Qualitative Heuristics for Balancing the Pros and Cons

Jean-François Bonnefon (bonnefon@univ-tlse2.fr)
LTC-CNRS, Toulouse

Didier Dubois and Hélène Fargier
IRIT-CNRS, Toulouse

Sylvie Leblois
LTC-CNRS, Toulouse

Abstract. Balancing the pros and cons of two options is undoubtedly a very appealing decision procedure, but one that has received scarce scientific attention so far, either formally or empirically. We describe a formal framework for pros and cons decisions, where the arguments under consideration can be of varying importance, but whose importance cannot be precisely quantified. We then define 8 heuristics for balancing these pros and cons, and compare the predictions of these to the choices made by 62 human participants on a selection of 33 situations. The Levelwise Tallying heuristic clearly emerges as a winner in this competition. Further refinements of this heuristic are considered in the discussion, as well as its relation to Take the Best and Cumulative Prospect Theory.

Keywords: decision heuristics; bipolar information; qualitative information; behavioral data; Take the Best; Cumulative Prospect Theory

1. Introduction

Balancing the pros and the cons is certainly among the most intuitive approaches one might take to decision making. It was already at the core of Benjamin Franklin's "moral algebra" (explained in his famous 1772 letter to Joseph Priestley), and it has certainly not fallen from grace since then, witness the 93, 000, 000 web pages featuring both the words "pro" and "con" as of November 2006. One likely feature of this kind of decision is that the decision maker will be unable to precisely quantify how important a given pro or con is, although she may be able to give a qualitative assessment of this importance.

We can assume that many decision makers attempt to reach a decision by balancing pros and cons, after roughly sorting them out in different levels of importance. But however appealing this qualitative balancing act might sound, it has inspired only few mathematical explorations to date, and even fewer psychological investigations. In section **??**, we introduce a formal framework for pros and cons decisions. In section **??**, we describe eight heuristics for balancing pros and cons. The predictions of these heuristics are then compared to choices made by human participants in 33 situations, chosen to emphasize the differences between the heuristics.

2. Pros & Cons Decisions

Tonight, Emma is going to the cinema, and considers watching one of two movies. She has listed the pros and cons of each choice. Movie 1 is one by her favorite director (a strong pro); it will be dubbed, which she hates, and the movie has attracted terrible critics (two strong cons). Movie 2 is only given in a remote theater, and she considers it a strong con that she would need a taxi to get there. On the other hand, movie 2 is a comedy (a genre she likes), it features an actress she likes, and it is inspired by a book she enjoyed reading. These are three pros, but Emma does not see them as very decisive: they do matter, but not as much as the other arguments she listed.

Note that Emma can only give a rough evaluation of how strong a pro or a con is. She can only say that (a) her liking the director, her hating dubbed movies, the terrible critics, and movie 2 being given in a remote theater are four arguments of comparable importance; and that (b) movie 2's genre, leading actress, and source of inspiration are three arguments of comparable importance, but not as important as the previous ones.

Before we try to predict what Emma's decision might be, let us formalize her problem. Each option (movie) is assessed by a finite subset of arguments taken from X, the set of all possible arguments. Comparing two options then amounts to comparing two subsets U, V of 2^X. X can be divided in three disjoint subsets: X^+ the set of pros, X^- the set of cons, and X^0 the set of irrelevant arguments (which do not count as a pro or a con). Any $U \subseteq X$ can likewise be partitioned: let $U^+ = U \cap X^+$, $U^- = U \cap X^-$, $U^0 = U \cap X^0$ be respectively the pros, the cons, and the irrelevant arguments relatively to U.

As in our movie example, all arguments are not equally important, although it is generally impossible to precisely quantify the importance of a given argument. Thus, in a purely qualitative, ordinal approach, the importance of arguments is described on a totally ordered scale of magnitude $L = [0_L, 1_L]$, by a function $\pi : X \mapsto L = [0_L, 1_L]$. $\pi(x) = 0_L$ 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. Intermediate values are attached to arguments of intermediate importance. For any level $\alpha \in L$, let $U_\alpha = \{x \in U, \pi(x) = \alpha_L\}$, $U^+_\alpha = U_\alpha \cap X^+$, and $U^-_\alpha = U_\alpha \cap X^-$.

