			Originality			Presentation quality		
	j_1	j_2	j_3	j_1	j_2	j_3	j_1	j_2	j_3	j_1	j_2	j_3
p_1	4	7	5	4	7	6	4	7	6	4	7	5
p_2	/	1	6	/	1	7	/	1	8	/	3	9
p_3	6	6	7	8	9	7	6	7	7	6	6	9
p_4	8	9	9	7	8	6	8	8	7	8	6	7
p_5	8	6	8	8	7	9	8	5	7	8	8	8
p_6	5	5	5	5	7	5	5	5	5	5	7	6
p_7	6	5	6	7	8	7	6	5	5	8	8	5
p_8	4	/	5	4	/	5	4	/	7	7	/	10
p_9	/	/	5	/	/	5	/	/	7	/	/	10
p_{10}	9	9	8	9	9	9	9	9	9	9	10	10
p_{11}	6	9	8	6	8	6	6	9	7	8	9	8
p_{12}	4	5	7	4	5	7	4	3	7	4	5	3
p_{13}	4	8	8	4	8	8	4	6	7	4	9	9

commensurability of the jury members' ordinal performance evaluations, a further serious problem represents the fact that not all five jury members did provide marks for all the competing posters. Jury member j_1, for instance, did not mark posters p_2 and p_9, whereas j_2 did not mark posters p_8 and again p_9 (see the slash (/) denotation in Table 1). This lack of information results from the fact that some jury members, due to availability constraints during the conference days, could not attend the public presentation of one or the other poster. All the posters in competition were, however, evaluated by at least two members of the award jury (see Appendix 13).

Finally, we may have a detailed look at the actual decision aiding process that was guiding the selection procedure of the EBPA jury.

1.3 The historical decision aid process

In order to assist the EBPA jury in selecting the winner among the competing posters, the chair of the EBPA jury, a professional decision aid specialist, deployed a standard multiple criteria base decision aid procedure. Four steps of this historical procedure are worthwhile to be reported here: – guiding the individual evaluation process of the jury members; – aggregating the collected individual preferences into a global pairwise outranking relation; – building a unique best poster recommendation (BCR) for the EBPA jury; – and, evaluating the robustness of the proposed BCR[6].

Guiding the evaluations of the jury members.
To harmonize as far as possible the evaluation process, a common evaluation sheet template (see Figure 2) was distributed to all the EBPA jury members. The main purpose of this template was to guide the jury members in their individual grading of the competing posters.

The filled in evaluation sheet for jury member 2 is shown in Figure 3. Please notice that some posters were not evaluated by all the jury members. Jury member 2, for instance, did not provide evaluations for posters 8 and 9 (see Figure 3).

Constructing an overall pairwise outranking relation.
From the five eventually gathered evaluations sheets (see Appendix 13), it becomes readily apparent that the collected gradings were all expressed on, in principle, non commensurable ordinal grading scales with eleven

[6]A glossary with abbreviations and symbols is provided at the end of the chapter.

jury member 2

EURO Best Poster Award 2004

Evaluation sheet

Evaluator: jury member 2

Evaluation marks should be given from 0 (very weak) – 10 (excellent)

authID	Scientific quality	Contribution to theory or practice of OR	Originality	Presentation quality
P_01	7	7	6	7
P_02	1	1	1	3
P_03	6	9	7	6
P_04	9	8	8	6
P_05	6	7	5	8
P_06	5	7	5	7
P_07	5	8	5	8
P_08	NA	NA	NA	NA
P_09	NA	NA	NA	NA
P_10	9	9	9	10
P_11	9	8	9	9
P_12	5	5	3	5
P_13	8	8	6	9

Figure 3. The evaluation sheet used by the jury members

grades from 0 (very weak) to 10 (excellent). The chair of the EBPA, being a specialist in the aggregation of non compensating, ordinal and possibly partial preference statements (Bisdorff, 2002), the construction into a global preference on the level of the jury as a whole was done with a specially adapted outranking approach, a forerunner of the Rubis method (Bisdorff et al., 2008). In accordance with the recommended

	p01	p02	p03	p04	p05	p06	p07	p08	p09	p10	p11	p12	p13
p01	0.00	16.00	-52.00	-76.00	-8.00	36.00	-32.00	52.00	30.00	-92.00	26.00	-84.00	-44.00
p02	-16.00	0.00	-32.00	-32.00	-32.00	-16.00	-16.00	-20.00	22.00	-52.00	-10.00	-32.00	-48.00
p03	64.00	48.00	0.00	8.00	32.00	96.00	72.00	92.00	38.00	-68.00	62.00	16.00	-8.00
p04	96.00	36.00	24.00	0.00	52.00	96.00	64.00	96.00	38.00	-44.00	50.00	40.00	8.00
p05	28.00	36.00	44.00	-4.00	0.00	100.00	56.00	96.00	38.00	-48.00	50.00	4.00	0.00
p06	8.00	16.00	-80.00	-80.00	-32.00	0.00	-16.00	72.00	30.00	-100.00	26.00	-92.00	-60.00
p07	36.00	44.00	-36.00	-8.00	-40.00	72.00	0.00	64.00	30.00	-80.00	38.00	0.00	-28.00
p08	0.00	20.00	-68.00	-60.00	-40.00	32.00	-20.00	0.00	42.00	-96.00	34.00	-68.00	-48.00
p09	-14.00	-2.00	-30.00	-30.00	-18.00	-2.00	-26.00	-2.00	0.00	-38.00	-22.00	-26.00	-30.00
p10	100.00	60.00	100.00	68.00	100.00	100.00	80.00	100.00	42.00	0.00	62.00	76.00	60.00
p11	42.00	22.00	-26.00	-42.00	-42.00	-14.00	-22.00	18.00	38.00	-62.00	0.00	-38.00	-14.00
p12	96.00	32.00	40.00	24.00	28.00	92.00	56.00	88.00	38.00	-36.00	50.00	0.00	12.00
p13	100.00	52.00	40.00	20.00	60.00	60.00	40.00	92.00	38.00	-4.00	62.00	28.00	0.00

Figure 4. Valued pairwise outranking characterisation

ranking of the selection criteria, a significance of 4 points was given to the *Scientific Quality*, 3 points to the *Contribution to OR Theory and/or Practice*, 2 point to the *Originality*, and, 1 point to the *Presentation Quality*. All jury members were considered equi-significant. The resulting historical global outranking relation is shown in Figure 4. The positive figures denote pairwise "*at least as good as*" situations that are validated by a weighted majority of jury members, whereas the negative figures denote non validated situations.

Building the best poster recommendation.
To find now the best poster to recommend for the EBPA, the EBPA jury chair was looking for the smallest subset of posters such that: – every non selected one was positively outranked by at least one of the selected poster (external stability); – and, the selected posters do not outrank each other (internal stability). Such a best choice set corresponds to a dominating kernel of the outranking digraph (see Bisdorff et al., 2006).

The historical global outranking relation luckily delivered a unique outranking kernel: singleton $\{p_{10}\}$, in fact a CONDORCET winner (see Figure 4). Poster p_{10} hence represents in view of the global outranking relation the evident recommendation for the EBPA winner.

Evaluating the robustness of the recommendation.
It remained however to verify that the result obtained was not, in fact, an artifact of the chosen numerical significance weights vector. Luckily again, the jury chair could prove that his apparent best choice recommendation was only depending on the officially recommended preorder of the significance of the selection criteria, but not on the effective numerical values used for the construction of the global valued outranking relation (see Bisdorff, 2004).

In the next section, we present and discuss now in detail, the construction of the global preferences of the EBPA jury with the help of an outranking approach.

2. Models of apparent preferences

Due to both the ordinal character of the performance scales and the presence of missing values, it was not possible, in the limited time span available at the EURO 2004 Conference, to construct the overall preferences of the jury members with a value oriented decision aid approach. To transform the ordinal marks into commensurable values would have needed a sophisticated preference elicitation procedure involving time consuming interviews of each jury member. Instead, the chair of the jury adopted a more descriptive, order statistics approach, inspired from social choice theory and generally promoted under the name "outranking approach" (see Roy and Bouyssou, 1993). In this section we are presenting in detail how this approach may allow us to model the apparent preferences of the EBPA jury members.

2.1 Pairwise "*at least as good as*" situations

Defining marginal "*at least as good as*" statements.
Let F denote the set of four selection criteria to be taken into account and let J denote the set of five jury members. If we consider, for instance, the evaluation of the posters for jury member j in J with respect to their scientific quality (sq), that is the main criteria for selecting the best poster, we may qualify the validation of pairwise "poster x is at least as good as poster y" situations, denoted $x \geqslant^j_{sq} y$, with the help of a bipolar[7] characteristic function $r(x \geqslant^j_{sq} y)$ defined for all couple of posters (x, y) as follows:

$$r(x \geqslant^j_{sq} y) := \begin{cases} +1 \quad \text{if} \quad g^j_{sq}(x) \geqslant g^j_{sq}(y), \\ -1 \quad \text{if} \quad g^j_{sq}(x) < g^j_{sq}(y), \\ 0 \quad \text{otherwise, i.e. when} \\ \qquad\qquad g^j_{sq}(x) = \text{'/'} \;\text{ or }\; g^j_{sq}(y) = \text{'/'}. \end{cases} \tag{1}$$

In Formula 1, $g^j_{sq}(x)$ and $g^j_{sq}(y)$ represent jury member j's performance evaluation of posters x, respectively y, with respect to preference viewpoint sq. In Table 2, we may read for instance that for jury member

Table 2. Pairwise performance comparisons by jury member j_1 on criterion sq

$r(\geqslant^1_{sq})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	-	0	−1	−1	−1	−1	−1	+1	0	−1	+1	−1	+1
p_2	0	-	0	0	0	0	0	0	0	0	0	0	0
p_3	+1	0	-	−1	−1	+1	+1	+1	0	−1	+1	+1	+1
p_4	+1	0	+1	-	+1	+1	+1	+1	0	−1	+1	+1	+1
p_5	+1	0	+1	+1	-	+1	+1	+1	0	−1	+1	+1	+1
p_6	+1	0	−1	−1	−1	-	−1	+1	0	−1	+1	−1	+1
p_7	+1	0	+1	−1	−1	+1	-	+1	0	−1	+1	+1	+1
p_8	+1	0	−1	−1	−1	−1	−1	-	0	−1	+1	−1	+1
p_9	0	0	0	0	0	0	0	0	-	0	0	0	0
p_{10}	+1	0	+1	+1	+1	+1	+1	+1	0	-	+1	+1	+1
p_{11}	+1	0	−1	−1	−1	−1	−1	+1	0	−1	-	−1	+1
p_{12}	+1	0	+1	−1	−1	+1	+1	+1	0	−1	+1	-	+1
p_{13}	+1	0	−1	−1	−1	−1	−1	+1	0	−1	+1	−1	-

j_1, posters p_1 and p_8 are each one judged at least as good as the other

[7] See (Bisdorff, 2002; Bisdorff et al., 2008).

($r(p_1 \geqslant_{sq}^{j_1} p_8) = r(p_8 \geqslant_{sq}^{j_1} p_1) = +1$). In Table 1, we see indeed that j_1 evaluated them equally with value $g_{sq}^{j_1}(p_1) = g_{sq}^{j_1}(p_8) = 4$. We also may note that poster p_1 is in fact not judged at least as good as poster p_3 ($r(p_3 \geqslant_{sq}^{j_1} p_1) = +1$ and $r(p_1 \geqslant_{sq}^{j_1} p_3) = -1$). This time, $g_{sq}^{j_1}(p_3) = 6$ against $g_{sq}^{j_1}(p_1) = 4$. It is also noteworthy that posters p_2 and p_9, as they were not evaluated by this jury member, may not be compared to any of the other posters ($r(x \geqslant_{sq}^{j_1} y) = 0$ for $x \in \{p_2, p_9\}$ and $y \in A - \{p_2, p_p\}$). The trivial reflexive comparison is globally ignored in this analysis.

In general, three different preferential situations may thus be characterised:

- Poster x is *better than* poster y (strict preference): $r(x \geqslant_{sq}^{j} y) = +1$ and $r(y \geqslant_{sq}^{j} x) = -1$.
- Poster x is *as good as* poster y (indifference): $r(x \geqslant_{sq}^{j} y) = +1$ and $r(y \geqslant_{sq}^{j} x) = +1$.
- Posters x and y are *mutually incomparable*: $r(x \geqslant_{sq}^{j} y) = 0$ and $r(y \geqslant_{sq}^{j} x) = 0$. Neither a preference, nor an indifference can be validated (indeterminacy).

It is important to notice that the bipolar characterisation $r(\geqslant)$ from A^2 to $\{-1, 0, +1\}$ forgets the actual performance values. The very magnitude of the performance differences is thereby ignored. Only the sign of the difference or a null difference are discriminated. We completely respect hence the purely ordinal character of the given performance measure scales. One may however consider that it is not always sure that a one point difference on a 0 to 10 points scale is really signifying a preference situation for sure. In many real decision aid cases, it is therefore opportune to analyze the actual preference discriminating power of the underlying performance measure scales.

Discriminating non equivalent performances.
In Equation 1 we have implicitly assumed that, for all the award jury members $j \in J$, a positive difference of one point on all the performance scales, indicates a clearly better performing situation. Indeed, with $g_{sq}^{j_2}(p_6) = 5$ and $g_{sq}^{j_2}(p_7) = 7$ jury member j_2 may validate that p_7 is at least as good as poster p_6 ($r(p_7 \geqslant_{sq}^{j_2} p_6) = +1$), but, may be, not the converse situation. ($r(p_6 \geqslant_{sq}^{j_2} p_7) = -1$). Indeed, in the context of solely ordinal performance evaluations, the actual confirmed preference discrimination threshold is commonly set equal to one ordinal level difference. For the decision aid practice, it may be opportune, the case given, to assume that a clearly warranted preference situation

is only given when a positive difference of at least two ordinal levels is observed. Depending on the actual discrimination of the ordinal performance evaluations, a one level difference may some time be seen as a still more or less equivalent performance, either, supporting, an indifference statement, or, indicating the hesitation between an indifference or a preference statement (see Bisdorff, 2002; Bisdorff et al., 2008).

