Comparing sets of positive and negative arguments: Empirical assessment of seven qualitative rules

Jean-François Bonnefon [1] and **Hélène Fargier**[2]

Abstract. Many decisions can be represented as bipolar, qualitative sets of arguments: Arguments can be pros or cons, and ranked according to their importance, but not numerically evaluated. The problem is then to compare these qualitative, bipolar sets. In this paper (a collaboration between a computer scientist and a psychologist), seven procedures for such a comparison are empirically evaluated, by matching their predictions to choices made by 62 human participants on a selection of 33 situations. Results favor cardinality-based procedures, and in particular one that allows for the internal cancellation of positive and negative arguments within a decision.

1 Introduction

Would you rather go to Rome or Beijin next spring? There are many arguments for or against each choice, and some may be more compelling than others, although it will be difficult to scale them on a common metrics. Maybe you really want to see the Coliseum, as much as you want to spend a night at Beijin opera house. You are somewhat concerned that you cannot speak Chinese. Language will also be an issue in Rome, but arguably less of one. How to decide?

Comparing the two decisions amounts to comparing two sets of arguments. Putting aside arguments that are irrelevant to the decision at hand, arguments can be positive (pros – typically, the good consequences of the decision), or negative (cons – typically, the bad consequences of the decision). Some of these arguments will be stronger or more compelling than others; however, decisions often have to be made on the basis of an ordinal ranking of the strength of the arguments, rather than on a numerical evaluation: some decision processes are of a qualitative nature. What is needed is thus a qualitative, bipolar approach to the comparison of sets of arguments. Meeting these two requirements would improve the cognitive plausibility of the method, as psychologists have argued that human choice processes are likely to be qualitative [11] and bipolar [5, 6, 15].

Ordinal ranking procedures from bipolar information have received scarce attention so far. Amgoud et al. [1] compare decisions in terms of positive and negative arguments, using a complex scheme for evaluating the strength of arguments (which possess both a level of importance and a degree of certainty, and involve criteria whose satisfaction is a matter of degree). They then compare decisions using simple optimistic or pessimistic rules, independently of the polarity of the arguments. Benferhat and Kaci [12, 3] propose to merge all positive affects into a degree of satisfaction (using the max rule). If high, this degree does not play any role and the decision is based on negative affects (using Wald's principle, see below). If low, it is understood as a negative affect and merged with the other ones. Finally, Dubois and Fargier [8] introduce a number of qualitative, bipolar rules for comparing sets of arguments; all these rule follow the prospective work of [17] on order of magnitude calculus and thus receive an interpretation in terms of order of magnitude probabilities.

This article, which is a collaboration between a computer scientist and a psychologist, will present an extensive empirical assessment of the descriptive validity of these rules. It will not, however, thoroughly describe all their properties or include their axiomatic characterization, for which the reader can refer to [8].

2 The rules

[8] consider the simple situation where each possible decision d is assessed by a finite subset of arguments $C(d) \subseteq X$. X is the set of all possible arguments. Comparing decisions then amounts to comparing sets of arguments, i.e. subsets A, B of 2^X. X can be divided in three disjoint subsets: X^+ is the set of positive arguments, X^- the set of negative arguments, X^0 the set of indifferent ones. Any $A \subseteq X$ can likewise be partitioned: let $A^+ = A \cap X^+$, $A^- = A \cap X^-$, $A^0 = A \cap X^0$ be respectively the positive, negative and indifferent arguments of A.

Arguments can be of varying importance. In a purely qualitative, ordinal approach, the importance of arguments can be described on a totally ordered scale of magnitude $L = [0_L, 1_L]$, for example by the following function π:

$$\pi : X \mapsto L = [0_L, 1_L]$$

$\pi(x) = 0_L$ means that the decision maker is indifferent to argument x: this argument will not affect the decision process. The order of magnitude 1_L (the highest level of importance) is attached to the most compelling arguments that the decision-maker can consider. Finally, the *order of magnitude* $\mathrm{OM}(A)$ of a set A is defined as the highest of the order of magnitude of its elements:

$$\forall A \subseteq X, \mathrm{OM}(A) = \max_{x \in A} \pi(x)$$

2.1 The qualitative Pareto comparison $\succeq^{\mathrm{Pareto}}$

This rule compares the two sets of arguments as a problem of bicriteria decision. The first criterion compares negative arguments according to Wald's [16] rule: A is better than B on the negative side iff $\mathrm{OM}(A^-) \leq \mathrm{OM}(B^-)$. The second criterion compares positive arguments according to the optimistic counterpart of Wald's rule.