Finally, it will be useful to define the order of magnitude $M(U)$ of a set U as the highest of the order of magnitude of its elements:

$$\forall U \subseteq X, M(U) = \max_{x \in U} \pi(x).$$

Note that $M(U)$ is a possibility measure on X—for a recent review on qualitative possibility theory, see **?** (**?**). We can now reformulate the Emma case as the comparison between two options U (movie 1) and V (movie 2).[1] Option U has one argument in U^+_α and two arguments in U^-_α; and option V has one argument in V^-_α and three arguments in V^-_β, where $\alpha > \beta$. As we will see in the next section, the different heuristics that can be defined for balancing pros and cons have quite diverging views on the Emma case. Some will prefer movie 1, some will prefer movie 2, some will regard the two movies as equally attractive, and some will find it impossible to compare the merits of the two movies.

3. Pros & Cons Heuristics

Qualitative heuristics for balancing the pros and the cons (or, technically, ordinal ranking procedures from bipolar information) have received scarce attention so far. Most work on such procedures has come from the field of Artificial Intelligence, following the renewed interest in argumentative models of choice and inference (**?**; **?**).

The heuristics we describe in this section have been formally examined and axiomatized (**?**; **?**). Since the present article takes an empirical rather than analytical approach to pros and cons heuristics, we will not restate all the formal properties of the heuristics nor give their axiomatization. We will nevertheless comment on important properties such as completeness or transitivity.

3.1. Focus heuristics F1, F2, and F3

The heuristics in the "Focus" family concentrate on the most important arguments available for the decision, and disregard arguments of lesser importance.

[1] In the rest of this article, we will denote an option by the subset of arguments which is used to assess this option. E.g., movie 1 is assessed by the subset U of arguments, and will thus be denoted "Option U."

3.1.1. *The straw and the beam* ($\succeq^{\mathrm{F1}}$)
With this heuristic, U is at least as good as V if and only if, at level $M(U \cup V)$ (i.e., the highest level of importance in the current comparison), the presence of arguments for V is cancelled by the existence of arguments for U, and the existence of arguments against U is cancelled by the existence of arguments against V. Formally, $U \succeq^{\mathrm{F1}} V$ if and only if:

$$\begin{aligned} & M(U \cup V) = M(V^+) \;\Rightarrow\; M(U \cup V) = M(U^+) \\ \text{and} \quad & M(U \cup V) = M(U^-) \;\Rightarrow\; M(U \cup V) = M(V^-) \end{aligned}$$

What will Emma do? It turns out that $M(U \cup V) = M(U^-)$; that is, the strongest con is attached to U. However, it is also true that $M(U \cup V) = M(V^-)$. The second condition is thus satisfied: the existence of a strong argument against V offsets the existence of a strong argument against U. The first condition is satisfied because $M(U \cup V)$ is not $M(V^+)$. Consequently, it holds that $U \succeq^{\mathrm{F1}} V$. Now, it does not hold that $V \succeq^{\mathrm{F1}} U$, because while $M(U \cup V) = M(U^+)$, it is not the case that $M(U \cup V) = M(V^+)$. There is no strong argument for V to offset the existence of a strong argument for U. Emma will go and see movie 1.

The relation $\succeq^{\mathrm{F1}}$ is transitive but incomplete. As soon as an option has both a pro and a con at the highest importance level, it becomes incomparable to any other option whose description does not feature pros or cons at this highest importance level.