For each jury member $j \in J$ and each preference point of view $f \in F$, we denote h^j_f, respectively p^j_f (with $0 \leqslant h_f < p_f \leqslant 10$), the indifference, respectively preference, threshold we may observe on the performance scale of preference point of view f for jury member j. For all couple (x, y) of decision alternatives where we dispose of valid performance evaluations $g^j_f(x)$ and $g^j_f(y)$, we may thus extend the definition of the bipolar-valued characteristic function r of the pairwise "at least as good as " $(x \geqslant^j_f y)$ comparison as follows:

$$r(x \geqslant^j_f y) := \begin{cases} -1 & \text{if} \quad g^j_f(x) - g^j_f(y) \leqslant -p^j_f \\ +1 & \text{if} \quad g^j_f(x) - g^j_f(y) \geqslant -h^j_f \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

If $g^j_f(x)$ or $g^j_f(y)$ are not available, we put $r(x \geqslant^j_f y)$ to the neutral value 0.

In Table 1 one may notice for instance that jury member j_2 has evaluated posters p_3, p_4 and p_5 on the criterion *Originality* with 7, 8, respectively 5 points. Suppose now that jury member j_2 admitted on his performance scale a preference threshold of 2 points and an indifference threshold of 0 points. In this case, $r(p_3 \geqslant^{j_2}_{or} p_4)$ becomes 0 as $g^{j_2}_{or}(p_3) - g^{j_2}_{or}(p_4) = 7 - 8 = -1$ which is higher than the negative preference threshold, but lower than the indifference threshold. Whereas, $r(p_5 \geqslant^{j_2}_{or} p_3)$ becomes -1 as $g^{j_2}_{or}(p_5) - g^{j_2}_{or}(p_3) = 5 - 7 = -2$ which is equal to the negative preference threshold. Similarly, $r(p_4 \geqslant^{j_2}_{or} p_3)$ or $r(p_4 \geqslant^{j_2}_{or} p_5)$ would both become $+1$. In case we encounter missing evaluations, as noticed before in Table 1, the bipolar characteristic function r will qualify any involved pairwise comparison as indeterminate, i.e. r will always take the neutral value 0.

Considering that we have to take into account the preferences of the five jury members on each one of the four selection criteria, we are in fact confronted in this decision aid problem with $5 \times 4 = 20$ individual "*at least as good as*" characterisations similar to the one shown in Table 3. How to aggregate this information into an overall global preference model will be described step by step hereafter.

2.2 Aggregating per viewpoint or per jury member

We will start by marginally aggregating the opinions of all the jury members concerning one specific preference point of view, namely the apparent *Scientific Quality* (sq) of the competing posters.

Appreciating the posters from the Scientific Quality viewpoint. Inspired by social choice theory (Fishburn, 1977; Sen, 1986; Arrow and Raynaud, 1986), we shall take the individual $r(\geqslant^j_{sq})$ characterisations of the five members of the award jury as a kind of pairwise voting result and balance the pro votes $(+1)$ against the con votes (-1) of a given "*at least as good*" statement. 0-valued characterisations are counted as abstentions. We thus obtain for the *Scientific Quality* criterion a bipolar characteristic function r of the overall "*poster x is at least as good as poster y*" statement with respect to the *Scientific Quality* viewpoint, denoted $(x \succsim_{sq} y)$ and defined as follows on each couple $(x, y) \in A^2$:

$$r(x \succsim_{sq} y) := \sum_{j \in J} \left(\frac{r(x \geqslant^j_{sq} y)}{|J|} \right) \tag{3}$$

The result of this aggregation operator r is shown in Table 3 and which admits the following semantics:

- A value of $+1.0$, respectively -1.0, means that, from the sq point of view, all five jury members *unanimously* judge poster x at least as good as poster y, respectively *not* at least as good as poster y.
- A positive value means that, from the sq point of view, *more* jury members judge poster x *at least as good as* poster y *than not*.
- A negative value signifies that, from the sq point of view, *more* jury members judge poster x *not at least as good as* poster y *than not*.
- The *null* value indicates an *indeterminate* situation, where the positive and the negative votes concerning the pairwise comparison of their scientific quality are balanced, and where no overall pro or con judgment hence can be made apparent.

This bipolar-valued characterisation $r(\succsim_{sq})$ has a nice order statistical property (see Barbut, 1980). $r(\succsim)$ represents in fact a *median characterisation* between a disjunctive $(\max_{j \in J}[r(\geqslant^j_{sq})])$ and a conjunctive $(\min_{j \in J}[r(\geqslant^j_{sq})])$ aggregation of the individual characterisations $r(\geqslant^j_{sq})$.

From an exploratory and descriptive data analysis point of view, the so characterised $\succsim_{sq}$ relation represents a compromise relation which is at minimal ordinal disagreement with all the individual "*at least as good as*" statements expressed. It gives us a convincing and reliable *central* model of the preferences of the EBPA jury from the *Scientific Quality* point of view. In Table 3 we show the result for all pairwise comparisons

Table 3. Comparing the posters from the *Scientific Quality* (*sq*) point of view

$r(\succsim_{sq})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	-	+.2	−.2	−.2.	−.2	+.2	−.4	+.6	+.4	−.2.	+.2	−.2.	−.6
p2	**−.2**	-	**−.6**	**−.6**	**−.6**	**−.2**	**−.2**	**−.2**	**0**	**−.6**	**−.4**	**−.6**	**−.6**
p_3	+.2	+.6	-	−.6	−.2	+.2	+.8	+.2	+.4	−.2.	+.6	−.2	+.6
p4	**+.2**	**+.6**	**+.6**	-	**+.6**	**+.2**	**+.8**	**+.2**	**+.4**	**+.2**	**+.6**	**+.6**	**+.2**
p_5	+.2	+.6	+.2	−.2	-	+.2	+.8	+.2	+.4	−.2	+.6	+.2	+.2
p_6	+.2	+.2	−.6	−.6	−.6	-	+.8	+.2	+.4	−.2.	+.2	−.2.	−.6
p_7	+.4	+.6	−.4	−.4	−.8	+.8	-	+.8	+.4	−.8	+.2	−.4	−.4
p_8	+.2	+.2	−.6	−.6	−.6	+.6	0	-	+.4	−.2.	+.2	−.2.	−.6
p_9	0	0	−.4	−.4	−.4	0	−.4	0	-	−.4	−.4	−.4	−.4
p10	**+.2**	**+.6**	**+.2**	**+.2**	**+.2**	**+.2**	**+.8**	**+.2**	**+.4**	-	**+.6**	**+.6**	**+.6**
p_{11}	+.6	+.4	−.2	−.6	−.6	−.2	−.2	+.2	+.4	−.6	-	−.6	−.2
p_{12}	+.2	+.6	+.2	−.2	+.2	+.2	+.8	+.2	+.4	+.2	+.6	-	+.2
p_{13}	+.2	+.6	+.6	−.2	+.6	+.6	+.4	+.2	+.4	+.2	+.6	+.2	-

of the posters. Take for instance the comparison of posters p_4 and p_1, where we notice that $r(p_4 \succsim_{sq} p_1) = +0.2$ and $r(p_1 \succsim_{sq} p_4) = -0.2$. Poster p_4 is judged having a better scientific quality than poster p_1 by a majority of jury members. Conversely, poster p_1 is judged having a better scientific quality than poster p_1 only by a minority of jury members. Moreover, as $r(p_4 \succsim_{sq} p_i) > 0$ for $p_{i \neq 4} \in A$, poster p_4 shows a positive majority margin with all the other posters. A majority of jury members expresses thereby that poster p_4 is at least as good as any of the other posters. A same situation may be verified for poster p_{10}.

Poster p_2 compares, on the contrary, negatively with all the other posters, except poster p_9, with which it appears mutually incomparable. So a majority of jury members express that poster p_2 is *not* at least as good as all the other posters, except for poster p_9.

It is worthwhile noticing here that we obtain this clear result even when jury members j_1 and j_2, did not provide performance evaluations for posters p_2 and p_9. Three jury members out of five have actually not conjointly evaluated this pair of posters and the remaining two jury members are divided in their opinions. Hence we obtain an indeterminate situation: $r(p_2 \succsim_{sq} p_9) = 0$ and $r(p_9 \succsim_{sq} p_2) = 0$. Both posters appear incomparable under the available information.

We may in fact compute such a bipolar-valued characterisation of the overall result on each of the four selection criteria and analyze the partial results from each preference point of view. However, we are now more interested in making apparent the overall preferences of each individual jury member by aggregating the comparisons over all the four imposed preference points of view.

Aggregating individual opinions.
Instead of aggregating the opinions of all jury members with respect to one preference dimension, as we did before, we may also aggregate the opinions of each jury member on all the selection criteria. To do so, we must take into account the hierarchy of significance that the decision problem organizer, i.e. the EURO 2004 Programme and Organising Committees, wished to give the four imposed selection criteria, i.e. $sq \succ tp \succ or \succ pq$. With no precise indications from the jury, the decision aid analyst fixed somehow arbitrarily the corresponding normalized numerical significance weights to: $w_{sq} = 0.4$, $w_{tp} = 0.3$, $w_{or} = 0.2$, and $w_{pq} = 0.1$. Hence, the total significance is, as required for a normalized significance weight vector, equal to 1.0.

We may now characterise a global "*at least as good as*" relation for each jury member, denoted $\succcurlyeq^j$ for $j = 1$ to 5 and defined in the following way:

$$r(x \succcurlyeq^j y) \;:=\; \sum_{f \in F} \Big(r(x \geqslant^j_f y) \cdot w_f \Big) \tag{4}$$

Table 4. Overall pairwise comparisons for jury member j_3

$r(\succcurlyeq^j)$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1		**−1.**	−1.	−.4	−.4	+.8	−.4	+.4	+.4	**−1.**	−.8	−.4	−1.
p_2	**+1.**		+.2	**+.2**	+.2	+1.	+1.	+.8	+.8	**−.6**	+.2	+.2	−.4
p_3	**+1.**	**+.6**		+.2	+.2	+1.	+1.	+.8	+.8	**−1.**	+1.	+.2	−.4
p_4	**+1.**	+.2	+.2		+.8	+1.	+1.	+.8	+.8	**−.2**	+.4	+.8	+.2
p_5	**+.4**	−.2	+.2	+.4		+1.	+.4	+.8	+.8	**−.2**	+.4	+.4	+.2
p_6	0	−1.	−1.	−1.	−.4		+.4	+.4	+.4	**−1.**	−.8	−1.	−1.
p_7	**+.6**	+.4	−.4	−.4	−.4	+.8		+.4	+.4	**−1.**	−.2	−.4	−1.
p_8	**+.4**	−.8	−.4	−.4	+.2	+1.	−.4		+1.	**−.8**	+.4	−.4	−.4
p_9	**+.4**	−.8	−.4	−.4	+.2	+1.	−.4	+1.		**−.8**	−.4	−.4	−.4
p_{10}	**+1.**	**+1.**	**+1.**	**+.2**	**+1.**	**+1.**	**+1.**	**+1.**	**+1.**		**+1.**	**+1.**	**+1.**
p_{11}	**+.8**	+.4	+.8	0	0	+.8	+.8	+.8	+.8	**−1.**		0	−.6
p_{12}	**+1.**	−.2	+.2	+.2	+1.	+1.	+.4	+.8	+.8	**−.2**	+.4		+.2
p_{13}	**+1.**	+.6	+1.	+.2	+1.	+1.	+1.	+.8	+.8	**−.2**	+.2	+1.	

Similar to the previous marginal aggregation concerning only the scientific quality of the competing posters, Equation 4 delivers again a median characterisation in between the disjunction or the conjunction of the individual $r(x \succcurlyeq_f^j y)$ characterisations along all selection criteria $f \in F$, but weighted by their significance w_f. Again, the so characterised relation $\succcurlyeq^j$ represents a significant compromise of the individual preference statements of a jury member j taking into account the specific significance of each preference dimension f in $F = \{sq, tp, or, pq\}$.

In Table 4 is shown the result for jury member j_3 for instance. One may notice in the upper left corner that $r(p_1 \succcurlyeq^{j_3} p_2) = -1$ and $r(p_2 \succcurlyeq^{j_3} p_1) = +1$. Which signifies that poster p_2 is *performing better than* poster p_1 on all four selection criteria. For posters p_3 and p_2, however,

$$\begin{aligned} r(p_3 \succcurlyeq^{j_3} p_2) &= 0.4 \cdot r(p_3 \succcurlyeq_{sq}^{j_3} p_2) \ + \ 0.3 \cdot r(p_3 \succcurlyeq_{tp}^{j_3} p_2) \\ &\quad + 0.2 \cdot r(p_3 \succcurlyeq_{or}^{j_3} p_2) \ + \ 0.1 \cdot r(p_3 \succcurlyeq_{pq}^{j_3} p_2) \\ &= 0.4 \cdot (+1) \ + \ 0.3 \cdot (+1) \ + \ 0.2 \cdot (-1) \ + \ 0.1 \cdot (+1) \\ &= +0.6 \end{aligned}$$

Positive validation of the "*poster p_3 is at least as good as poster p_2*" statement from the *Scientific Quality* (+0.4), *Contribution to OR Theory and/or practice* (0.3) and *Presentation Quality* (0.1) points of view is counter-balanced by the negative validation from the *Originality* (.2) point of view. Conversely, positive validation of the "*poster p_2 is at least as good as poster p_3*" statement from the *Contribution to OR Theory and/or practice* (0.3), the *Originality* (+0.2), and the *Presentation Quality* (0.1) points of view is counter-balanced by the negative validation from the *Scientific Quality* (0.4) point of view. Globally there are therefore appreciated to be more or less equally good.

As $r(p_i \succcurlyeq^{j_3} p_1) \geq 0$ for all $p_{i \neq 1} \in A$ (see Table 4: Column p_1), all posters are considered by j_3 to be as *at least as good as* p_1. On the contrary, as $r(p_i \succcurlyeq^{j_3} p_{10}) < 0)$ (see Table 4: Column p_{10}), all posters $p_{i \neq 10}$ are *not* considered to be *at least as good as* poster p_{10}. Furthermore, as $r(p_{10} \succcurlyeq^{j_3} p_i) > 0$ for all $p_{i \neq 10} \in A$, poster p_{10} is considered by j_3 to be *globally better than* any other poster. In social choice theory terms, p_{10} gives a CONDORCET *winner* for jury member j_3 .

We may compute such a global "*at least as good as*" characterisation $r(\succcurlyeq^j)$ for all five jury members. In order to analyze now the potential disagreements between the individual jury members' global preferences, we may use a Kendall like distance (see Bisdorff, 2008), denoted K, between the bipolar-valued characteristics of their apparent "*at least as good as*" statements . Let r and s be two jury member.

$$K(\succcurlyeq^r, \succcurlyeq^s) = \sum_{x \neq y \in A} \left(\frac{|r(x \succcurlyeq^r y) - r(x \succcurlyeq^s y)|}{2 \cdot |A|(|A| - 1)} \right) \quad (5)$$

As the bipolar characterisations may take values from −1.0 to +1.0, the disagreement distance K varies between 0.0 (no disagreement at all) and 1.0 (total disagreement). In case one relation is completely indeterminate and the other completely determined (with solely +1.0 or −1.0 values), one would obtain a K value of 0.5.