$$A \succeq^{\mathrm{Pareto}} B \iff \text{and} \begin{array}{l} \mathrm{OM}(A^+) \geq \mathrm{OM}(B^+) \\ \mathrm{OM}(A^-) \leq \mathrm{OM}(B^-) \end{array}$$

[1] LTC-CNRS, 5 allées Antonio Machado 31058 Toulouse Cedex 9, France. E-mail: bonnefon@univ-tlse2.fr

[2] IRIT-CNRS, 118 Route de Narbonne 31062 Toulouse Cedex, France. E-mail: fargier@irit.fr

$\succeq^{\text{Pareto}}$ is reflexive and transitive, but incomplete. A is strictly preferred to B in two cases: $\text{OM}(A^+) \geq \text{OM}(B^+)$ and $\text{OM}(A^-) < \text{OM}(B^-)$, or $\text{OM}(A^+) > \text{OM}(B^+)$ and $\text{OM}(A^-) \leq \text{OM}(B^-)$. A and B are indifferent when $\text{OM}(A^+) = \text{OM}(B^+)$ and $\text{OM}(A^-) = \text{OM}(B^-)$. In other cases, A is not comparable with B.

2.2 The implication rule $\succeq^{\text{DPoss}}$

This rule focuses on the most important arguments in the situation. A is at least as good as B iff, at level $\text{OM}(A \cup B)$, the presence of arguments for B is cancelled by the existence of arguments for A, and the existence of arguments against A is cancelled by the existence of arguments against B. Formally:

$$A \succeq^{\text{DPoss}} B \text{ iff :}$$

$$\begin{array}{rl} & \text{OM}(A \cup B) = \text{OM}(B^+) \Rightarrow \text{OM}(A \cup B) = \text{OM}(A^+) \\ \text{and} & \text{OM}(A \cup B) = \text{OM}(A^-) \Rightarrow \text{OM}(A \cup B) = \text{OM}(B^-) \end{array}$$

$\succeq^{\text{DPoss}}$ is reflexive, transitive and incomplete. E.g., a set with a positive and a negative argument in 1_L is incomparable to any other set.

2.3 The ordinal bipolar rule $\succeq^{\text{Poss}}$

This rule is simpler but less decisive. It considers any argument against A as an argument for B; any argument for A as an argument against B; and reciprocally. Then, the decision supported by the strongest argument(s) is preferred:

$$A \succeq^{\text{Poss}} B \text{ iff :}$$

$$\max(\text{OM}(A^+), \text{OM}(B^-)) \geq \max(\text{OM}(B^+), \text{OM}(A^-))$$

This rule is complete but quasi-transitive only. That is, $\succ^{\text{Poss}}$ is transitive but the indifference relation is not necessarily transitive. For example, when $\text{OM}(B^+) = \text{OM}(B^-)$, it is possible that indifference obtains between A and B, and between B and C as well, but not however between A and C.

Notice that $\succ^{\text{Poss}}$ is the rule identified in [17] as a principle for order of magnitude reasoning. It is less decisive than $\succeq^{\text{DPoss}}$ in the sense that $\succeq^{\text{DPoss}}$ is a refinement of $\succeq^{\text{Poss}}$:

$$A \succ^{\text{Poss}} B \Rightarrow A \succ^{\text{DPoss}} B.$$

2.4 Discriminating arguments: $\succeq^{\text{Discri}}$ and $\succeq^{\text{DDiscri}}$

$\succeq^{\text{Poss}}$ and (to a lower extent) $\succeq^{\text{DPoss}}$ and $\succeq^{\text{Pareto}}$ suffer from a severe drowning effect [2, 7, 9] that is often found in purely possibilistic frameworks. For example, when B is included in A, and even if all of the proper elements of A are positive, A is not necessarily strictly preferred to B.