3.1.2. *My enemy's enemies* ($\succeq^{\mathrm{F2}}$)
This heuristic treats all arguments against V as arguments for U, and all arguments for V as arguments against U (and reciprocally). It then selects the option that is supported by the arguments at the highest level. Formally, $U \succeq^{\mathrm{F2}} V$ if and only if:

$$\max(M(U^+), M(V^-)) \geq \max(M(U^+), M(V^-)).$$

In the Emma case, $\max(M(U^+), M(V^-)) = \max(M(U^+), M(V^-)) = \alpha$. Emma is indifferent, she can toss a coin to decide on a movie. The relation $\succeq^{\mathrm{F2}}$ is simpler than the relation $\succeq^{\mathrm{F1}}$, and has the advantage of being complete, but it is only quasi-transitive: $\succ^{\mathrm{F2}}$ itself is transitive, but the corresponding indifference relation is not (e.g., when $M(V^+) = M(V^-)$, it is possible to have both $U \sim^{\mathrm{F2}} V$ and $V \sim^{\mathrm{F2}} W$, while $U \sim^{\mathrm{F2}} W$ may not hold).

Furthermore, while $\succeq^{\mathrm{F2}}$ is complete, it is not as decisive as $\succeq^{\mathrm{F1}}$: it yields indifference more often than $\succeq^{\mathrm{F1}}$ does, as illustrated by the Emma case. Indeed, $\succeq^{\mathrm{F1}}$ is a refinement of $\succeq^{\mathrm{F2}}$: $U \succ^{\mathrm{F2}} V \Rightarrow U \succ^{\mathrm{F1}} V$.

3.1.3. *Pareto dominance ($\succeq^{\mathrm{F3}}$)*
This heuristic looks for the option that wins on both the positive and negative sides. This rule compares the two sets of arguments as a problem of bi-criteria decision. The first criterion compares negative arguments according to Wald's rule (**?**): U is better than V on the negative side if and only if $M(U^-) \leq M(V^-)$. The second criterion compares positive arguments according to the optimistic counterpart of Wald's rule. Formally, $U \succeq^{\mathrm{F3}} V$ if and only if:

$$\begin{aligned} & M(U^+) \geq M(V^+) \\ \text{and} \quad & M(U^-) \leq M(V^-) \end{aligned}$$

To Emma, there is a strong argument for U, but only weak arguments for V: $M(U^+) = \alpha > M(V^+) = \beta$. In parallel, there are strong arguments both against U and against V: $M(U^-) = \alpha = M(V^-)$. Emma will go and see movie 1.

The relation $\succeq^{\mathrm{F3}}$ is transitive but not complete. For example, as soon as an option has both pros and cons, whatever their importance, it becomes incomparable to the null option (no pro, no con).

3.2. Inclusion heuristics I1 and I2

While the heuristics we have considered so far have some intuitive appeal, they all suffer from a notable shortcoming— that is, they do not satisfy the principle of preferential independence. This principle states that if U is preferred to V, then this preference should not change when the descriptions of U and V are enriched by the exact same set of arguments. Formally: $\forall U, V, W$ such that $(U \cup V) \cap W = \emptyset$, $U \succeq V \iff U \cup W \succeq V \cup W$.

Consider for example the case of Emma's cousin, Francine. Francine must decide whether she will go and see movie 3, about which she knows nothing, or movie 4, which features an actress she likes (a weak pro). All three Focus heuristics would (reasonably) predict that Francine will go and see movie 4. But let us now add the information that *both* movies are by Francine's favorite director (a strong pro in each case). Now, all three Focus heuristics predict that Francine will be indifferent between the two movies—a disputable prediction, and a violation of the principle of preferential independence.