Table 5. Average disagreement between jury members and between preference viewpoints (in %)

$K(\succcurlyeq^{j_r}, \succcurlyeq^{j_s})$	j_2	j_3	j_4	j_5	$K(\succcurlyeq_{f_r}, \succcurlyeq_{f_s})$	tp	or	pq
j_1	31.3	32.2	33.1	31.3	sq	12.9	12.9	18.6
j_2		27.9	28.5	20.2	tp		11.8	19.5
j_3			29.2	25.7	or			19.1
j_4				37.1				

In the left part of Table 5 are shown the disagreement distances we obtain between the jury members. One may notice here that jury member j_1 is more or less equally distant to all the other jury members ($K(\succcurlyeq^1, \succcurlyeq^j) \approx 32\%$ for $j \neq 1 \in J$). Whereas, jury member j_5 shows a more differentiated situation with most disagreements with j_4 ($K(\succcurlyeq^5, \succcurlyeq^4) = 37.1\%$) and less with j_2 ($K(\succcurlyeq^5, \succcurlyeq^2) = 20.2\%$). In fact, we may verify with the help of this K measure, that the views of the jury members on the competing posters significantly disagree one from the other. This observation guarantees somehow that the jury members have indeed expressed each one independantly their own personal view on the competing posters.[8]

Similarly, the right part of Table 5 shows the disagreement distances between the four preference viewpoints[9] Most disagreement (19.5%) is here observed between the *Scientific Quality* (sq) and the *Presentation Quality* (pq). The least disagreement (11.8%)is observed between the *Contribution to Or Theory and/or practice* (tp) and *Originality* (or). It is worthwhile noticing that the disagreements between the preference viewpoints appear less important than those between the jury members. It seams as if, for *Scientific Quality*, *Contribution to OR Theory and/or practice* and *Originality*, high and low appreciations have been somehow more correlated[10] in the performance evaluations.

[8]The very short time period available between the posters' evaluation and the final selection of the best posters is another procedural circumstance of the decision making process which made it rather difficult for the jury members to coordinate before the final selection procedure.
[9]see global outranking per preference viewpoint in the Appendix.
[10]A common misunderstanding holds this apparent statistical correlation as the sign of a violation of the required preferential independence hypothesis. Consider, however, two certainly preferentially independent selection criteria: *Cost* and *Space* in a car selection problem. A

Finally, we still have to aggregate these individual global preference statements of the jury members on all the weighted selection criteria into an overall global characterisation of the pairwise outranking situations which become apparent between the competing posters.

2.3 Aggregating into a global "*outranking*" statement

Three strategies are available for the overall aggregation of the preferences:

1) First, aggregate the jury members' opinions on each preference point of view and then only, aggregate over the selection criteria.

2) Aggregate first over the selection criteria for each jury member and then only, aggregate to a global consensus among the jury members.

3) Or, we may directly aggregate all individual opinions over all the preference points of view.

We propose here to follow the third strategy.

Aggregating all opinions from every point of view.
All five award jury members, renowned experts in the field of Operations Research, are by nomination to be considered equal in significance for aggregating the preferential information[11]. Mixing this equi-significance of the five jury members with the imposed differentiated significance of the four selection criteria, we consider now to be in the presence of $5 \times 4 = 20$ criteria that we may gather into four equi-significance classes listed hereafter in decreasing order of importance: $\{sq^j \mid j \in J\}$, $\{tp^j \mid j \in J\}$, $\{or^j \mid j \in J\}$, and $\{pq^j \mid j \in J\}$.

Following a similar numerical weighting strategy as in the preceding Section, we associate the following normalized significance weight vector $\mathbf{w}$ with these four equivalence classes: $w^j_{sq} = 0.4/5$, $w^j_{tp} = 0.3/5$, $w^j_{or} = 0.2/5$ and $w^j_{pq} = 0.1/5$, for $j = 1$ to 5. Note that we recover hence the same relative weights $w_{sq} = 0.4$, $w_{tp} = 0.3$, $w_{or} = 0.2$ and $w_{pq} = 0.1$ for each preference dimension as before. We will use this property

large car is generally more expensive than a smaller one. Nevertheless, both selection criteria are per se independent.

[11] except perhaps the chair of the award jury, who may influence the final balance if an indeterminate situation arises. This was not the case here. On the contrary a clear and convincing solution appeared, as we will see later on.

when discussing alternative overall aggregation strategies in the next Paragraph.

All the 20 criteria in our case here are by design preferentially independent, exhaustive and consistent. We are hence in the presence of a coherent family of criteria[12] and a weighted additive aggregation of the individual criterion based characterisations is meaningful. As done before already, and considering the given significance vector $\mathbf{w}$, we may therefore compute the characterisation r of a global "poster x is at least as good as poster y" statement, denoted $(x \succsim_w y)$, as follows:

$$r(x \succsim_{\mathbf{w}} y) \;:=\; \sum_{f \in F \wedge j \in J} \Big(r(x \geqslant_f^j y) \cdot w_f^j \Big) \tag{6}$$

This $r(\succsim_{\mathbf{w}})$ function, defined on all the couples of posters, characterises what is commonly called a *pairwise outranking* situation (see Roy, 1991). As all the bipolar characteristic function before, It takes value in a rational[13] interval $[-1.0, +1.0]$ with the following semantics:

i) $r(x \succsim_{\mathbf{w}} y) = +1.0$: all jury members *unanimously validate* on all the selection criteria the statement that poster x is at least as good as poster y on all selection criteria.

ii) $+1.0 > r(x \succsim_{\mathbf{w}} y) > 0.0$: a *significant weighted majority* of jury members *validates* the statement that poster x is at least as good as poster y. For short we say that poster x outranks[14] poster y.

iii) $r(x \succsim_{\mathbf{w}} y) = -1.0$: No jury member validates on any preference dimension the statement that poster x is at least as good as poster y. In negative terms, all jury members *unanimously invalidate* such a statement.

iv) $-1.0 < r(x \succsim_{\mathbf{w}} y) < 0.0$: Under the given significance vector $\mathbf{w}$, a*significant weighted minority* of jury members only *validates* the statement that poster x is at least as good as poster y. In negative terms, a *significant weighted majority* of jury members in fact *invalidates* this statement. Symmetrically to the positive case, we say here for short that poster x does not outrank poster y.

---

[12] A family of criteria is coherent .... (see Roy and Bouyssou, 1993)

[13] One may admit without loss of generality, that it is always possible to express the significance weights with a set of integer numbers. Here the corresponding integer numbers would be 4 for the *Scientific Quality*, 3 for the *Contribution to OR Theory and/or practice*, 2 to the *Originality*, and 1 to the *Presentation Quality* point of view.

[14] Explain the connection with the classical outranking relations used in the Electre IS or in the Promethee approaches.

v) $r(x \succsim_{\mathbf{w}} y) = 0.0$: Under the given significance vector $\mathbf{w}$, the statement that poster x is at least as good as poster y may *neither be validated, nor, invalidated*. The overall weighted preferential judgment is, so to say, *suspended*.

Table 6. Global outranking of the posters considering significance vector $\mathbf{w}$

$r(\succsim_{\mathbf{w}})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	-	+.16	−.52	−.76	+.08	+.36	−.32	+.52	+.30	**−.92**	+.26	−.84	−.44
p_2	−.16	-	−.32	−.32	−.32	−.16	−.16	−.16	+.22	**−.52**	−.10	−.32	−.48
p_3	+.64	+.48	-	+.08	+.32	+.96	+.72	+.92	+.38	**−.68**	+.62	+.16	−.08
p_4	+.96	+.36	+.24	-	+.52	+.96	+.64	+.96	+.38	**−.44**	+.50	+.40	+.08
p_5	+.28	+.36	+.44	−.04	-	**+1.0**	+.56	+.96	+.38	**−.48**	+.50	+.04	**±0.0**
p_6	+.08	+.16	−.80	−.80	−.32	-	−.16	+.72	+.30	**−1.0**	+.26	−.92	−.60
p_7	+.36	+.44	−.36	−.08	−.40	+.72	-	+.64	+.30	**−.80**	+.38	±0.0	−.28
p_8	**±0.0**	+.20	−.68	−.60	−.40	+.32	−.20	-	+.42	**−.96**	+.34	−.68	**−.48**
p_9	−.14	−.02	−.30	−.30	−.18	−.02	−.26	−.02	-	**−.38**	−.22	−.26	−.30
p_{10}	**+1.0**	**+.60**	**+1.0**	**+.68**	**+1.0**	**+1.0**	**+.80**	**+1.0**	**+.42**	-	**+.62**	**+.76**	**+.60**
p_{11}	+.42	+.22	−.26	−.42	−.42	−.14	−.22	+.18	+.38	**−.62**	-	−.38	−.14
p_{12}	+.96	+.32	+.40	+.24	+.28	+.92	+.56	+.88	+.38	**−.36**	+.50	-	+.12
p_{13}	**+1.0**	+.52	+.40	+.20	+.60	+.60	+.40	+.92	+.38	**−.04**	+.62	+.28	-

In Table 6 is shown the bipolar characterisation of the "*global outranking*" statement on all the pairs of posters in the EBPA competition.

The bipolar-valued characteristic function $r(\succsim)$ still preserves the nice order statistical property we have mentioned in the previous Section when appreciating the posters from a single point of view, and when aggregating the opinions of a jury member. In the disagreement K-distance sense (see Equation 5), $r(\succsim)$ represents indeed again a weighted *median* characterisation between the all *disjunctive* ($\max_{f \in F, j \in J}[r(\geqslant^j_f) \cdot w^j_f]$) and the all *conjunctive* ($\min_{f \in F, j \in J}[r(\geqslant^j_f) \cdot w^j_f]$) aggregation of the individual "*at least as good as*" characterisations (see Barbut, 1980).

From an exploratory and descriptive data analysis point of view, the global outranking relation represents therefore a convincing compromise which is at minimal ordinal disagreement distance with all the individual "*at least as good as*" relations. It gives us hence a convincing and reliable *central* model of the global preferences of the EBPA jury.

Stability with respect to marginal aggregation strategies.
As mentioned before, we could have followed two alternate strategies for aggregating the individual preferences:

i) First aggregate the opinions of the jury members on each preferential point of view, and then, propose a global compromise viewpoint;

ii) Or, first, aggregate all preference points of view for each individual jury member, and then, propose the consensus opinion of the whole jury.

However, it is easy to verify that the additive formulation (see Equation 6) of the bipolar characterisation of the global "$x \succsim_{\mathbf{w}} y$" statement, coupled with the consistent choice of the individual weights w_f^j, induces in fact the equivalence of all three potential aggregation strategies.

Proposition 1.

$$r(x \succsim_{\mathbf{w}} y) \;=\; \sum_{j \in J} \big(r(x \succsim^j y)/|J| \big) \;=\; \sum_{f \in F} \big(r(x \succsim_f y) \cdot w_f \big). \qquad (7)$$

Proof. Note that $r(x \succsim^j y) = \sum_{f \in F} \big[r(x \succsim_f^j y) \cdot w_f \big]$ and that $w_f = \sum_{j \in J} w_f^j = \sum_{j \in J} w_j/|J|$. Similarly, $r(x \succsim_f y) = \sum_{j \in J} \big[r(x \succsim_f^j y) \cdot w^j \big]$ and that $w_j = \sum_{f \in F} w_f^j$. □

All three strategies lead hence naturally to the same weighted bipolar characterisation of the global pairwise "*outranking*" statement. It is evident that this result is mainly dependent on the effective verification of the preferential independence and significance of points of view, as well as, of the individual jury members.

Having herewith modelled the overall preferences of the award jury on all the competing posters, we are now prepared for rebuilding the historical best poster recommendation submitting to the decision of the EBPA jury members.

3. Rebuilding the best poster recommendation

As mentioned in the methodological part (see Section XXXX reference to Denis' Chapter), a global outranking relation as constructed in the preceding Section, apart from being trivially reflexive, a fact that we ignore deliberately in this case study, has commonly no further structural properties that would allow to implement a simple choice function for determining the globally best decision alternative. In this case here, however, we are lucky. A clear winner is appearing as we will discover soon.

3.1 Exploiting the Condorcet graph

The semantics of the bipolar-valued characterisation of the global outranking relation give access to a crisp graph called Condorcet[15] graph and denoted $C(A, \mathrm{S})$, where A represents the set of competing posters and S[16] represents a crisp outranking relation defined on A as follows:

$$(x\,\mathrm{S}\,y) \quad \text{is} \quad \begin{cases} \textit{true}\ (+) & \text{when} \quad r(x \succsim_{\mathbf{w}} y) > 0, \\ \textit{false}\ (-) & \text{when} \quad r(x \succsim_{\mathbf{w}} y) < 0, \\ \textit{indeterminate}\ (0) & \text{when} \quad r(x \succsim_{\mathbf{w}} y) = 0. \end{cases} \tag{8}$$

The S relation may thus only render a partially defined crisp outranking on A. In Table 6, we see, for instance, that $r(p_8 \succsim_{\mathbf{w}} p_1) = 0$. Consequently, $p_8\,\mathrm{S}\,p_1$ is indeterminate, i.e. the global outranking situation between p_8 and p_1 appears neither validated, nor invalidated. The cumulative significance of weighted positive (validating) arguments is here exactly counterbalanced by the cumulative significance of the weighted negative (invalidating) arguments and no global conclusion concerning the validation or not of the outranking situation in question can be drawn. This is not a symmetrical situation, however. The converse global outranking situation, where we see that $r(p_1 \succsim_{\mathbf{w}} p_8) = +.52$), is, however, strongly validated with more than 75% significance[17]. Several other similar indeterminate cases do appear in the Condorcet graph under review (see Table 6).

Based on this Condorcet graph (see the + and − denotation in Table 7), three progressively extended exploitation approaches become available:

i) Determine the Condorcet *winner*, if there is one, or

ii) Determine its *maximal strong components* if the Condorcet graph shows a transitive global outranking, or

iii) Determine its *outranking kernel*, if there is one.

[15] We follow here a suggestion made by Barbut (1980) who calls a median cut graph after Marie Jean Antoine Nicolas de Caritat, marquis de Condorcet (1743–1794). A French mathematician and political scientist who is the inventor of the pairwise voting procedure named after him.

[16] The S notation comes from the French term "*surclasser*" (to *outrank*).