This drowning problem is related to the fact that the three rules $\succeq^{\text{Poss}}$, $\succeq^{\text{DPoss}}$, and $\succeq^{\text{Pareto}}$ do not satisfy the condition of preferential independence, a very natural principle of decision making used when the criteria (or arguments) are non-interactive:

$$\forall A, B, C \text{ such that } (A \cup B) \cap C = \emptyset : A \succeq B \iff A \cup C \succeq B \cup C$$

$\succeq^{\text{Discri}}$ and $\succeq^{\text{DDiscri}}$ incorporate the principle of preferential independence, by cancelling elements that appear in both sets before applying $\succeq^{\text{Poss}}$ or $\succeq^{\text{DPoss}}$:[3]

$$\begin{array}{lllllll} A & \succeq^{\text{Discri}} & B & \iff & A \setminus B & \succeq^{\text{Poss}} & B \setminus A \\ A & \succeq^{\text{DDiscri}} & B & \iff & A \setminus B & \succeq^{\text{DPoss}} & B \setminus A \end{array}$$

[3] A similar variant of the Pareto rule might also be proposed ($A \succeq^{\text{DiscriPareto}} B \iff A \setminus B \succeq^{\text{Pareto}} A \setminus B$). We did not study any further the refinements of $\succeq^{\text{Pareto}}$, since it rapidly appeared that the rule was counter intuitive.

$\succeq^{\text{Discri}}$ is complete (as is $\succeq^{\text{Poss}}$) but not transitive, although its strict part $\succ^{\text{Discri}}$ is. $\succeq^{\text{DDiscri}}$ is partial (as is $\succeq^{\text{DPoss}}$) but not transitive, although $\succ^{\text{DDiscri}}$ is.

2.5 Cardinality rules $\succeq^{\text{Bilexi}}$ and $\succeq^{\text{Lexi}}$

These rules are based on a levelwise comparison by cardinality. The arguments in A and B are scanned top-down, until a level is reached such that there is a difference either in the number of arguments for A and B, or in the number of arguments against A and B. At this point, the set that presents the lower number of negative arguments and the greater number of positive ones is preferred. Formally: for any level $i \in L$, let

$$\begin{array}{lll} A_i & = & \{x \in A, \pi(x) = i\} \\ A_i^+ & = & A_i \cap X^+ \\ A_i^- & = & A_i \cap X^- \end{array} .$$

Let δ be the highest value of i s.t. $|A_i^+| \neq |B_i^+|$ or $|A_i^-| \neq |B_i^-|$. Then:

$$A \succeq^{\text{Bilexi}} B \iff |A_\delta^+| \geq |B_\delta^+| \text{ and } |A_\delta^-| \leq |B_\delta^-|$$

$\succeq^{\text{Bilexi}}$ is reflexive, transitive, but not complete. Indeed, if (at the decisive level) one of the set wins on the positive side, and the other on the negative side, the rule concludes to incomparability. A more decisive variant of $\succeq^{\text{BiLexi}}$ loses information about this conflict. The idea behind $\succeq^{\text{Lexi}}$ is to simplify each set before the comparison, accepting that one positive and one negative argument of A can cancel each other. In other terms, at each level j, A is assigned the score $|A_j^+| - |A_j^-|$. A top-down comparison of the scores is then performed:

$$A \succeq^{\text{Lexi}} B \iff \exists i \in L \text{ such that:}$$

$$\begin{array}{llll} \forall j > i, & |A_j^+| - |A_j^-| & = & |B_j^+| - |B_j^-| \\ \text{and} & |A_i^+| - |A_i^-| & > & |B_i^+| - |B_i^-| \end{array} .$$

3 Empirical Tests

3.1 Methods

How well does each rule predict the actual choices made by lay persons? To answer this question, we elaborated 33 situations of choice between two options, each option being represented as a list of positive and negative features of varying importance (see Table 1; the set of situations was designed to maximize the differentiability of the different rules, as well as to check some basic hypotheses about the qualitative nature of the decisions). Participants were 62 adults (31 men, 31 women, mean age = 24). The decisions involved 'Poldevian' stamps (a fictive nation). Stamp collection provided us with a clear-cut situation of qualitative comparison. Insofar as information about the monetary value of the stamps was unavailable, they were sorted in two broad groups: the rare, coveted stamps on the one hand and the common ones on the other. This was explicitly explained to participants:

> "Poldevian stamps come in two types, **rare** and **common**. Rare stamps are difficult to find, and they are treasured by collectors. Common stamps are much easier to find, and add much less value to a collection. Among the many Poldevian stamps, we will only be interested today in four rare and four common stamps. The rare stamps are called ARBON, BANTA, CASSA, and DIDOT. The common stamps are called WIV, XYL, YER, and ZAM."