The heuristics in the "Inclusion" family are variants of the Focus heuristics, which satisfy the principle of preferential independence. They do so by first cancelling the arguments that appear in the descriptions of both options, before applying one of the Focus heuristics. Formally:

$$\begin{aligned} U \succeq^{\mathrm{I1}} V &\iff U \setminus V \succeq^{\mathrm{F1}} V \setminus U \\ U \succeq^{\mathrm{I2}} V &\iff U \setminus V \succeq^{\mathrm{F2}} V \setminus U \end{aligned}$$

Obviously, each of the Inclusion heuristics refines its parent Focus heuristics—i.e., it follows the strict preference of the parent heuristic (if any) but may have a strict preference where the parent heuristic does not:

$$U \succ^{\mathrm{F1}} V \;\Rightarrow\; U \succ^{\mathrm{I1}} V$$
$$U \succ^{\mathrm{F2}} V \;\Rightarrow\; U \succ^{\mathrm{I2}} V$$

Just as its parent $\succeq^{\mathrm{F1}}$, the relation $\succeq^{\mathrm{I1}}$ is complete but its indifference part is not transitive. Likewise, just as its parent $\succeq^{\mathrm{F2}}$, the relation $\succeq^{\mathrm{I2}}$ is transitive but partial.

In the Emma case, $\succeq^{\mathrm{I1}}$ and $\succeq^{\mathrm{I2}}$ make the same predictions as $\succeq^{\mathrm{F1}}$ and $\succeq^{\mathrm{F2}}$, respectively, because there is nothing to simplify there—there is no argument that simultaneously appears in the descriptions of movie 1 and movie 2. In the Francine case, though, the Inclusion heuristics make a different prediction than the Focus heuristics. Remember that Francine is hesitating between two movies by her favorite director (strong pro in each case), one of which features an actress she likes (weak pro). According to the Focus heuristics, she should be indifferent between the two movies. What about the Inclusion heuristics?

Let us apply $\succeq^{\mathrm{I1}}$ to the Francine situation. First, the two options must be simplified by removing the arguments they have in common—here, the director. Now $\succeq^{\mathrm{F1}}$ must be applied to the simplified options. Here the choice is now between a movie to which no pros or cons are attached, and a movie with a weak pro (the actress). According to $\succeq^{\mathrm{F1}}$, Francine will prefer the latter movie. (The same applies to $\succeq^{\mathrm{F2}}$, and therefore to $\succeq^{\mathrm{I2}}$.)

3.3. Cardinality heuristics C1, C2 and C3

The Inclusion heuristics only cancel arguments that appear in the exact same form in the description of the two options. Further simplification methods are possible, of which we will consider three.

3.3.1. *Tallying* ($\succeq^{\mathrm{C1}}$)

This heuristic disregards the relative importance of arguments, and simply computes a score for each option by adding up the number of its pro then subtracting the number of its cons. The option with the best net score wins. Formally, $U \succeq^{\mathrm{C1}} V$ if and only if $|U^+| - |U^-| \geq |V^+| - |V^-|$. The relation $\succeq^{\mathrm{C1}}$ is complete and transitive.

In the Emma case, option U (movie 1) has one pro and two cons, while option V (movie 2) has three pros and one con. Thus $|U^+| - |U^-| = -1 < |V^+| - |V^-| = +2$. Emma will go and see movie 2.

Note that the Tallying heuristic is the only one in this article that does not take into account the relative importance of arguments. Our main reason for considering Tallying in the present article is its frequent appearance in the literature on decision heuristics (**?**; **?**; **?**; **?**).

3.3.2. *Bivariate levelwise tallying* ($\succeq^{C2}$)

Unlike the previous one, this heuristic takes into account that some arguments are more important than others. This heuristic first considers arguments at the highest level of importance, and checks whether some option achieves bipolar dominance by cardinality. That is, the arguments in U and V are scanned top-down, until a level is reached such that there is a difference either in the number of arguments for U and V, or in the number of arguments against U and V. At this point, the set that presents the lower number of cons and the greater number of pros is preferred. Formally, let δ be the highest value of α such that $|U^+_\alpha| \neq |V^+_\alpha|$ or $|U^-_\alpha| \neq |V^-_\alpha|$. Then $U \succeq^{C1} V$ if and only if:

$$\begin{array}{rrcl} & |U^+_\delta| & \geq & |V^+_\delta| \\ \text{and} & |U^-_\delta| & \leq & |V^-_\delta| \end{array}$$

What will Emma do? She stops at the highest importance level α, because it is already true that $|U^-_\alpha| > |V^-_\alpha|$: there are two strong arguments against U, but only one against V. However, at this same level α, there is one argument for U and no argument for V, and thus $|U^+_\alpha| > |V^+_\alpha|$. Therefore, Emma is stuck and finds herself unable to decide between the two movies. Note that this is not the same as being indifferent: In the present situation, she would not agree to make the decision by tossing a coin.