[17] Passing from the r characteristic function to classic election style majority percentages is readily achieved by shifting the r value up by 1.0 and dividing the result by 2.0. For instance, $r(p_{10} \succsim_{\mathbf{w}} p_9) = 0.42$, which gives in percentages: $(0.42 + 1.0)/2.0 = 0.71 = 71\%$.

Table 7. The CONDORCET outranking relation

x S y	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1		+	−	−	+	+	−	+	+	−	+	−	−
p_2	−		−	−	−	−	−	−	+	−	−	−	−
p_3	+	+		+	+	+	+	+	+	−	+	+	−
p_4	+	+	+		+	+	+	+	+	−	+	+	+
p_5	+	+	+	−		+	+	+	+	−	+	+	0
p_6	+	+	−	−	−		−	+	+	−	+	−	−
p_7	+	+	−	−	−	+		+	+	−	+	0	−
p_8	0	+	−	−	−	+	−		+	−	+	−	−
p_9	−	−	−	−	−	−	−	−		−	−	−	−
$\mathbf{p_{10}}$	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**		**+**	**+**	**+**
p_{11}	+	+	−	−	−	−	−	+	+	−		−	−
p_{12}	+	+	+	+	+	+	+	+	+	−	+		+
p_{13}	+	+	+	+	+	+	+	+	+	−	+	+	

The CONDORCET winner.
In CONDORCET's method, the winner of an election is a decision alternative that, if it exists, outranks all the other competing posters. Note that, as Condorcet was essentially considering a strict preference model, the CONDORCET winner, if it existed, was necessarily unique. As the outranking relation here is not an asymmetric relation, we may find, the case given, several such CONDORCET winners in a global outranking graph $C(A, \succsim)$.

Careful inspection, now, of the Table 6 – line by line – makes it apparent that poster p_{10} represents obviously such a CONDORCET winner. It outranks positively (see Table 7) all other posters with a comfortable minimal weighted significance of 71% (see line p_{10} in Table 6).

We are lucky in the case here. No other competing poster is in a similar good situation, and p_{10} may thus be recommended, on the basis of the given outranking graph as the winner of the EURO 2004 BPA competition.

The top strong component.
Already Condorcet himself noticed that his pairwise voting approach could end up with cyclic strict global preferences, an apparent *social choice paradox*, named after him. In the multiple criteria based aggregation of "*at least as good as*" statements, such potentially cyclic outrankings are, however, *not* considered to be *paradoxical* or even problematic at all. They simply show, the case given, that each preference dimension may well express cyclically opposed preferential opinions, so that no global consensus on a unique linearly ordered common point of view is possible. Roy and Bouyssou (1993) propose therefore, in the

ELECTRE method (see the methodological Chapter by Denis XXXX), to collapse the strong components (maximal cycles in fact) of the CONDORCET graph into potential equivalence classes of decision alternatives, and to exploit the so-reduced CONDORCET graph for building the best choice recommendation.

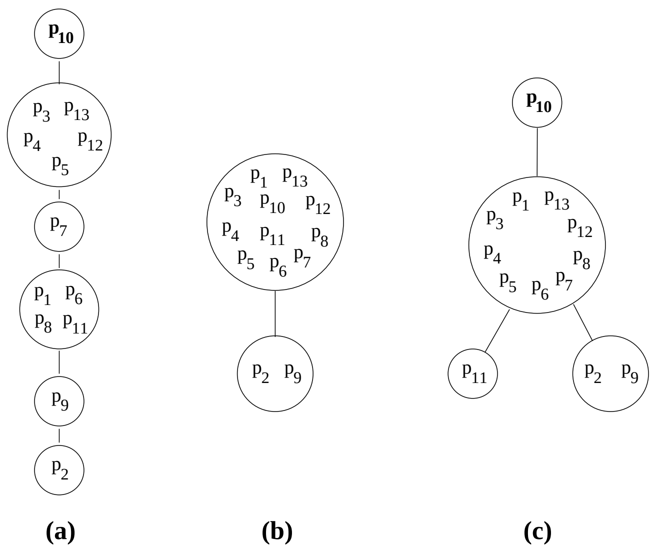

Figure 5. Hasse diagrams of strong components reduced outranking graph with 1 level preference discrimination (**a**), with 2 levels preference discrimination (**b**), and with additional veto effects (**c**)

From our crisp $C(A, \mathrm{S})$ graph, we obtain the following, linearly ordered, strong components (see Figure 5.a): $\{\mathbf{p_{10}}\} \succ \{p_3, p_4, p_5, p_{12}, p_{13}\} \succ \{p_7\} \succ \{p_1, p_6, p_8, p_{11}\} \succ \{p_2\} \succ \{p_9\}$. A CONDORCET winner, the case given, necessarily appears as best singleton strong component. We find here again confirmed that poster p_{10} clearly dominates indeed all the other competing posters. Considering furthermore the depth of the linear ordering of the strong components, we may notice that the five jury members do share apparently loads of preferential opinions. A significant majority, for instance, shares the opinion that p_9 is the worst candidate (see Table 7 line p_9). The global outranking gives therefore this weak ordering of the competing posters. This opportune situation is, however, not generally given. There might appear for instance several CONDORCET winners due, for instance, to a less precise discrimination of the individual performances.

Table 8. Global outranking of the posters with a preference discrimination threshold of two points

$r(\succsim_{\mathbf{w}})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	-	+.32	−.32	−.40	±0.0	+.66	+.04	+.70	+.34	**−.84**	+.36	−.62	−.26
p_2	−.16	-	−.24	−.32	−.32	−.16	−.16	−.18	+.32	**−.48**	−.02	−.32	−.40
p_3	+.82	+.54	-	+.16	+.50	+.98	+.72	+.96	+.40	**−.40**	+.62	+.34	+.14
p_4	+.98	+.46	+.60	-	+.66	+.98	+.70	+.96	+.38	**−.06**	+.56	+.68	+.16
p_5	+.58	+.42	+.46	+.18	-	+1.0	+.62	+.96	+.38	**−.22**	+.50	+.22	+.16
p_6	+.40	+.24	−.42	−.56	−.16	-	−.18	+.78	+.30	**−.96**	+.26	−.48	−.46
p_7	+.48	+.44	+.10	−.04	−.16	+.74	-	+.64	+.30	**−.62**	+.46	+.10	−.06
p_8	+.20	+.32	−.40	−.50	−.22	+.54	−.06	-	+.42	**−.86**	+.34	−.52	−.42
p_9	−.08	+.10	−.30	−.24	−.18	−.02	−.18	−.02	-	**−.34**	−.22	−.20	−.30
p_{10}	**+1.0**	**+.60**	**+1.0**	**+.76**	**+1.0**	**+1.0**	**+.80**	**+1.0**	**+.42**	-	**+.62**	**+.88**	**+.64**
p_{11}	+.42	+.26	−.25	−.28	−.34	+.18	−.04	+.32	+.38	**−.50**	-	−.24	±0.0
p_{12}	+.96	+.44	+.60	+.46	+.42	+.94	+.68	+.90	+.38	**−.02**	+.56	-	+.28
p_{13}	**+1.0**	**+.56**	**+.50**	**+.36**	**+.60**	**+.80**	**+.40**	**+.94**	**+.40**	**+.24**	**+.62**	**+.40**	-

Let's consider for the moment that a real better performing situation is only warranted when we observe a positive difference of at least two ordinal levels, a not unreasonable working hypothesis. We consequently obtain a much less clear cut global outranking picture. Two strong components only remain (see Figure 5.b), where $\{p_2, p_9\}$ is the less preferred of both. All the other competing posters are now considered to be equally preferred. The top strong component gathers under this working hypothesis, eleven of the thirteen best choice candidates. It obviously does not represent anymore a satisfactory potential best choice recommendation.

However, two CONDORCET winners do appear now: poster p_{10}, as well as poster p_{13}, outrank all the other candidates in this revised CONDORCET graph (See lines p_{10} and p_{13} in Table 8). By recognizing these CONDORCET winners directly from the values of the weighted bipolar characterisation $r(\succsim_{\mathbf{w}})$, one would readily notice that poster p_{10} gives a CONDORCET winner with at least 71% significance support in both cases (see line p_{10} both in Table 6 and in Table 8). Whereas poster p_{13}, actually a CONDORCET winner only in the reduced preference discrimination case (see Table 8) and with at least 62% of significance only, gives a somehow less convincing best poster candidate.

The outranking kernel.
It becomes apparent from the preceding considerations, that, in order to be suitable in a decision aid problem, a best choice recommendation should correspond to a *maximal* or, if not available, to a somehow *initial* node of the global outranking relation. A CONDORCET winner, if it

exists, fulfills ideally this condition. If the CONDORCET graph is a, perhaps partial, weak order, the maximal equivalence class, or classes the case given, give a potential set of somehow equivalent best poster candidates. The overall aggregation may, however, yield a Condorcet graph which generally shows neither a transitive nor a complete crisp outranking. An extension of the maximality condition (see Roy, 1985), leads therefore to the following three conditions, a suitable best choice recommendation should fulfill:

1 All decision alternatives, not retained as candidate for the best choice, should be rejected with objective reasons. The best choice recommendation should outrank the rejected alternatives, a fact called "*externally stable*".

2 The recommended set of potential best alternatives should be as limited in cardinality as possible, ideally a singleton.

3 The best choice candidates retained in a choice recommendation should be perceived either equivalent or incomparable, a fact called "*internally stable*".

A choice recommendation fulfilling these three conditions is actually called an *outranking kernel*[18]. And, both CONDORCET winners, mentioned in the previous section, represent two such outranking kernels observed in the corresponding CONDORCET graphs. However, as we have already mentioned, they are not one as well determined as the other. This insight gives us the hint that the Condorcet graph, due to its crisp polarization effect, is not well suited for discriminating between several best poster candidates. And we might, instead, have advantage in formulating a best poster recommendation directly from the bipolar-valued outranking characterisation.

3.2 The RUBIS best choice method

This approach has been promoted under the name RUBIS by Bisdorff et al. (2008). It results from general mathematical and algorithmic results obtained for computing best choice recommendations in bipolar-valued directed graphs (see Bisdorff et al., 2006). We shall briefly outline the main theoretical concepts and formulas.

Similar to the bipolar characterisation of a pairwise outranking situation between competing posters, one may also thus characterise the

[18] See glossary entry "Kernel" at the end of the Chapter.

more or less validation of the fact that a given subset of decision alternatives represents a suitable best choice recommendation. We need only to adequately characterise the statement that the corresponding choice is internally and externally stable.

Let $Y \subseteq A$ be a non empty subset of potential best choice candidates. We denote $\Delta^{\text{ind}}(Y)$, respectively $\Delta^{\text{dom}}(Y)$ or $\Delta^{\text{abs}}(Y)$, the bipolar characteristic value we may attribute to the statement that "Y is internally stable", respectively "Y is *dominantly stable*" or "Y is *absorbingly* stable". Formally, we define these values for all couples $(x, y) \in A^2$ $(x \neq y)$ of posters as follows:

$$\Delta^{\text{ind}}(Y) := \begin{cases} 1.0 \quad \text{if} \quad |Y| = 1, \\ \min_{(x \neq y) \in Y^2} \left[r(x \succsim_{\mathbf{w}} y) \right] \quad \text{otherwise.} \end{cases} \tag{9}$$

$$\Delta^{\text{dom}}(Y) := \begin{cases} 1.0 \quad \text{if} \quad Y = A, \\ \min_{x \notin Y} \left[\max_{y \in Y} r(y \succsim_{\mathbf{w}} x) \right] \end{cases} \tag{10}$$

$$\Delta^{\text{abs}}(Y) := \begin{cases} 1.0 \quad \text{if} \quad Y = A, \\ \min_{x \notin Y} \left[\max_{y \in Y} r(x \succsim_{\mathbf{w}} y) \right] \end{cases} \tag{11}$$

We get the same semantics as with the bipolar characterisation of the preferential statements. With stab $\in \{\text{ind}, \text{dom}, \text{abs}\}$,

i) $\Delta^{\text{stab}}(Y) = 1.0$ signifies that it is *certainly validated* that Y yields a choice recommendation which verifies the respective stability condition.

ii) $\Delta^{\text{stab}}(Y) > 0.0$ signifies that it is *more validated than invalidated* that Y yields a choice recommendation which verifies the respective stability condition.

iii) $\Delta^{\text{stab}}(Y) = -1.0$ signifies that it is *certainly invalidated* that Y yields a choice recommendation which verifies the respective stability conditions.

iv) $\Delta^{\text{stab}}(Y) < 0.0$ signifies that it is *more invalidated than validated* that Y yields a choice recommendation which verifies the respective stability conditions.

v) $\Delta^{\text{stab}}(Y) = 0.0$ signifies as usual an indeterminate situation where neither the validation, nor the invalidation may be assumed.

We have shown in Bisdorff et al. (2006), that the kernels of the Condorcet graph correspond bijectively to the choice sets that are internally and dominantly stable.