Table 1. The 33 situations, with choices yielded by the 7 rules. Option $a^{++}(xy)^{-}$ has one very positive feature a and two mildly negative features x and y. $\emptyset$ is the null option. $\succ$, $\prec$, $=$, and $\sim$ resp. read 'prefer option 1', 'prefer option 2', 'indifferent', 'options are incomparable'.

	Option1	Option2	Pareto	DPoss	Poss	DDiscri	Discri	BiLexi	Lexi
1	$a^{++}(wxyz)^{-}$	$\emptyset$	$\sim$	$\succ$	$\succ$	$\succ$	$\succ$	$\succ$	$\succ$
2	$(wxyz)^{+}b^{--}$	$\emptyset$	$\sim$	$\prec$	$\prec$	$\prec$	$\prec$	$\prec$	$\prec$
3	$c^{++}d^{--}$	$\emptyset$	$\sim$	$\sim$	$=$	$\sim$	$=$	$\sim$	$=$
4	$a^{++}z^{+}b^{--}$	$\emptyset$	$\sim$	$\sim$	$=$	$\sim$	$=$	$\sim$	$\succ$
5	$a^{++}b^{--}z^{-}$	$\emptyset$	$\sim$	$\sim$	$=$	$\sim$	$=$	$\sim$	$\prec$
6	$b^{++}a^{--}$	$b^{++}(wxyz)^{-}$	$\prec$	$\prec$	$=$	$\prec$	$\prec$	$\prec$	$\prec$
7	$a^{++}c^{--}$	$d^{++}(wxyz)^{-}$	$\prec$	$\prec$	$=$	$\prec$	$=$	$\prec$	$\prec$
8	$a^{++}d^{--}$	$(wxyz)^{+}d^{--}$	$\succ$	$\succ$	$=$	$\succ$	$\succ$	$\succ$	$\succ$
9	$d^{++}c^{--}$	$(wxyz)^{+}a^{--}$	$\succ$	$\succ$	$=$	$\succ$	$=$	$\succ$	$\succ$
10	$d^{++}b^{--}$	w^{+}	$\sim$	$\sim$	$=$	$\sim$	$=$	$\sim$	$\prec$
11	w^{-}	$a^{++}c^{--}$	$\sim$	$\sim$	$=$	$\sim$	$=$	$\sim$	$\prec$
12	$c^{++}(wxyz)^{-}$	$(bc)^{++}a^{--}$	$\succ$	$\succ$	$=$	$\succ$	$=$	$\sim$	$\prec$
13	$d^{++}(wxyz)^{-}$	$(ab)^{++}c^{--}$	$\succ$	$\succ$	$=$	$\succ$	$=$	$\sim$	$\prec$
14	$b^{++}(ad)^{--}$	$(wxyz)^{+}d^{--}$	$\succ$	$\succ$	$=$	$\succ$	$=$	$\sim$	$\prec$
15	$a^{++}(cd)^{--}$	$(wxyz)^{+}b^{--}$	$\succ$	$\succ$	$=$	$\succ$	$=$	$\sim$	$\prec$
16	a^{++}	$(wxyz)^{+}$	$\succ$	$\succ$	$\succ$	$\succ$	$\succ$	$\succ$	$\succ$
17	b^{++}	$b^{++}z^{+}$	$=$	$=$	$=$	$\prec$	$\prec$	$\prec$	$\prec$
18	c^{++}	$d^{++}z^{+}$	$=$	$=$	$=$	$=$	$=$	$\prec$	$\prec$
19	$(bd)^{++}$	$(ab)^{++}w^{+}$	$=$	$=$	$=$	$=$	$=$	$\prec$	$\prec$
20	$(bc)^{++}$	$d^{++}(wxyz)^{+}$	$=$	$=$	$=$	$=$	$=$	$\succ$	$\succ$
21	a^{--}	$(wxyz)^{-}$	$\prec$	$\prec$	$\prec$	$\prec$	$\prec$	$\prec$	$\prec$
22	b^{--}	$b^{--}x^{-}$	$=$	$=$	$=$	$\succ$	$\succ$	$\succ$	$\succ$
23	c^{--}	$d^{--}w^{-}$	$=$	$=$	$=$	$=$	$=$	$\succ$	$\succ$
24	$(bd)^{--}$	$(ab)^{--}w^{-}$	$=$	$=$	$=$	$=$	$=$	$\succ$	$\succ$
25	$(bd)^{--}$	$a^{--}(wxyz)^{-}$	$=$	$=$	$=$	$=$	$=$	$\prec$	$\prec$
26	$(ab)^{++}(wxyz)^{-}$	a^{++}	$\prec$	$=$	$=$	$\succ$	$\succ$	$\succ$	$\succ$
27	$(bd)^{++}(wxyz)^{-}$	c^{++}	$\prec$	$=$	$=$	$=$	$=$	$\succ$	$\succ$
28	a^{--}	$(wxyz)^{+}(ac)^{--}$	$\prec$	$=$	$=$	$\succ$	$\succ$	$\succ$	$\succ$
29	c^{--}	$(wxyz)^{+}(bd)^{--}$	$\prec$	$=$	$=$	$=$	$=$	$\succ$	$\succ$
30	$d^{++}w^{-}$	d^{++}	$\prec$	$=$	$=$	$\prec$	$\prec$	$\prec$	$\prec$
31	$b^{++}w^{-}$	a^{++}	$\prec$	$=$	$=$	$=$	$=$	$\prec$	$\prec$
32	$c^{--}w^{+}$	c^{--}	$\succ$	$=$	$=$	$\succ$	$\succ$	$\succ$	$\succ$
33	$d^{--}w^{+}$	a^{--}	$\succ$	$=$	$=$	$=$	$=$	$\succ$	$\succ$