As shown by the Emma example, the relation $\succeq^{C2}$ is not complete (it is transitive, though). For example, it may happen that, at the decisive level, one option wins on the positive side but loses on the negative side. A way around this obstacle is to allow within-option simplification of the arguments before the comparison takes place. This is what is done in our last heuristic.

3.3.3. *(Univariate) Levelwise tallying* ($\succeq^{C3}$)

The *Levelwise Tallying* heuristic, just as its bivariate cousin above, first considers arguments at the highest level of importance. For each option, it computes a score by adding up the number of its pros (at this level) and then subtracting the number of its cons (still at this level). The option with the highest net score wins; if there is a tie, the procedure is repeated at the next level of argument importance. Formally, $U \succeq^{C3} V$ if and only if $|U^+_\beta| - |U^-_\beta| > |V^+_\beta| - |V^-_\beta|$, and for all $\alpha > \beta$, $|U^+_\alpha| - |U^-_\alpha| = |V^+_\alpha| - |V^-_\alpha|$. The relation $\succeq^{C3}$ is complete and transitive.

Emma begins with arguments at the highest level of importance α. It turns out that $|U_\alpha^+| - |U_\alpha^-| = -1 = |V_\alpha^+| - |V_\alpha^-|$. Emma goes down one level of importance and considers arguments of importance β. She finds out that $|U_\beta^+| - |U_\beta^-| = 0 < |V_\beta^+| - |V_\beta^-| = +3$. Thus, it holds that $V \succeq^{\mathrm{C3}} U$. Emma will go and see movie 2.

Just as $\succeq^{\mathrm{C2}}$ and $\succeq^{\mathrm{I2}}$, the levelwise tallying heuristic obeys the principle of preferential independence. It is also a refinement of $\succeq^{\mathrm{F2}}$: it follows the strict preference of $\succeq^{\mathrm{F2}}$ if there is one, but it is more decisive (yields less ties) than $\succeq^{\mathrm{F2}}$. In fact, the heuristics that derive from $\succeq^{\mathrm{F2}}$ can be ranked from the least to the most decisive, as follows:

$$A \succ^{\mathrm{F2}} B \Rightarrow A \succ^{\mathrm{I2}} B \Rightarrow A \succ^{\mathrm{C2}} B \Rightarrow A \succ^{\mathrm{C3}} B$$

4. Method

To assess the descriptive validity of these eight heuristics, we elaborated 33 situations of choice between two options, and compared in each case the predictions of the heuristics to the choices made by a sample of 62 adult volunteers (31 men, 31 women, mean age = 24.0, *SD* = 8.9).

The decisions all involved the trading of "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, unremarkable stamps 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."

In the 33 situations, the two options were described as a list of pros and cons, couched in terms of stamps. Receiving a rare stamp was a strong pro, receiving a common stamp was a weak pro; giving away a rare stamp was a strong con, and giving away a common stamp was a weak con. The Poldevian stamp equivalent to the Emma case is found in situation 15: Option U is to receive ARBON but give away DIDOT and CASSA (all rare stamps); and option V is to receive WIV, XYL, YER, and ZAM (all common stamps), but give away BANTA (a rare stamp).

Table I. Predictions of the 8 heuristics in the 33 choice situations. An option described as $a^{++}(xy)^{-}$ has one very positive feature a and two mildly negative features x and y. $\emptyset$ is the null option. U, V, $=$, and $\perp$ resp. read 'prefer U', 'prefer V', 'indifferent', 'options are incomparable'.