Table 9. Internal and external stability of potential best choice recommendations

choice (Y)	Δ^{ind}	Δ^{dom}	Δ^{abs}	choice (Y)	Δ^{ind}	Δ^{dom}	Δ^{abs}
$\{\mathbf{p_{10}}\}$	**+1.0**	**+.42** (71%)	**-1.0**	$\{p_{05}\}$	+1.0	-.48	-.42
$\{\mathbf{p_{13}}\}$	+1.0	**-.04** (48%)	-.60	$\{p_{02}\}$	+1.0	-.52	-.02
$\{\mathbf{p_9}\}$	**+1.0**	**-.38**	**+.22**	$\{p_{11}\}$	+1.0	-.62	-.22
$\{\mathbf{p_9}, \mathbf{p_{10}}\}$	**-.42**	**+.60** (80%)	**+.22**	$\{p_{03}\}$	+1.0	-.68	-.80
$\{p_{10}, p_{13}\}$	-.60	+.42	-.60	$\{p_{07}\}$	+1.0	-.80	-.32
$\{p_{12}\}$	+1.0	-.36	-.92	$\{p_{01}\}$	+1.0	-.92	-.16
$\{p_{10}, p_{12}\}$	-.76	+.42	-.92	$\{p_{08}\}$	+1.0	-.96	-.20
$\{p_{04}\}$	+1.0	-.44	-.80	$\{p_{06}\}$	+1.0	-1.0	-.16

In Table 9 we show the evaluation of the stability conditions for all potential singletons and some pairs of posters. It appears, that poster p_{10}, with $\Delta^{ind}(p_{10}) = 1.0$, $\Delta^{dom}(p_{10}) = +.42$ and $\Delta^{abs}(p_{10}) = -1.0$, yields the unique internal and dominantly stable choice recommendation available in the bipolar valued global outranking relation $\succsim_{\mathbf{w}}$ defined on A. In Figure 6.a, we may see indeed that p_{10} outranks all other compet-

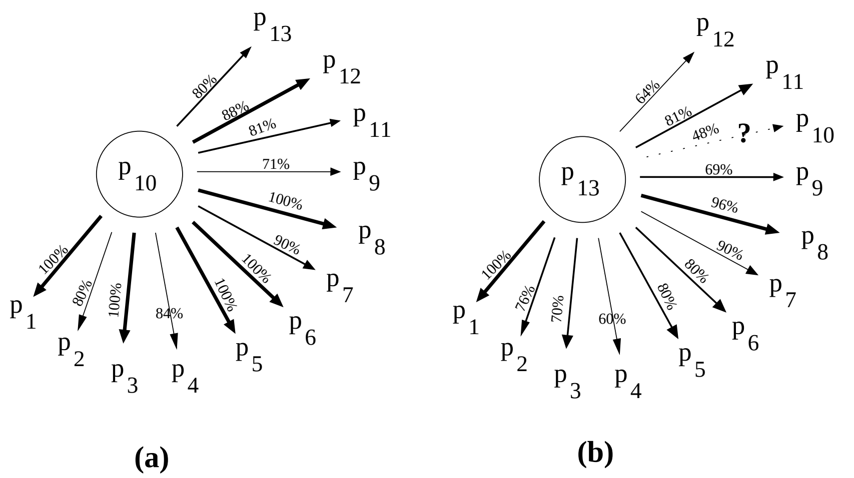

Figure 6. Dominance stability of singletons $\{p_{10}\}$ and $\{p_{13}\}$ in the global weighted outranking graph (Label fig:p10domination)

ing posters with a minimum significance of 71%[19] and is not outranked for certain by any other candidate. It is worthwhile mentionning that poster p_{13} appears as second potential choice recommendation as it also outranks all other competing posters, except poster p_{10}, with a minimal

[19]The conversion formula for percentages is $(\Delta^{stab}(Y) + 1.0)/2.0$.

significance of 64% (see Figure 6.b). It also gets apparent in Table 9, that poster p_9 yields the unique absorbingly stable choice with 61% of significance. This candidate is outranked by all the other competing posters with a minimum signficance of 61%. Finally, it is worthwhile noticing that, the pair $\{p_9, p_{10}\}$ would potentially yield a higly outranking (80% signficance), but, at the same time, outranked (61% of significance) choice recommendation. This recommendation would, however, not be stable. Poster p_9 is indeed outranked by poster p_{10} with a significance of 76% (see Table 6). Similarly, all other pairs are not internally stable.

Besides these singletons and pairs mentioned so far, no other small subset of competing posters is convincingly outranking all the others. At the sight of the results shown in Table 9, we may hence conclude that poster p_{10} represents definitely the best candidate that we may recommended the EBPA jury to attribute the EBPA.

But is this result not an artifact of our preference modelling strategy? Isn't this result not an anecdotal consequence of the numerical significance weights we are using in the computation of the bipolar valued characterisation of the global outranking situations? Answering these questions is the subject of the following, last, Section.

3.3 Robustness analysis

Three strategies for testing the stability of the previous result with respect to some variants of the preference model construction are proposed hereafter:

1 Taking into account large positive and negative performance differences. This is the specialty of the Electre outranking concept.

2 Requiring a qualified – high – significance level for the validation of an outranking statement.

3 Testing the stability of the Condorcet graph with respect to the numerical significance weights.

Taking into account very large performance differences.
In the Electre methods, Roy (1991); Roy and Słowiński (2008) suggest polarizing the global outranking situation by, on the one hand, cutting those arcs in the Condorcet graph where a worst performance (a veto) is challenging an otherwise significant "*at least as good as*" situation and, on the other hand, reinforcing those significant "*at least as good as*" situations where a best performance (a counter-veto) may be observed. Let

us set somehow arbitrarily such veto or counter-veto threshold, denoted v_j for $j \in J$, to the maximum spread of the performances given by each jury member minus one point[20]. From Table 13 in Appendix 13, we see that for j_1 it is 4 points, for j_2 it is 7 points, for j_3 it is 6 points, for j_4 it is 5 points, and for j_5 it is 7 points.

In order to detect these veto situations, denoted $\lll^j_f$, we are using again a bipolar characteristic function $r(\lll^j_f)$, defined as follows on all couples (x, y) in A^2:

$$r(x \lll^j_f y) = \begin{cases} +1 & \text{if} \quad g^j_f(x) - g^j_f(y) \leqslant -v_j, \\ -1 & \text{if} \quad g^j_f(x) - g^j_f(y) \geqslant +v_j, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Note that the bipolar symmetric negation of a serious worst performance (veto) situation $\lll^j_f$, namely changing the sign of its $r(\lll^j_f)$ value, gives the characteristic value of the corresponding very best performance (counter-veto) situations, denoted $\ggg^j_f$.

Extending the ideas of Roy (1991), we may now describe a pairwise global outranking situation[21], denoted $x \, \widetilde{\mathrm{S}} \, y$, with the help of the following bipolar characteristic function r:

$$r(x \, \widetilde{\mathrm{S}} \, y) = \begin{cases} r(x \succsim_{\mathbf{w}} y) & \text{if} \quad \forall f \in F : \big(r(x \lll^j_f y) = -1\big) \\ & \qquad \wedge \big(r(x \ggg^j_f y) = -1\big), \\ +1 \quad \text{if} \quad r(x \succsim y) > 0.0 \text{ and } \exists f \in F : \big(r(x \ggg^j_f y) = +1\big) & \\ \qquad\qquad \text{and} \quad \nexists f \in F : \big(r(x \lll^j_f y) = +1\big), & \\ -1 \quad \text{if} \quad r(x \succsim y) < 0.0 \text{ and } \exists f \in F : \big(r(x \lll^j_f y) = +1\big) & \\ \qquad\qquad \text{and} \quad \nexists f \in F : \big(r(x \ggg^j_f y) = +1\big), & \\ 0 \quad \text{otherwise.} & \end{cases} \tag{13}$$

The resulting semantics are the following:

- $r(\widetilde{\mathrm{S}})$ remains unchanged with $r(\succsim_{\mathbf{w}})$ in case we do not observe any veto or counter-veto situation.

[20]The EBPA 2004 jury members did not feel any need to consider such veto effects when deliberating on the final best poster choice. This is certainly related to the easy dominating situation of poster p_{10}.

[21]The classical outranking relation, as used in the various ELECTRE methods, differs slightly from our bipolar definition here in the sense that the large performance difference polarization is solely operated for a veto situation, but not for a counter-veto situation. This unipolar handling may induce, however, abusive strict invalidation of otherwise more or less validated outranking situations.

- We get a positively polarized validation, i.e. $r(x \gtrsim_{\mathbf{w}} y) \rightarrow +1.0$, in case we observe a significant positive outranking coupled with a very best performance (a counter-veto) on at least one criterion and we observe no serious counter-performance raising a veto on some other criterion.
- We get a negatively polarized invalidation, i.e. $r(x \gtrsim_{\mathbf{w}} y) \rightarrow -1.0$, in case we have a significant negative outranking coupled with a serious worst performance on at least one criterion and we observe no counter-veto on some other criterion.
- In all the other cases, i.e. when we observe conjointly best and worst performances, or a positive validation coupled with a serious worst performance, or a negative validation coupled with a very best performance, we admit the neutral zero value. Neither a validation, nor an invalidation of the global outranking situation may then be assumed and we suspend the validation of the corresponding outranking statement.

Table 10. Global outranking with veto and counter-veto effects

$r(\widetilde{\mathrm{S}})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	-	+.32	−**1.0**	−1.0	−1.0	+.66	±0.0	+.70	+1.0	−**1.0**	+1.0	−1.0	−1.0
p_2	−.16	-	−1.0	−1.0	−1.0	−0.16	−1.0	−.18	+.32	−**1.0**	±0.0	−1.0	−1.0
p_3	+**1.0**	+1.0	-	+.16	+.50	+.98	+.72	+1.0	+.40	−.**40**	+1.0	+.34	+1.0
p_4	+1.0	+1.0	+.60	-	+.66	+.98	+.70	+1.0	+.38	−.**06**	+1.0	+.68	+1.0
p_5	+1.0	+1.0	+.46	+.18	-	+1.0	+.62	+1.0	+1.0	−.**22**	+1.0	+.22	±0.0
p_6	+.40	+.24	−.42	−.56	−.16	-	+.18	+.78	+1.0	−**1.0**	+1.0	−.48	−1.0
p_7	+1.0	+1.0	+.10	−.04	−.16	+.74	-	+.64	+1.0	−**1.0**	+1.0	+.10	±0.0
p_8	+.20	+.32	−1.00	−1.00	−1.0	+.54	−.06	-	+.42	−**1.0**	+1.0	−.52	−1.0
p_9	−1.0	+.10	−.30	−.24	−1.0	−1.0	−1.0	−.02	-	−**1.0**	±0.0	−1.0	−1.0
p_{10}	+**1.0**	+**1.0**	+**1.0**	+.**76**	+**1.0**	+**1.0**	+**1.0**	+**1.0**	+**1.0**	-	+**1.0**	+**1.0**	+**1.0**
p_{11}	±0.0	±0.0	−1.0	−1.0	−1.0	±0.0	−1.0	±0.0	±0.0	−**1.0**	-	−1.0	−1.0
p_{12}	+1.0	+1.0	+.60	+.46	+.42	+.94	+.68	+.90	+1.0	−**1.0**	+1.0	-	±0.0
p_{13}	+1.0	+1.0	±0.0	±0.0	±0.0	+1.0	±0.0	+1.0	+1.0	±**0.0**	+1.0	±0.0	-

In Table 10 we may see the effect of this veto and counter-veto polarization. Many pairwise outrankings, like $p_3 \, \widetilde{\mathrm{S}} \, p_1$ or $p_1 \, \widetilde{\mathrm{S}} \, p_3$ appear now, either certainly validated, or, certainly invalidated. Take poster p_{10} for instance. It outranks now the other posters with certitude, except poster p_4, where, nevertheless, the polarized validation is highly significant (88%). In this large-performance-differences (LPD) polarized outranking graph, poster p_{10} becomes on the one hand, an even more convincing Condorcet winner. Whereas, poster p_{13} on the other hand, does no more positively outrank poster p_{10} ($r(p_{13} \, \widetilde{\mathrm{S}} \, p_{10}) = 0.0$), and so does no more qualify as second Condorcet winner.

The LPD[22] polarization induces four strong components:

$$\{p_{10}\} \succ \{p_1, p_3, p_4, p_5, p_6, p_7, p_8, p_{12}, p_{13}\} \succ \begin{cases} \{p_2, p_9\} \\ \{p_{11}\} \end{cases}$$

where poster p_{10} remains, as CONDORCET winner, the singleton top strong component. However, the so collapsed CONDORCET graph shows a partial weak order instead of the previous linear ordering (see Figure 5.c). The two worst strong components $\{p_2, p_9\}$ and $\{p_{11}\}$ appear now mutually incomparable.

The kernel extraction delivers now three solutions: $\{p_{10}\}$ as *outranking* kernel with a dominance significance of 88%, and two overlapping *outranked* kernels: $\{p_9, p_{11}\}$ and $\{p_2, p_{11}\}$ with absorbency significance of 61%, respectively 55%.

Poster p_{10} is, under these working hypotheses, even more convincingly to be recommended for getting the EBPA. But are we not fooled by weakly significant validations and invalidations of global weighted outranking situations?

Requiring a qualified significance level.
Until now, we have indeed considered that a simple majority of weighted significance is sufficient for validating, respectively invalidating, a given global outranking situation. Let us for one moment be more suspicious and require instead a qualified majority of at least 75% significance. Translated into bipolar characteristic terms, we will require a bipolar characteristic value of at least 0.5, respectively at most −0.5, for validating, respectively invalidating, the global weighted outranking relation, a situation we will denote $\mathrm{S}_{75\%}$. We have consequently to adapt the definition of the associated crisp outranking relation from Equation 8 on all the couples (x, y) of posters as follows:

$$x\,\mathrm{S}_{75\%}\,y \text{ is } \begin{cases} \textit{true}\ (+) & \text{if} \quad r(x \succsim_{\mathbf{w}} y) \geqslant 0.5, \\ \textit{false}\ (-) & \text{if} \quad r(x \succsim_{\mathbf{w}} y) \leqslant -0.5, \\ \textit{indeterminate}\ (0) & \text{if} \quad -0.5 < r(x \succsim_{\mathbf{w}} y) < 0.5\,. \end{cases} \tag{14}$$

In Table 11, we show the corresponding 75% qualified significance denotation we obtain on the global weighted outranking relation $\widetilde{\mathrm{S}}$ under the working hypothesis of widened preference thresholds (2 points) and by taking into account the polarizing effects of large performance differences. Compared to Table 7, much more outranking situations get

[22] See the glossary at the end of chapter.

Table 11. The 75% significance qualified outranking relation

$x\ \mathrm{S}_{75\%}\ y$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1		0	−	−	−	+	0	+	+	−	+	−	−
p_2	0		−	−	−	0	−	0	0	−	0	−	−
p_3	+	+		0	0	+	+	+	0	0	+	0	+
p_4	+	+	+		+	+	+	+	0	0	+	+	+
p_5	+	+	0	0		+	+	+	+	0	+	0	0
p_6	0	0	0	−	0		0	+	+	−	+	0	−
p_7	+	+	0	0	0	+		+	+	−	+	0	0
p_8	0	0	−	−	−	+	0		0	−	+	−	−
p_9	−	0	0	0	−	−	−	0		−	0	−	−
$\mathbf{p_{10}}$	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**	**+**		**+**	**+**	**+**
p_{11}	0	0	−	−	−	0	−	0	0	−		−	−
p_{12}	+	+	+	0	0	+	+	+	+	−	+		0
p_{13}	+	+	0	0	0	+	0	+	+	0	+	0	

indeterminate now, like $p_1 \,\mathrm{S}\, p_2$ and $p_2 \,\mathrm{S}\, p_1$ for instance. Only poster p_{10} outranks all the other posters with a sufficiently high significance. Notice that poster p_{13} does not anymore yield an alternative Condorcet winner besides p_{10}. Indeed, $p_{13} \,\mathrm{S}\, p_{10}$ can now no more be validated. We thus obtain one outranking singleton kernel: $\{p_{10}\}$, and one outranked kernel $\{p_2, p_9, p_{11}\}$. Posters p_2, p_9 and p_{11} appear clearly outranked at this qualified significance level.