It is worth noting that using an abstract setting like this one, with little appeal to participant's world knowledge and little import to their personal life, is a routine procedure in experimental psychology. It allows to escape the idiosyncracies of a given application domain (e.g., choice of an apartment, or a car), and to study a mental process under rigorously controlled conditions.

Participants' task consisted in a series of 33 choices between two options (in choices 1 to 5, the second option was the status quo). E.g., in Situation 15, participants were asked which of Club 1 or 2 they would join: Club 1 offered ARBON as a welcome gift, but required DIDOT and CASSA as a membership fee. Club 2, on the other hand, offered WIV, XYL, YER, and ZAM as a welcome gift, but required BANTA as a membership fee. Participants could choose one of four responses: (a) choose Club 1, (b) choose Club 2, (c) indifferent, one or the other, would agree to choose randomly, and (d) unable to make a decision, would not agree to choose randomly. While the third response suggests indifference between the two options, the fourth response indicates incomparability.

3.2 Results

3.2.1 Preliminaries

Choices made by participants in all 33 situations are reported in Table 2. Before we turn to the results proper, three general remarks are in order. First, a great majority of the sample (80 to 90%) used a pure qualitative ranking of the arguments, as we expected. It remains that some participants (10 to 20%) appear to go beyond a strictly qualitative evaluation of the arguments: Situations 16 and 21 (and possibly Situation 1) make it clear that these participants value four common stamps more than one rare stamp.

Second, some participants (28 to 34%) seem to re-scale negative arguments, giving them more weight than positive arguments: getting a rare stamp is less important than loosing a rare one. E.g., these 'qualitative pessimists' prefer the null option in Situation 3, or option 1 in Situation 11. Finally, some misunderstandings apparently could not be avoided, especially in the third tier of the task. E.g., a small number of participants (7 to 10%) manifest a strict preference for Option 2 in Situations 22, 23, 24, or for Option 1 in Situation 25. This rate nevertheless stays low (below 10% in any case).