	U	V	F1	F2	F3	I1	I2	C1	C2	C3
1	$a^{++}(wxyz)^{-}$	$\emptyset$	U	U	$\perp$	U	U	V	U	U
2	$(wxyz)^{+}b^{--}$	$\emptyset$	V	V	$\perp$	V	V	U	V	V
3	$c^{++}d^{--}$	$\emptyset$	$\perp$	$=$	$\perp$	$\perp$	$=$	$=$	$\perp$	$=$
4	$a^{++}z^{+}b^{--}$	$\emptyset$	$\perp$	$=$	$\perp$	$\perp$	$=$	U	$\perp$	U
5	$a^{++}b^{--}z^{-}$	$\emptyset$	$\perp$	$=$	$\perp$	$\perp$	$=$	V	$\perp$	V
6	$b^{++}a^{--}$	$b^{++}(wxyz)^{-}$	V	$=$	V	V	V	U	V	V
7	$a^{++}c^{--}$	$d^{++}(wxyz)^{-}$	V	$=$	V	V	$=$	U	V	V
8	$a^{++}d^{--}$	$(wxyz)^{+}d^{--}$	U	$=$	U	U	U	V	U	U
9	$d^{++}c^{--}$	$(wxyz)^{+}a^{--}$	U	$=$	U	U	$=$	V	U	U
10	$d^{++}b^{--}$	w^{+}	$\perp$	$=$	$\perp$	$\perp$	$=$	V	$\perp$	V
11	w^{-}	$a^{++}c^{--}$	$\perp$	$=$	$\perp$	$\perp$	$=$	V	$\perp$	V
12	$c^{++}(wxyz)^{-}$	$(bc)^{++}a^{--}$	U	$=$	U	U	$=$	V	$\perp$	V
13	$d^{++}(wxyz)^{-}$	$(ab)^{++}c^{--}$	U	$=$	U	U	$=$	V	$\perp$	V
14	$b^{++}(ad)^{--}$	$(wxyz)^{+}d^{--}$	U	$=$	U	U	$=$	V	$\perp$	V
15	$a^{++}(cd)^{--}$	$(wxyz)^{+}b^{--}$	U	$=$	U	U	$=$	V	$\perp$	V
16	a^{++}	$(wxyz)^{+}$	U	U	U	U	U	V	U	U
17	b^{++}	$b^{++}z^{+}$	$=$	$=$	$=$	V	V	V	V	V
18	c^{++}	$d^{++}z^{+}$	$=$	$=$	$=$	$=$	$=$	V	V	V
19	$(bd)^{++}$	$(ab)^{++}w^{+}$	$=$	$=$	$=$	$=$	$=$	V	V	V
20	$(bc)^{++}$	$d^{++}(wxyz)^{+}$	$=$	$=$	$=$	$=$	$=$	V	U	U
21	a^{--}	$(wxyz)^{-}$	V	V	V	V	V	U	V	V
22	b^{--}	$b^{--}x^{-}$	$=$	$=$	$=$	U	U	U	U	U
23	c^{--}	$d^{--}w^{-}$	$=$	$=$	$=$	$=$	$=$	U	U	U
24	$(bd)^{--}$	$(ab)^{--}w^{-}$	$=$	$=$	$=$	$=$	$=$	U	U	U
25	$(bd)^{--}$	$a^{--}(wxyz)^{-}$	$=$	$=$	$=$	$=$	$=$	U	V	V
26	$(ab)^{++}(wxyz)^{-}$	a^{++}	$=$	$=$	V	U	U	U	U	U
27	$(bd)^{++}(wxyz)^{-}$	c^{++}	$=$	$=$	V	$=$	$=$	U	U	U
28	a^{--}	$(wxyz)^{+}(ac)^{--}$	$=$	$=$	V	U	U	V	U	U
29	c^{--}	$(wxyz)^{+}(bd)^{--}$	$=$	$=$	V	$=$	$=$	V	U	U
30	$d^{++}w^{-}$	d^{++}	$=$	$=$	V	V	V	V	V	V
31	$b^{++}w^{-}$	a^{++}	$=$	$=$	V	$=$	$=$	V	V	V
32	$c^{--}w^{+}$	c^{--}	$=$	$=$	U	U	U	U	U	U
33	$d^{--}w^{+}$	a^{--}	$=$	$=$	U	$=$	$=$	U	U	U