All the preceding analysis evidently depends on the numeric significance weight vector $\mathbf{w}$ we have chosen for computing the overall global outranking relation $\succsim_{\mathbf{w}}$. The organizers of the EBPA did not fix these weights, instead they only imposed a significance hierarchy of the selection criteria. Let us finally study the very impact of the choice of the significance weight vector $\mathbf{w}$ on our best choice recommendation.

Stability of the CONDORCET graph.
The question we must ask at this point is whether the bipolar characterisation of the global outranking may not appear as an artifact induced by our more or less arbitrarily chosen cardinal significance weights: $w_{sq} = 0.4$, $w_{tp} = 0.3$, $w_{or} = 0.2$,and $w_{pq} = 0.1$?

Let $\mathcal{W}$ denote the set of all possible weights vectors we may define on a family F of criteria. Let $\geqslant_{\mathbf{w}}$ be a significance preorder[23] associated with F via the natural $\geqslant$ relation on the significance values in the given weight vector $\mathbf{w}$. The symmetric part $=_{\mathbf{w}}$ of the relation $\geqslant_{\mathbf{w}}$ induces s ordered equi-significance classes, denoted $\Pi^{\mathbf{w}}_{(1)} <_{\mathbf{w}} \ldots <_{\mathbf{w}} \Pi^{\mathbf{w}}_{(s)}$, with

[23] As classically done, $>_{\mathbf{w}}$ denotes the asymmetric part of $\geqslant_{\mathbf{w}}$, whereas $=_{\mathbf{w}}$ denotes its symmetric part.

$1 \leq s \leq |F|$. The criteria gathered in each equi-significance class have the same weight in $\mathbf{w}$ and, for $1 \leqslant i < j \leqslant s$, those of equi-significance class $\Pi^{\mathbf{w}}_{(i)}$ have a higher weight than those of class $\Pi^{\mathbf{w}}_{(j)}$. In our case here, we observe in fact $s = 4$ such equi-significance classes: one for each preference viewpoint f in F gathering the equi-significant opinions of all the five jury members[24].

Let $\mathcal{W}_{=_\mathbf{w}} \subset \mathcal{W}$ denote the set of all significance weights vectors that are compatible with the equivalence part $=_\mathbf{w}$. Let $\mathcal{W}_{\geqslant_\mathbf{w}} \subset \mathcal{W}$ denote the set of all significance weights vectors that are compatible with $\geqslant_\mathbf{w}$,and let $\mathbf{w} \in \mathcal{W}$. The CONDORCET *robustness denotation* (Bisdorff, 2004) of $\succsim_\mathbf{w}$, denoted $[\![\succsim_\mathbf{w}]\!]$, is defined, for all $(x,y) \in A \times A$, as follows:

$$
[\![x \succsim_\mathbf{w} y]\!] := \begin{cases}
4 & \text{if } r(x \succsim_\mathbf{v} y) = +1.0 \;\; \forall \mathbf{v} \in \mathcal{W}\,; \\
3 & \text{if } r(x \succsim_\mathbf{v} y) > 0.0 \;\; \forall \mathbf{v} \in \mathcal{W}_{=_\mathbf{w}}\,; \\
2 & \text{if } \big[r(x \succsim_\mathbf{v} y) > 0.0 \;\; \forall \mathbf{v} \in \mathcal{W}_{\geqslant_\mathbf{w}}\big] \\
 & \quad \wedge \big[\,\exists \mathbf{v}' \in \mathcal{W} \;:\; r(x \succsim_{\mathbf{v}'} y) < +1.0\,\big]\,; \\
1 & \text{if } \big[r(x \succsim_\mathbf{w} y) > 0.0\big] \;\wedge\; \big[\,\exists \mathbf{v}' \in \mathcal{W}_{\geqslant_\mathbf{w}} \;:\; r(x \succsim_{\mathbf{v}'} y) \leqslant 0.0\,\big]\,; \\
0 & \text{if } r(x \succsim_\mathbf{w} y) = 0.0\,; \\
-1 & \text{if } \big[r(x \succsim_\mathbf{w} y) < 0.0\big] \;\wedge\; \big[\,\exists \mathbf{v}' \in \mathcal{W}_{\geqslant_\mathbf{w}} \;:\; r(x \succsim_{\mathbf{v}'} y) \geqslant 0.0\,\big]\,; \\
-2 & \text{if } \big[r(x \succsim_\mathbf{v} y) < 0.0 \;\; \forall \mathbf{v} \in \mathcal{W}_{\geqslant_\mathbf{w}}\big] \\
 & \quad \wedge \big[\,\exists \mathbf{v}' \in \mathcal{W} \;:\; r(x \succsim_{\mathbf{v}'} y) > -1.0\,\big]\,; \\
-3 & \text{if } r(x \succsim_\mathbf{v} y) < 0.0 \;\; \forall \mathbf{v} \in \mathcal{W}_{=_\mathbf{w}}\,; \\
-4 & \text{if } r(x \succsim_\mathbf{v} (x,y)) = -1.0 \;\; \forall \mathbf{v} \in \mathcal{W}\,;
\end{cases}
\tag{15}
$$

with the following semantics:

- $[\![x \succsim_\mathbf{w} y]\!] = \pm 4$ if the jury *unanimously validates* (resp. *invalidates*) the outranking situation between x and y on all the selection criteria;
- $[\![x \succsim_\mathbf{w} y]\!] = \pm 3$ if a significant majority of the jury *validates* (resp. *invalidates*) the outranking situation between x and y for any significance weights of the selection criteria;
- $[\![x \succsim_\mathbf{w} y]\!] = \pm 2$ if a *significant majority* of the jury *validates* (resp. *invalidates*) the outranking situation between x and y for all $\geqslant_\mathbf{w}$-compatible significance weights;
- $[\![x \succsim_\mathbf{w} y]\!] = \pm 1$ if a significant weighted majority of criteria *validates* (respectively *invalidates*) this outranking situation for $\mathbf{w}$ but not for all $\geqslant_W$-compatible weights;
- $[\![x \succsim_\mathbf{w} y]\!] = 0$ if the total significance of the warranting criteria is *exactly balanced* by the total significance of the not warranting criteria for $\mathbf{w}$.

[24] See the complete set of global outrankings from each preference point of view in the Appendix

Let us start by presenting the notation which allows us to detail the construction of the CONDORCET robustness denotation associated with a valued outranking relation $\succsim_{\mathbf{w}}$ and a significance weights vector $\mathbf{w}$.

We recall that $r(\succsim_f)$ represents the sum of the jury's members opinions on preference point of view f. When changing the sign of $r(\succsim_f)$, we may as well represent the sum of the jury's members *negated* opinions on this preference point of view f. From this fact it follows that $r(x \succsim_{\mathbf{w}} y) > 0.0$ is verified for all $\mathbf{w} \in \mathcal{W}_{\mathbf{w}}$ if and only if $r(x \succsim_{\mathbf{w}} y) - r(x \not\succsim_{\mathbf{w}} y) > r(x \not\succsim_{\mathbf{w}} y) - r(x \succsim_{\mathbf{w}} y)$ is also verified (Bisdorff, 2004). The latter inequality gives us the operational key for implementing a test for the presence of a Condorcet robustness of degree ± 2. The same weights w_f and $-w_f$, denoting the "*affirmative*", respectively the "*refutative*", significance of each preference point of view, appear on each side of these inequalities. Furthermore, the sum of the coefficients $r(x \succsim_f x)$ and $r(x \not\succsim_f x)$ – that constitute the terms $r(x \succsim_{\mathbf{w}} y)$ and $r(x \not\succsim_{\mathbf{w}} y)$ – are equal for all couples (x, y) of posters. These coefficients may appear therefore as some kind of "*credibility*" distribution on the set of positive and negative significance weights.

To illustrate this insight, let us order the sequence $F_{\pm}$ of negative and positive preference points of view from the most significant negative one to the most significant positive one: $F_{\pm} := [-sq, -td, -or, -pq, pq, or, td, sq]$. Let us furthermore denote $r(x \succsim_{f_{(k)}} y)$, respectively $r(x \not\succsim_{f_{(k)}} y)$, for $(k) = (1), ..., (2s)$ indexing the ordered entries in the sequence $F_{\pm}$, the bipolar characteristics of the individual outranking situations gathered in the same equi-significance class $\Pi^{\mathbf{w}}_{f_{(k)}}$.

In the first line of Table 12, we may for instance, observe the distribution of $r(p_{10} \succsim_f p_1)$ over the ordered sequence $F_{\pm}$. The jury unanimously validates the outranking on all the selection criteria. Hence, $[\![p_{10} \succsim_{\mathbf{w}} p_1]\!] = +4$. The positive outranking results remains indeed valid with any significance weight vector, even one where the jury members would be attributed different significance weights. Consider now in the second part of Table 12, the distribution of $r(p_{10} \succsim_f p_4)$ and $r(p_{10} \not\succsim_f p_4)$ over the ordered sequence $F_{\pm}$. The jury unanimously validates again this outranking situation, but only on three of the four selection criteria. From the *Scientific Quality* point of view, however, only a majority of 60% validates it. Hence, $[\![p_{10} \succsim_{\mathbf{w}} p_4]\!] = +3$. The positive outranking results remains indeed valid with any significance weight vector where the the jury members are consider equi-significant.

Furthermore, let $C^{\mathbf{w}}_{f_{(k)}}(x, y) := \sum_{i=1}^{k} \left[r(x \succsim_{f_{(i)}} y) \right]$ be the cumulative sum of "*outranking*" characteristics for all preference points of view having significance at least equal to the one associated to $f_{(k)}$, and let

Table 12. Repartition of the bipolar characterisation $r(\succsim)$ into negative and positive arguments

$F_{\pm}$	$-sq$	$-tp$	$-or$	$-pq$	$+pq$	$+or$	$+tp$	$+sq$
$p_{10} \succsim_{(f)} p_1$	0	0	0	0	**1.0**	**1.0**	**1.0**	**1.0**
$p_{10} \succsim_{(f)} p_4$	0	0	0	0	1.0	1.0	1.0	**0.2**
$p_{10} \not\succsim_{(f)} p_4$	**0.2**	1.0	1.0	1.0	0	0	0	0
$p_1 \succsim_{(f)} p_2$	0	0	0	0.2	0	0.2	0.2	0.2
$p_1 \not\succsim_{(f)} p_2$	0.2	0.2	0.2	0	0.2	0	0	0
$C^{\mathbf{w}}_{(k)}(p_1, p_2)$	0	0	0	**0.2**	**0.2**	**0.4**	**0.6**	**0.8**
$\overline{C^{\mathbf{w}}_{(k)}}(p_1, p_2)$	**0.2**	**0.4**	**0.6**	**0.6**	**0.8**	**0.8**	**0.8**	**0.8**
$p_4 \succsim_{(f)} p_{10}$	0	1.0	0.6	1.0	0	0	0	0.2
$p_4 \not\succsim_{(f)} p_{10}$	0.2	0	0	0	1.0	0.6	1.0	0
$C^{\mathbf{w}}_{(k)}(p_4, p_{10})$	**0**	**1.0**	1.6	2.6	2.6	2.6	2.6	2.8
$\overline{C^{\mathbf{w}}_{(k)}}(p_4, p_{10})$	**0.2**	**0.2**	0.2	0.2	1.2	1.8	2.8	2.8

$\overline{C^{\mathbf{w}}_{f(k)}}(x, y) := \sum_{i=1}^{k} \big[- r(x \succsim_{f_{(i)}} y) \big]$ be the same cumulative sum of the negation of these characteristics.

In the third part of Table 12, we may see these cumulative repartition for the comparison of posters p_1 and p_2. As $C^{\mathbf{w}}_{(k)}(p_1, p_2)$ for $f_{(k)}$ in $F_{\pm}$ is strictly lower than the cumulative repartition of $\overline{C^{\mathbf{w}}_{(k)}}(p_{10}, p_4)$, we are thus sure that $r(p_{10} \succsim_{\mathbf{w}} p_4)$ will stay strictly positive for all $\mathbf{v} \in \mathcal{W}_{\mathbf{w}}$. Hence, $[\![p_1 \succsim_{\mathbf{w}} p_2]\!] = +2$. This ± 2 denotation test of Proposition 2 corresponds in fact to the verification of *stochastic dominance*-like conditions (see Bisdorff, 2004). And, in the absence of a ± 4 or ± 3 denotation, the following proposition gives us the corresponding test for the presence of a ± 2 denotation:

Proposition 2 (Bisdorff (2004) Label prop:condorcetRobustness)**.** *Let $\succsim_{\mathbf{w}}$ represent the global weighted outranking relation obtained with significance weights vector $\mathbf{w}$.*

$$[\![x \succsim_{\mathbf{w}} y]\!] = +2 \quad \Leftrightarrow \quad \begin{cases} \forall k \in 1, ..., s : \; C^{\mathbf{w}}_{f(k)}(x, y) \; \leqslant \; \overline{C^{\mathbf{w}}_{f(k)}}(x, y) \,; \\ \exists k \in 1, ..., s : \; C^{\mathbf{w}}_{f(k)}(x, y) \; < \; \overline{C^{\mathbf{w}}_{f(k)}}(x, y). \end{cases} \qquad (16)$$

The respective negative degree $[\![x \succsim_{\mathbf{w}} y]\!] = -2$ may be checked with similar conditions using reversed inequalities.

A ± 1 Condorcet robustness denotation, corresponding to the observation of a weighted majority (resp. minority) in the absence of the

± 2 case, is simply verified as follows:

$$[\![x \succsim_{\mathbf{w}} y]\!] = \pm 1 \iff \big((x \succsim_{\mathbf{w}} y) \gtrless 0.0 \big) \wedge [\![x \succsim_{\mathbf{w}} y]\!] \neq \pm 2 \big). \quad (17)$$

This situation is illustrated in the fourth part of Table 12, where we may notice that the cumulative repartition of the bipolar characterisation of $p_4 \succsim p_{10}$ is neither strictly lower nor strictly greater than its negation. Hence, $[\![p_4 \succsim_{\mathbf{w}} p_{10}]\!] \neq \pm 2$. The apparent result that p_4 does not outranks poster p_{10} – $r(p_4 \succsim_{\mathbf{w}} p_{10}) = -0.44$ (see Table 6 in Section 2.3) is thus not stable for all **w**-order compatible significance vectors and $[\![p_4 \succsim_{\mathbf{w}} p_{10}]\!]$ is equal to -1.