3.2.2 Overall descriptive validity

As a first measure of overall descriptive validity, we computed the 'fit' of each procedure to the choices made by participants. The fit of a procedure is the *average percentage of answers it correctly predicts*, across participants. A good fit might not be a guarantee of cognitive plausibility, but a low fit is certainly an indicator of poor descriptive validity. A second measure of overall validity is what we call the 'restricted fit' of the procedure, i.e., the average percentage of answers it correctly predicts across participants, only taking into account the situations *for which it predicts a strict preference*. Table 3 reports the general and restricted fit of each rule. (We also tested the fit of the procedure suggested by Benferhat and Kaci [12, 3], which we mentioned in the Introduction – this fit was only 26%.)

$\succeq^{\text{DPoss}}$, $\succeq^{\text{Pareto}}$ and especially $\succeq^{\text{Poss}}$ fare badly in this first evaluation. The restricted fit of $\succeq^{\text{Poss}}$ may seem high, but it concerns 4 situations only (on the positive side, this shows that the principles of $\succ^{\text{Poss}}$ are considered as an obvious, minimal norm by the decision makers). On the contrary, $\succeq^{\text{DPoss}}$ and $\succeq^{\text{Pareto}}$ show some serious weakness with respect to restricted fit. Only $\succeq^{\text{Bilexi}}$ and $\succeq^{\text{Lexi}}$ achieve promising general

Table 2. Choices made by participants (in % of answers) in the 33 experimental Situations.

	Option1	Option2	$\succ$	$\prec$	$=$	$\sim$
1	$a^{++}(wxyz)^{-}$	$\emptyset$	79	21	–	–
2	$(wxyz)^{+}b^{--}$	$\emptyset$	–	86	7	7
3	$c^{++}d^{--}$	$\emptyset$	3	34	35	28
4	$a^{++}z^{+}b^{--}$	$\emptyset$	73	10	3	14
5	$a^{++}b^{--}z^{-}$	$\emptyset$	3	83	–	14
6	$b^{++}a^{--}$	$b^{++}(wxyz)^{-}$	7	83	3	7
7	$a^{++}c^{--}$	$d^{++}(wxyz)^{-}$	10	80	–	10
8	$a^{++}d^{--}$	$(wxyz)^{+}d^{--}$	83	3	3	11
9	$d^{++}c^{--}$	$(wxyz)^{+}a^{--}$	76	–	3	21
10	$d^{++}b^{--}$	w^{+}	14	76	7	3
11	w^{-}	$a^{++}c^{--}$	45	38	3	14
12	$c^{++}(wxyz)^{-}$	$(bc)^{++}a^{--}$	14	86	–	–
13	$d^{++}(wxyz)^{-}$	$(ab)^{++}c^{--}$	28	69	3	–
14	$b^{++}(ad)^{--}$	$(wxyz)^{+}d^{--}$	10	45	3	42
15	$a^{++}(cd)^{--}$	$(wxyz)^{+}b^{--}$	7	45	3	45
16	a^{++}	$(wxyz)^{+}$	90	10	–	–
17	b^{++}	$b^{++}z^{+}$	–	100	–	–
18	c^{++}	$d^{++}z^{+}$	–	100	–	–
19	$(bd)^{++}$	$(ab)^{++}w^{+}$	–	97	3	–
20	$(bc)^{++}$	$d^{++}(wxyz)^{+}$	83	17	–	–
21	a^{--}	$(wxyz)^{-}$	17	80	–	3
22	b^{--}	$b^{--}x^{-}$	73	10	7	10
23	c^{--}	$d^{--}w^{-}$	83	7	3	7
24	$(bd)^{--}$	$(ab)^{--}w^{-}$	73	10	3	14
25	$(bd)^{--}$	$a^{--}(wxyz)^{-}$	7	76	–	17
26	$(ab)^{++}(wxyz)^{-}$	a^{++}	72	28	–	–
27	$(bd)^{++}(wxyz)^{-}$	c^{++}	72	28	–	–
28	a^{--}	$(wxyz)^{+}(ac)^{--}$	90	3	–	7
29	c^{--}	$(wxyz)^{+}(bd)^{--}$	86	7	–	7
30	$d^{++}w^{-}$	d^{++}	–	97	3	–
31	$b^{++}w^{-}$	a^{++}	–	97	3	–
32	$c^{--}w^{+}$	c^{--}	90	–	3	7
33	$d^{--}w^{+}$	a^{--}	86	–	4	10

Table 3. Average % of answers predicted by each rule, overall or when restricted to situations when it predicts a strict preference

	Overall fit	Restricted fit
Poss	13	83
Dposs	23	57
Pareto	31	53
Discri	32	83
DDiscri	39	70
BiLexi	64	82
Lexi	77	79

fits of 64% and 77%. Although very decisive ($>^{\text{Lexi}}$ is decisive in 32 of the 33 situations), these two rules also have a good restricted fit.