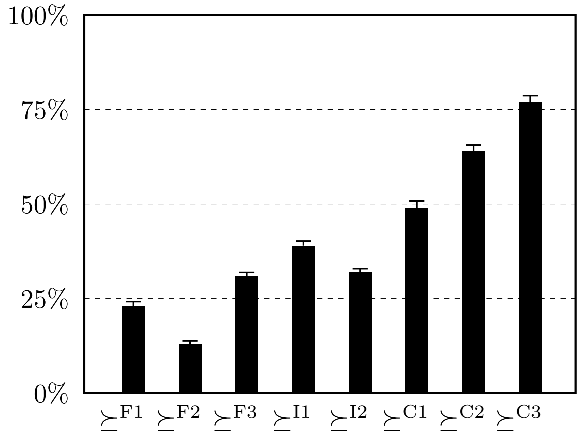

Figure 1. Average percentage of answers correctly predicted by each heuristic.

Participants could choose one of four responses: (a) choose option U, (b) choose option V, (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.

Table **??** displays the entire set of situations, together with the predictions of the eight heuristics in each situation. The 33 situations were selected to emphasize the different behaviors of the 8 heuristics. On average, the overlap in the predictions of any two heuristics was 41%. The overlap was slightly greater within each family of heuristics: 54% within the Focus family, 67% within the Inclusion family, and 58% within the Cardinality family. The overlap in predictions between the Focus and Inclusion families was also 58%, a high figure that makes sense since the Inclusion heuristics are refined versions of the Focus heuristics. On the contrary, there was little prediction overlap between the Inclusion and Cardinality families (35%), and even less between the Focus and Cardinality families (21%).

5. Results and Discussion

5.1. Overall Accuracy

Table **??** shows the repartition of participants' answers in the 33 situations. As a general index of descriptive validity, we computed the accuracy of each heuristic—that is, the average number of answers it correctly predicts, across participants. Figure **??** displays the accuracy of the 8 heuristics, in percentage form.

Figure **??** clearly shows the superiority of the Cardinality family of heuristics over the Focus and Inclusion families. Furthermore, it suggests that Levelwise Tallying ($\succeq^{C3}$) has by far the greatest descriptive validity, with an overall accuracy of 77%. Indeed, Levelwise Tallying always provided the most accurate predictions of *all* participants' choices, at the individual level. The second best heuristic was always Bivariate Levelwise Tallying $\succeq^{C2}$. On average, Levelwise Tallying predicted a reliably larger number of answers than did Bivariate Levelwise Tallying, $t(61) = 11.1$, $p < .001$, $d = 1.0$. Similar results held when comparing Bivariate Levelwise Tallying to all other procedures, all ts > 4.7, all ps $< .001$.

Let us insist on this general result before we proceed to more fine-grained considerations about participants' choices. It is *not* the case that some participants leaned towards one heuristic while some leaned towards another. The responses of *any one of the 62 participants* were always closer to the predictions of Levelwise Tallying than to the predictions of the other heuristics.

5.2. Focus Vs. Inclusion Vs. Cardinality

A closer look at the results confirms the overall impression given by the accuracy index. First, Focus heuristics are of limited validity because they too easily predict indifference, based on arguments at the highest level of importance—they suffer from what has been called a "drowning" problem, see **?** (?); **?** (?). Situation 17 provides a striking example of that shortcoming. In that situation, the choice is between getting the rare stamp BANTA, or getting this same rare stamp *plus* the common stamp ZAM. Unsurprisingly, participants unanimously preferred the second option. Focus heuristics cannot account for this preference, as they all disregard here the small bonus of getting a common stamp. Other illustrations of this discrepancy between the predictions of the Focus heuristics and the choices made by the participants can be found in situations 26–33.