The Condorcet robustness degrees of the global outranking statements $\succsim_{\mathbf{w}}$ for all couples of competing posters are shown in Table 13. We notice now that the previous best choice recommendation, namely

Table 13. Condorcet robustness degrees of the weighted outranking relation

$[\![x \succsim_{\mathbf{w}} y]\!]$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	$\mathbf{p_{10}}$	p_{11}	p_{12}	p_{13}
p_1		+2	−3	−3	−1	+2	−2	+2	+3	**−3**	+3	−3	−2
p_2	−2		−2	−2	−2	−2	−2	−3	+3	**−3**	−1	−2	−3
p_3	+3	+3		+1	+1	+3	+3	+3	+3	**−3**	+3	+1	−1
p_4	+3	+3	+2		+3	+3	+3	+3	+3	**−1**	+3	+2	+2
p_5	+3	+3	+2	−1		+4	+3	+3	+3	**−3**	+3	+1	+0
p_6	+1	+2	−3	−3	−2		−2	+3	+3	**−4**	+3	−3	−3
p_7	+2	+2	−3	−1	−3	+3		+3	+3	**−3**	+3	+0	−3
p_8	+0	+3	−3	−3	−3	+2	−3		+3	**−3**	+3	−2	−3
p_9	−3	−3	−3	−3	−3	−3	−2	−3		**−3**	−2	−2	−3
$\mathbf{p_{10}}$	**+4**	**+3**	**+4**	**+3**	**+4**	**+4**	**+3**	**+4**	**+3**		**+3**	**+3**	**+3**
p_{11}	+2	+2	−3	−3	−3	−2	−3	+2	+3	**−3**		−3	−2
p_{12}	+3	+2	+2	+1	+3	+3	+3	+2	+3	**−1**	+3		+2
p_{13}	+4	+3	+3	+1	+3	+3	+3	+3	+3	**−1**	+3	+2	

poster p_{10} becomes positively confirmed. Indeed, with a robustness degree of at least +3, i.e. *positively outranking with any* $=_{\mathbf{w}}$*-compatible weight vector* , i.e. even totally independent of any significant differentiation of the preference points of view, poster p_{10} is definitely confirmed as the unique *robust* Condorcet winner. There is even evidence that p_{10} *unanimously outranks* posters p_1, p_3, p_5, p_6 and p_8. Inspecting Column p_{10} of the same Table 13, we may furthermore notice that no other poster positively outranks p_{10}. The jury members even unanimously invalidate the statement that p_6 and p_{11} might outrank p_{10}. The Condorcet robustness analysis shows, by the way, that poster p_4 is, apart from p_{10}, positively outranking all other competing poster with any $\geqslant_{\mathbf{w}}$-compatible weight vector. Finally, poster p_9 is definitely confirmed to be outranked with any $\geqslant_{\mathbf{w}}$-order compatible weight vector, and so can be rejected with good reasons, even if it has not been evaluated by some jury members.

Conclusions

To conclude this long methodological study of the EBPA case, let us enumerate some remarks, first, on the case itself, and later more generally concerning our approach to MCDA decision aid practice.

1. **Auditing**: Regarding the output of the historical decision making process, the actual decision to attribute the EBPA to poster p_10 can be transparently legitimated, both from the preference modelling – , as well as, from the best unique choice recommendation viewpoint. Computing the underlying global outranking relation and the corresponding best choice recommendation requires, apart from the official set of selection criteria with its significance preorder pre-settled institutionally by the EBPA organizer, no further model parameters. The only clearly needed information here are in fact the individual ordinal performance assessments delivered by the jury members with respect to the officially recommended selection criteria. This cognitive task corresponds well, however, with their scientific and professional qualification. Indeed, the official nomination into the jury is precisely based on this reputation and guarantees therefore the official expert status of the jury members.
2. **MAVT approach**: Would a value theory based approach, in this case, do the same job? This is doubtful for two reasons:
   (a) The actual decision making process shows that the EBPA jury has to come to its conclusion after all the poster sessions have happened and before the actual closing session begins where the winner has to be announced. This leaves very little time – a single physical meeting of the jury members – to elicit all the cognitively complex model parameters like swing weights or value trade offs between the selection criteria which would, the case given, be needed for a conjoint measurement of the overall performances.
   (b) The presence of some missing evaluations represents furthermore an irreducible problem for all value theory approach. By nature, the value theory approach can indeed not positively take into account non existing evaluations. Either, partially evaluated posters would have to be dropped from the contest (a solution difficult to legitimate by the EBPA jury), or, a

fictive neutral value would have to be artificially fixed to fill in the missing values.

The social choice theory approach, that underlies the outranking methodology, fits, on the contrary, quite well here with the actual group decision problem the EBPA jury has to tackle. Achieving or validating decisions by implicit or explicit voting procedures is quite acceptable in our culture and the five jury members are by nature to be considered equi-significant for the selection of the best poster. A jury member, considered as a voter, may, by the way, abstain himself from delivering his opinion. This feature, present in every practical voting system, is effectively available in the Rubis outranking approach and naturally allows for coherently tackling missing performance assessments.

3. **Ranking versus selecting**: Finally, it is interesting to compare our approach with the case study concerning the choice of a cooling system for a power plant (see Pirlot et al., 2011, Chapter XX). There the best choice decision problem is treated as a ranking problem with the argument that the output of ranking methods is *richer*. Clearly, a value theory approach will not make the difference between selection and ranking procedures as a total ranking will anyway be available beforehand to the selection procedure via the global value assessments of the decision alternatives. In the outranking approach, however, things are more subtle in the sense that constructing a complete ranking may in practice need much more preferential information (in order to be richer as claimed before) than a direct best choice selection procedure like the Rubis method. But this additional preferential information, needed for a complete ranking of the decision alternatives, generally represents the most doubtful part of the preference modelling assumptions, especially when the multiple criteria don't reveal a trivial concordance for easily rendering a global ranking. Here we painfully recover Arrow's impossibility theorem in the sense that making globally concordant a family of, otherwise discordant, criteria definitely needs a strong arbiter, i.e. a dictator principle generally hidden behind non trivial model parameters, which induces insidiously the requested, complete and transitive, global preference. As put to the point by Roy[25], it is precisely the parsimony and

[25] "... The goal of our research was to design a resolution method ... that is easy to put into practice, that requires as few and reliable hypotheses as possible, and that meets the needs [of the decision maker] ..." Roy et al. (1966).

the simplicity of the preference modelling parameters that represent the practical advantage of an outranking approach, and in particular of the Rubis best choice method when applied to this case.

Appendix: The complete performance tableau

The evaluation of the 13 competing posters on all four selection criteria by all the jury members are expressed in Table A.1 on a common ordinal performance scale from 0 (lowest) to 10 (highest). The *slash* symbol (/) represents not available evaluations (see the corresponding paragraph in Section 1.3) at the moment where the award jury had to select the award winner.

Table A.1. Evaluation marks given by all the jury members on all the competing posters

Poster ID	Scientific quality					Theory or practice of OR					Originality					Presentation quality				
	j_1	j_2	j_3	j_4	j_5	j_1	j_2	j_3	j_4	j_5	j_1	j_2	j_3	j_4	j_5	j_1	j_2	j_3	j_4	j_5
p_1	4	7	5	5	3	4	7	6	5	3	4	6	6	7	3	4	7	5	6	2
p_2	/	1	6	2	/	/	1	7	3	/	/	1	8	3	/	/	3	9	7	/
p_3	6	6	7	6	2	8	9	7	6	4	6	7	7	7	5	6	6	9	7	5
p_4	8	9	9	8	6	7	8	6	7	4	8	8	7	7	4	8	6	7	7	6
p_5	8	6	8	7	2	8	7	9	7	0	8	5	7	7	2	8	8	8	6	5
p_6	5	5	5	6	2	5	7	5	5	0	5	5	5	6	2	5	7	6	5	5
p_7	6	5	6	6	/	7	8	7	6	/	6	5	5	6	/	8	8	5	3	/
p_8	4	/	5	6	2	4	/	5	6	0	4	/	7	5	2	7	/	10	5	4
p_9	/	/	5	3	/	/	/	5	3	/	/	/	7	3	/	/	/	10	3	/
p_{10}	9	9	8	8	4	9	9	9	7	6	9	9	9	7	7	9	10	10	8	7
p_{11}	6	9	8	7	5	6	8	6	6	5	6	9	7	8	5	8	9	8	7	3
p_{12}	4	5	7	5	/	4	5	7	5	/	4	3	7	5	/	4	5	3	3	/
p_{13}	4	8	8	8	8	4	8	8	7	10	4	6	7	7	8	4	9	9	8	10

The complete performance table is published in XMCDA-2.0 format under the name `bpaeuro20.xml` on the archive site of the Handbook[26]. An extract of the XMCDA-2.0 encoding is shown hereafter:

```
<?xml version="1.0" encoding="UTF-8"?>
<xmcda:XMCDA
    xmlns:xmcda=
    "http://www.decision-deck.org/2009/XMCDA-2.0.0"
    instanceID="void">

   <projectReference id="bpaeouro20" name="bpaeouro20.xml">
       <title>The EURO 20 Best Poster Award</title>
       <author>Raymond Bisdorff</author>
       <version>D2 MCDA Applications Book</version>
   </projectReference>

   <alternatives mcdaConcept="alternatives">
       <description>
         <subTitle>List of competing posters.</subTitle>
       </description>
       <alternative id="p01" name=""
                    mcdaConcept="potentialDecisionAction">
          <description>
```

[26]see http://decision-deck.org/MCDA_Applications_Handbook. XXX

```
          <comment>submitted poster</comment>
        </description>
        <type>real</type>
        <active>true</active>
      </alternative>
      ...
      ...
    </alternatives>

    <criteria mcdaConcept="criteria">
      <description>
        <subTitle>Family of criteria</subTitle>
      </description>

      <criterion id="or2"
                 name="or2"
                 mcdaConcept="criterion">
        <description>
          <comment>
            Originality evaluated by jury member j-2
          </comment>
          <version>performance</version>
        </description>
        <active>true</active>
        <criterionValue>
          <value><real>0.04</real></value>
        </criterionValue>
        <scale>
          <quantitative>
            <preferenceDirection>max</preferenceDirection>
            <minimum><real>0.0</real></minimum>
            <maximum><real>10.0</real></maximum>
          </quantitative>
        </scale>
        <thresholds>
          <threshold id="pref"
                     name="preference"
                     mcdaConcept=
           "performanceDiscriminationThreshold">
            <linear>
              <slope><real>0.0</real></slope>
              <intercept><real>1.0</real></intercept>
            </linear>
          </threshold>
        </thresholds>
      </criterion>
      ...
      ...
    </criteria>
    <performanceTable mcdaConcept="performanceTable">
      <description>
        <subTitle>Rubis Performance Table.</subTitle>
      </description>
      <alternativePerformances>
        <alternativeID>p01</alternativeID>
        <performance>
          <criterionID>or2</criterionID>
          <value><real>6.00</real></value>
        </performance>
        ...
        ...
      </alternativePerformances>
    </performanceTable>
</xmcda:XMCDA>
```

Appendix: Overall outranking per preference viewpoint

The following tables may be computed with the `digraphs` Python module shown in Listing B.1

Listing B.1. Computing with the Python `digraphs` module

```
#!/usr/bin/env python
# RB March 2010
# MCDA Applications Handbook
# Chapter on BPA Euro 2004: Appendix B
#----------------------------------------

from digraphs import XMCDA2PerformanceTableau, BipolarOutrankingDigraph

# load the complete performance tableau from file bpaeuro20.xml
T = XMCDA2PerformanceTableau('bpaeuro20')

# gather the individual preference viewpoints from all 5 jury members
Sq = ['sq1', 'sq2', 'sq3', 'sq4', 'sq5']
Tp = ['tp1', 'tp2', 'tp3', 'tp4', 'tp5']
Or = ['or1', 'or2', 'or3', 'or4', 'or5']
Pq = ['pq1', 'pq2', 'pq3', 'pq4', 'pq5']

# gather the family F of preference viewpoints
F = [Sq, Tp, Or, Pq]

# generate and show the outranking relation Sf per viewpoint f in F
for f in F:
    print 'Global outranking relation from viewpoint %s' %f
    Sf = BipolarOutrankingDigraph(T, coalition=f)
    Sf.recodeValuation(-1.0,1.0)
    Sf.showRelationTable()
```

Table B.1. Compairing the posters from the *Scientific Quality* (*sq*) point of view

$r(\succsim_{sq})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	-	+0.2	−0.2	−1.0	−0.2	+0.2	−0.4	+0.6	+0.4	−1.0	+0.2	−1.0	−0.6
p_2	−0.2	-	−0.6	−0.6	−0.6	−0.2	−0.2	−0.2	0	−0.6	−0.4	−0.6	−0.6
p_3	+0.2	+0.6	-	−0.6	−0.2	+1.0	+0.8	+1.0	+0.4	−1.0	+0.6	−0.2	−0.6
p_4	+1.0	+0.6	+0.6	-	+0.6	+1.0	+0.8	+1.0	+0.4	+0.2	+0.6	+0.6	+0.2
p_5	+0.2	+0.6	+1.0	−0.2	-	+1.0	+0.8	+1.0	+0.4	−0.2	+0.6	+0.2	+0.2
p_6	+0.2	+0.2	−0.6	−0.6	−0.6	-	0	+1.0	+0.4	−1.0	+0.2	−1.0	−0.6
p_7	+0.4	+0.6	−0.4	−0.4	−0.8	+0.8	-	+0.8	+0.4	−0.8	+0.2	−0.4	−0.4
p_8	+0.2	+0.2	−0.6	−0.6	−0.6	+0.6	0	-	+0.4	−1.0	+0.2	−1.0	−0.6
p_9	0	0	−0.4	−0.4	−0.4	0	−0.4	0	-	−0.4	−0.4	−0.4	−0.4
p_{10}	+1.0	+0.6	+1.0	+0.2	+1.0	+1.0	+0.8	+1.0	+0.4	-	+0.6	+0.6	+0.6
p_{11}	+0.6	+0.4	−0.2	−0.6	−0.6	−0.2	−0.2	+0.2	+0.4	−0.6	-	−0.6	−0.2
p_{12}	+1.0	+0.6	+1.0	−0.2	+0.2	+1.0	+0.8	+1.0	+0.4	+0.2	+0.6	-	+0.2
p_{13}	+1.0	+0.6	+0.6	−0.2	+0.6	+0.6	+0.4	+1.0	+0.4	+0.2	+0.6	+0.2	-

Table B.2. Compairing the posters from the *Contribution to OR Theory and/or practice* (*tp*) point of view