$\succeq^{\text{Lexi}}$ has the best fit with the choices made by any single one of the participants – it always provides the best description for all individual patterns of decisions. The second best fit is always $\succeq^{\text{Bilexi}}$'s. The difference between $\succeq^{\text{Lexi}}$'s and $\succeq^{\text{Bilexi}}$'s fits is statistically reliable, as assessed by Student's $t(61) = 11.1$, $p < .001$: The probability of observing such a difference in our sample, were the fits actually similar in the general population, is less than one in a thousand. The same result holds when comparing the fit of $\succeq^{\text{Bilexi}}$ to the fits of all other procedures. >From these initial considerations, only $\succeq^{\text{Lexi}}$ and $\succeq^{\text{Bilexi}}$ emerge as serious candidates for predicting participants' choices. We will now consider in more detail the case that can be made from our data in support of these two cardinality-based procedures.

3.2.3 Escaping the drowning effect by cancellation principles

Twelve situations (17—18, 22—23, 26—33) were designed so that they could elicit a drowning effect. Depending on the situation, 66 to 95% participants escape drowning. This strongly suggests that human decision-makers know their way out the classic drawback of purely possibilistic procedures like $\succeq^{\text{Poss}}$, $\succeq^{\text{DPoss}}$ or $\succeq^{\text{Pareto}}$.

$\succeq^{\text{DDiscri}}$ and $\succeq^{\text{Discri}}$ (on the one hand), and $\succeq^{\text{Bilexi}}$ and $\succeq^{\text{Lexi}}$ (on the other hand), represent two different ways to overcome the drowning effect. $\succeq^{\text{DDiscri}}$ and $\succeq^{\text{Discri}}$ do it by cancelling the arguments that appear in both options, thus realizing a refinement by inclusion of $\succeq^{\text{DPoss}}$ and $\succeq^{\text{Poss}}$. The two rules $\succeq^{\text{Bilexi}}$ and $\succeq^{\text{Lexi}}$ lexi rules also allow the cancellation of positive (resp. negative) arguments of the same level, an additional refinement by cardinality (see [8] for more details).

The proportion of participants who refine by inclusion and not by cardinality is very low, less than 10 % in any case. To a couple of exceptions only, anytime a participant drowns on a situation that could be refined by cardinality, this participant had already failed to refine by inclusion on the corresponding version of the problem. Thus, in a very large majority of cases, refinement appears to operate through cardinality rather than simply inclusion.

Finally, evidence of *internal cancellation* (within a given option, a positive and a negative feature of the same level cancel each other) would count as an argument for favoring $\succeq^{\text{Lexi}}$ over $\succeq^{\text{Bilexi}}$. In all relevant situations (3—5 and 10—15), the modal answer is *always* the one predicted by $\succeq^{\text{Lexi}}$, and never the one predicted by $\succeq^{\text{Bilexi}}$.

4 Conclusion

We proposed seven procedures for choosing among options represented as bipolar sets of ordinal-ranked arguments. We then asked the question of their descriptive value: How well does each of them predict the choices human decision-makers would make? For many of these procedures, the answer is: not very well, with the noticeable exception of the two cardinality rules that seem to predict human choices fairly well, and $\succeq^{\text{Lexi}}$ in particular has the highest empirical validity among our procedures.

An interesting parallel can be made in that regard between $\succeq^{\text{Lexi}}$ and the *Take the Best* decision heuristic that has been intensively studied by psychologists [4, 10, 13, 14]. When two options evaluated with respect to a series of strictly ordered criteria (in the sense that the criteria are supposed to be of very different order of magnitude so that they can be ranked lexicographically), Take the Best can be applied and amounts to choosing the option that is prefered by the most important of the criteria that makes a difference (going down one level in case of a tie, and so forth). Applied on such a situation of lexicographically ranked criteria, the discri and lexi bipolar rules are formally equivalent to Take the Best. But they are able to account for other decision situations, e.g., several criteria can share the same degree of importance.