$r(\succcurlyeq_{sq})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	-	+0.2	−1.0	−0.6	+0.2	+0.6	−0.8	+0.6	+0.4	−1.0	+0.2	−0.6	−0.6
p_2	−0.2	-	−0.2	−0.2	−0.2	−0.2	−0.2	−0.2	+0.4	−0.6	0	−0.2	−0.6
p_3	+1.0	+0.6	-	+1.0	+1.0	+1.0	+0.8	+1.0	+0.4	−0.2	+0.6	+0.6	+0.2
p_4	+1.0	+0.2	−0.6	-	+0.2	+1.0	+0.4	+1.0	+0.4	−1.0	+0.2	+0.6	−0.2
p_5	+0.2	+0.2	−0.2	−0.2	-	+1.0	0	+1.0	+0.4	−0.6	+0.2	−0.2	−0.2
p_6	+0.2	+0.2	−1.0	−1.0	+0.2	-	−0.8	+0.6	+0.4	−1.0	+0.2	−1.0	−0.6
p_7	+0.8	+0.6	−0.4	+0.8	0	+0.8	-	+0.8	+0.4	−0.8	+0.6	+0.8	0
p_8	−0.2	+0.2	−1.0	−0.6	−0.2	+0.2	−0.4	-	+0.4	−1.0	+0.2	−0.6	−0.6
p_9	−0.4	0	−0.4	−0.4	0	0	−0.4	0	-	−0.4	−0.4	−0.4	−0.4
p_{10}	+1.0	+0.6	+1.0	+1.0	+1.0	+1.0	+0.8	+1.0	+0.4	-	+0.6	+1.0	+0.6
p_{11}	+0.6	+0.4	−0.2	−0.2	−0.2	+0.2	−0.2	+0.2	+0.4	−0.6	-	−0.2	−0.2
p_{12}	+1.0	+0.2	−0.6	+0.6	+0.2	+1.0	0	+1.0	+0.4	−1.0	+0.2	-	−0.2
p_{13}	+1.0	+0.6	+0.2	+0.6	+0.6	+0.6	+0.4	+1.0	+0.4	−0.2	+0.6	+0.6	-

Table B.3. Compairing the posters from the *Originality* (*or*) point of view

$r(\succcurlyeq_{sq})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	-	+0.2	−0.6	−0.6	+0.2	+0.6	+0.4	+0.6	0	−0.6	+0.2	−1.0	+0.2
p_2	−0.2	-	−0.2	−0.2	−0.2	−0.2	−0.2	−0.2	+0.4	−0.2	0	−0.2	−0.2
p_3	+1.0	+0.2	-	+0.2	+0.6	+1.0	+0.8	+1.0	+0.4	−0.6	+0.6	+0.2	+0.6
p_4	+1.0	+0.2	+0.6	-	+1.0	+1.0	+0.8	+1.0	+0.4	−0.6	+0.6	−0.2	+0.6
p_5	+0.2	+0.2	+0.2	+0.2	-	+1.0	+0.8	+1.0	+0.4	−0.6	+0.6	−0.2	+0.2
p_6	−0.6	+0.2	−1.0	−1.0	−0.2	-	+0.4	+0.6	0	−1.0	+0.2	−1.0	−0.6
p_7	−0.4	+0.2	−0.4	−0.8	−0.4	+0.8	-	+0.4	0	−0.8	+0.2	−0.4	−0.4
p_8	−0.2	+0.2	−0.6	−0.6	−0.2	−0.2	−0.4	-	+0.4	−1.0	+0.6	−0.6	−0.2
p_9	0	0	0	0	0	0	0	-	0	−0.4	0	0	0
p_{10}	+1.0	+0.6	+1.0	+1.0	+1.0	+1.0	+0.8	+1.0	+0.4	-	+0.6	+0.6	+0.6
p_{11}	+0.2	0	−0.2	−0.2	−0.2	−0.2	−0.2	+0.6	+0.4	−0.6	-	−0.2	+0.2
p_{12}	+1.0	+0.2	+1.0	+0.6	+0.6	+1.0	+0.8	+1.0	+0.4	−0.2	+0.6	-	+0.6
p_{13}	+1.0	+0.2	+0.2	+0.2	+0.6	+0.6	+0.4	+1.0	+0.4	−0.2	+0.6	−0.2	-

Table B.4. Compairing the posters from the *Presentation Quality* (*pq*) point of view

$r(\succcurlyeq_{sq})$	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}
p_1	0	−0.2	−0.2	−0.6	−1.0	−0.2	0	−0.2	+0.2	−1.0	+0.8	−0.6	−0.6
p_2	+0.2	0	+0.2	+0.2	+0.2	+0.2	+0.2	−0.2	+0.2	−0.6	+0.6	+0.2	−0.2
p_3	+0.6	+0.2	0	−0.2	−0.2	+0.6	0	+0.2	+0.2	−1.0	+0.8	+0.2	−0.2
p_4	+0.6	+0.2	+0.6	0	+0.2	+0.6	+0.4	+0.6	+0.2	−1.0	+0.8	+0.2	−0.6
p_5	+1.0	+0.2	+0.6	+0.6	0	+1.0	+0.8	+0.6	+0.2	−1.0	+0.8	+0.6	−0.6
p_6	+0.6	−0.2	−0.6	−0.6	−1.0	0	0	+0.2	+0.2	−1.0	+0.8	−0.2	−0.6
p_7	+0.4	−0.2	0	0	0	0	0	0	+0.2	−0.8	+0.8	0	−0.4
p_8	+0.2	+0.2	−0.2	−0.6	−0.6	+0.6	0	0	+0.6	−0.6	+0.8	+0.2	−0.2
p_9	−0.2	−0.2	−0.2	−0.2	−0.2	−0.2	+0.2	−0.2	0	−0.2	+0.6	+0.2	−0.2
p_{10}	+1.0	+0.6	+1.0	+1.0	+1.0	+1.0	+0.8	+1.0	+0.6	0	+0.8	+1.0	+0.6
p_{11}	−0.4	−0.6	−0.8	−0.8	−0.8	−0.8	−0.4	−0.8	+0.2	−0.8	0	−0.4	−0.4
p_{12}	+0.6	−0.2	−0.2	+0.2	+0.2	+0.2	+0.8	−0.2	+0.2	−1.0	+0.8	0	−0.2
p_{13}	+1.0	+0.6	+0.6	+0.6	+0.6	+0.6	+0.4	+0.2	+0.2	−0.2	+0.8	+0.6	0

GLOSSARY

Abbreviations and terms

BCR : Best choice recommendation promoted by the Rubis decision aid methodology (see Bisdorff et al. (2008)). It consists in a minimal, strict outranking, stable and maximal determined bipolar-valued choice set, computed directly on the given bipolar-valued outranking graph.

DECISION DECK : The Decision Deck project aims at collaboratively developing Open Source software tools implementing Multiple Criteria Decision Aid (MCDA). Its purpose is to provide effective tools for three types of users:
– practitioners who use MCDA tools to support actual decision makers involved in real world decision problems;
– teachers who present MCDA methods in courses, for didactic purposes;
– researchers who want to test and compare methods or to develop new ones.
More information may be found on the the official web site of the Decision Deck Project: http://www.decision-deck.org.

EBPA : Euro 2004 Best Poster Award (see Bisdorff (2004)).

ELECTRE : Multiple criteria decision aid methods, originating from seminal work of B. Roy (early seventies) with contributions from numerous PhD students and senior researchers that collaborated with him during the eighties and nineties when visiting his Laboratoire d'Analyse et de Modélisation des Systèmes d'aide à la Décision (Lamsade) (see http://www.lamsade.dauphine.fr) at Université Paris-Dauphine.

EURO : The Association of European OR Societies (see http://www.euro-online.org).

Kernel : An *outranking* (respectively *outranked*) *kernel* in a directed outranking graph corresponds to a *dominant* (respectively an *absorbent*) and *independent* subset of decision alternatives. The dominant version, also called *game solution*, is due to Von Neumann and Morgenstern (1944), whereas the absorbent version is due to Berge. The name "noyau" (kernel) proposed by Berge (1958), probably stems from the zero values (algebraic kernels) of the Grundy function which deliver internally and externally stable solutions for NIM like game (see Berge, 1962).

LPD : Large Performance Differences. The LPD polarized characterisation of an outranking situation allows to take into account large performance differences when assessing its validation.

RUBIS : Multiple criteria decision aid method for selecting the unique best decision alternative from a bipolar-valued outranking digraph (see Bisdorff et al., 2008).

XMCDA : Standard XML encoding norm for MCDA Applications data (see the corresponding chapter in this handbook).

Symbols

A : set of conference posters competing for the best poster award.

J : Index set for the award jury members j_k, $k = 1, \ldots, 5$, nominated for selecting the best poster.

F : Official set of EBPA selection criteria: *Scientific Quality* (sq), *Contribution to OR theory and/or practice* (tp), *Originality* (or), and *Presentation Quality* (pq). Each $f \in F = \{sq, tp, or, pq\}$ is also called a *preference viewpoint*.

$\mathbf{x \geqslant^{j}_{f} y}$: Individual statement of jury member j that poster x is at least as good as poster y with respect to preference viewpoint f (see Equation 2).

$\mathbf{x} \succcurlyeq^{\mathbf{j}} \mathbf{y}$: Overall statement of jury member j that poster x is at least as good as poster y with respect to all the given selection criteria (see Equation 4).

$\mathbf{x} \succcurlyeq_{\mathbf{f}} \mathbf{y}$: Global statement of all jury members that poster x is at least as good as poster y with respect to preference point of view f (see Equation 3).

$\mathbf{x} \succsim \mathbf{y}$: Global outranking statement of all jury members that poster x is at least as good as poster y with respect to all the given selection criteria. This situation is commonly referred as poster x globally *outranks* poster y (see Equation 6).

$\mathbf{C}(\mathbf{A}, \mathrm{S})$: The CONDORCET graph, i.e. the median cut crisp directed graph, associated with the bipolar-valued characterisation of the global outranking statements. $x \,\mathrm{S}\, y$, for $(x, y) \in A^2$, is true (respectively false) if the bipolar characteristic function $r(x \succsim_{\mathbf{w}} y)$ of the global outranking situation $x \succsim_{\mathbf{w}} y$ shows a significant weighted majority (respectively minority) of epistemic support considering the significance weights vector $\mathbf{w}$.

$\mathbf{x} \widetilde{\mathrm{S}} \mathbf{y}$: ELECTRE like outranking statement, polarizing the global outranking statement with *veto* and *counter-veto* effects (see Equations 12 and 13, Roy and Słowiński (2008)).

$\mathbf{r}(\mathbf{x} \,\mathrm{R}\, \mathbf{y})$: bipolar characteristic function of pairwise relational statement $x \,\mathrm{R}\, y$ with $\mathrm{R} \in \{\geqslant^j_f, \succcurlyeq_f, \succcurlyeq^j, \succsim, \widetilde{\mathrm{S}}\}$ defined on A and taking values in the rational interval $[-1.0, +1, 0]$. Positive values validate, whereas negative values invalidate, the relational $x \,\mathrm{R}\, y$ statement. The zero value signifies an indeterminate situation, i.e. where the relational $x \,\mathrm{R}\, y$ statement appears neither validated nor invalidated.

$[\![\mathbf{x} \succsim_{\mathbf{w}} \mathbf{y}]\!]$: CONDORCET robustness denotation associated with global weighted outranking relation $\succsim_{\mathbf{w}}$. See Definition 15.

Bibliography

Arrow, K. J. and Raynaud, H. (1986). *Social Choice and Multicriterion Decision-Making*, volume 1 of *MIT Press Books*. The MIT Press. 20

Barbut, M. (1980). Médianes, Condorcet et Kendall. *Mathématiques et Sciences Humaines*, 69:9–13. 20, 27, 29

Berge, C. (1962). *The Theory of Graphs (original title: The theory of graphs and its applications)*. Dover 2001 (originally published by Wiley 1962). 53

Bisdorff, R. (2002). Logical foundation of multicriteria preference aggregation. In Bouyssou, D., Jacquet-Lagrèze, E., Perny, P., Słowiński, R., Vanderpooten, D., and Vincke, P., editors, *Aiding Decisions with Multiple Criteria*, pages 379–403. Kluwer Academic Publishers. 15, 17, 19

Bisdorff, R. (2004). Concordant outranking with multiple criteria of ordinal significance. *4OR*, 2(4):293–308. 16, 41, 42, 43, 53

Bisdorff, R. (2008). On clustering the criteria in an outranking based decision aid approach. In Thi, H. A. L., Bouvry, P., and Pham, D., editors, *Computation and Optimization in Information Systems and Management Sciences*, CCIS, pages 409–418. Springer. 23

Bisdorff, R., Meyer, P., and Roubens, M. (2008). Rubis: a bipolar-valued outranking method for the best choice decision problem. *4OR: A Quarterly Journal of Operations Research*, 6(2):143 – 165. 8, 15, 17, 19, 33, 53

Bisdorff, R., Pirlot, M., and Roubens, M. (2006). Choices and kernels from bipolar valued digraphs. *European Journal of Operational Research*, 175:155–170. 8, 16, 33, 34

Fishburn, P. C. (1977). Condorcet social choice functions. *SIAM Journal on Applied Mathematics*, 33(3):469–489. 20

Pirlot, M., Teghem, J., Ulungu, B., Duvivier, L., Bulens, P., and Goffin, G. (2011). Choosing a cooling system for a power plant in belgium. In Bisdorff, R., Dias, L., Mousseau, V., and Pirlot, M., editors, *Evaluation and Decision Models with Multiple Criteria: Case Studies*. Spinger. 46

Roy, B. (1985). *Méthodologie multicritère d'aide à la décision*. Economica, Paris. 33

Roy, B. (1991). The outranking approach and the foundations of electre methods. *Theory and Decision*, 31(1):49–73. 26, 36, 37

Roy, B., Benayoun, R., and B., S. (1966). Electre: une méthode pour guider le choix en présence de points de vue multiples. Technical Report 49, Société d'Economie et de Mathématique Appliqué, Direction Scientifique. 46

Roy, B. and Bouyssou, D. (1993). *Aide Multicritère à la Décision : Méthodes et Cas*. Economica, Paris. 16, 26, 30

Roy, B. and Słowiński, R. (2008). Handling effects of reinforced preference and counter-veto in credibility of outranking. *European Journal of Operational Research*, 188(1):185–190. 36, 54

Sen, A. K. (1986). Social choice theory. In *in K.J. Arrow and M.D. Intriligator (Eds.), Handbook of Mathematical Economics Vol. III*, pages 1073–1181. 20