In this sense, $\succeq^{\text{Lexi}}$ is a natural extension of the widely popular (though controversial) rule advocated by psychologists.[4] We believe that decision rules for artificial agents should be, whenever possible, as descriptively valid as they are formally sound – from the results we have reported, we see the present paper as a step towards showing that $\succeq^{\text{Lexi}}$ is blessed indeed with these two virtues.

REFERENCES

[1] L. Amgoud, J. F. Bonnefon, and H. Prade, 'An argumentation-based approach for multiple criteria decision', in *ECSQARU'05 - LNAI 3571*, ed., L. Godo, Berlin, (2005). Springer Verlag.

[2] S. Benferhat, C. Cayrol, D. Dubois, J. Lang, and H. Prade, 'Inconsistency management and prioritized syntax-based entailment', in *IJCAI*, pp. 640–647, (1993).

[3] S. Benferhat, D. Dubois, S. Kaci, and H. Prade, 'Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions', *International Journal on Information Fusion*, (to appear).

[4] A. Bröder, 'Assessing the empirical validity of the "Take-The- Best" heuristic as a model of human probabilistic inference', *Journal of Experimental Psychology: Learning. Memory, and Cognition*, **26**, 1332–1346, (2000).

[5] J. T. Cacioppo and G. G. Berntson, 'Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates', *Psychological Bulletin*, **115**, 401–423, (1994).

[6] J. T. Cacioppo, W. L. Gardner, and G. G. Berntson, 'The affect system has parallel and integrative processing components: Form follows function', *Journal of Personality and Social Psychology*, **76**, 839–855, (1999).

[7] H. Fargier D. Dubois and H. Prade, 'Refinements of the maximin approach to decision-making in fuzzy environment', *Fuzzy Sets and Systems*, **81**, 103–122, (1996).

[8] D. Dubois and H. Fargier, 'On the qualitative comparison of sets of positive and negative affects', in *ECSQARU'05 - LNAI 3571*, ed., L. Godo, pp. 305–316, Berlin, (2005). Springer-Verlag.

[9] H. Fargier and R. Sabbadin, 'Qualitative decision under uncertainty: Back to expected utility', in *IJCAI*, pp. 303–308, (2003).

[10] G. Gigerenzer, , and D. G. Goldstein, 'Reasoning the fast and frugal way: Models of bounded rationality', *Psychological Review*, **103**, 650–669, (1996).

[11] G. Gigerenzer, P. M. Todd, and the ABC group, *Simple heuristics that make us smart*, Oxford University Press, 1999.

[12] S. Kaci and H. Prade, 'Bipolar goals. a possibilisitc logic characterization odf', in *Proceedings of the ECAI'04 Workshop on local computation for logics and uncertainty*, pp. 13–18, (2004).

[13] B. R. Newell, , and D. R. Shanks, 'Take-the-best or look at the rest? Factors influencing 'one-reason' decision making', *Journal of Experimental Psychology: Learning. Memory, and Cognition*, **29**, 53–65, (2003).

[14] B. R. Newell, N. J. Weston, and D. R. Shanks, 'Empirical tests of a fast and frugal heuristic: Not everyone "takes-the-best"', *Organizational Behavior and Human Decision Processes*, **91**, 82–96, (2003).

[15] P. Slovic, M. Finucane, E. Peters, and D. G. MacGregor, 'Rational actors or rational fools? implications of the affect heuristic for behavioral economics', *The Journal of Socio-Economics*, **31**, 329–342, (2002).

[16] A. Wald, *Statistical Decision Functions*, Wiley, 1950.

[17] N. Wilson, 'An order of magnitude calculus', in *Proceedings UAI-95*, pp. 548–555, San Francisco, CA, (1995).

[4] While we will not push this comparison any further in the present paper, we do plan to give it due consideration in future work. In like vein, we plan to consider how other strategies identified in the psychological literature can relate to the rules we have already explored, or can be given a formal counterpart within our approach.