Inclusion heuristics solve that problem by dismissing arguments that appear in both options. That way, the choice between BANTA plus ZAM or BANTA alone comes down to the choice between ZAM or nothing at all. This simplification does capture participants' preferences in situation 17 (as well as in situations 26, 28, 30, and 32), but it fails to handle situations such as 18. In that situation, the choice is in between getting the rare stamp CASSA or getting the rare stamp DIDOT plus the common stamp ZAM. Again, participants unanimously preferred the

Table II. Choices made by participants (in % of answers) in the 33 experimental situations. An option described as $a^{++}(xy)^{-}$ has one very positive feature a and two mildly negative features x and y. $\emptyset$ is the null option. U, V, $=$, and $\perp$ resp. read 'prefer U', 'prefer V', 'indifferent', 'options are incomparable'.

	Option U	Option V	U	V	$=$	$\perp$
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

second option, although the Inclusion heuristics predicted indifference: Since the arguments at the highest importance level are not the same, they cannot be simplified, and the heuristics cannot consider the weaker argument in favor of the second option. Inclusion heuristics thus cannot capture the choices made by participants in that situation, as it is also the case in situations 27, 29, 31, and 33.

Cardinality heuristics do much better on these situations—bar the notable failure of the Tallying heuristic to predict participants' preference in situation 29. Since the Tallying heuristic does not make any difference between a strong argument and a weak one, it prefers an option with two strong cons and four weak pros (i.e., a net score of +2) to an option with one strong con only (i.e., a net score of −1), while 86% of participants show the opposite preference. This blindness to the importance of arguments is responsible for the poor accuracy of the Tallying heuristic in situations 1–2, 6–9, 16, 20, 21, 25, and 28, and more generally for the disappointing performance of this heuristic compared to the other Cardinality heuristics.

5.3. Bivariate Vs. Univariate Levelwise Tallying

Only two heuristics $\succeq^{C2}$ and $\succeq^{C3}$, with their levelwise and cardinality-based approach, appear to predict participants' choices reasonably well. What is more, the Levelwise Tallying heuristic $\succeq^{C3}$ fares much better than its bivariate version $\succeq^{C2}$. Not only $\succeq^{C3}$ has greater overall accuracy than $\succeq^{C2}$, not only $\succeq^{C3}$ predicts *each and every* participant's choices better than $\succeq^{C2}$, but $\succeq^{C3}$ beats $\succeq^{C2}$ in all the situations that were selected to compare them (3–5 and 10–15).

Situations 3–5 and 10–15 were designed so that Levelwise Tallying $\succeq^{C3}$ predicted a strict preference, whilst Bivariate Levelwise Tallying $\succeq^{C2}$ predicted incomparability between the two options. For example, in situation 12, option U has one strong pro and four less important cons, while option V has two strong pros and one strong con. Bivariate Levelwise Tallying went no further than the highest level of argument importance, because there was already a difference in cardinality between the two options, on the positive side as well as on the negative side. The problem was that V won on the positive side (two strong pros against one strong pro), whilst U won on the negative side (no strong con against one strong con). Thus, no decision could be made. In contrast, Levelwise Tallying considered that the options were matched at the highest level of importance, as both had a net score of −1 at this level. Thus, it stepped down one level of importance, and took into account the fact that, at this less important level, U had four cons whilst V had none. It thus predicted a preference for V, which was indeed manifested by 86% of the participants. Similar results were observed for all other relevant situations (although they were not as clear cut with respect to situations 14 and 15).

5.4. Unexpected Answers

Considerations of overall accuracy as well as detailed examination of each situation suggest that Levelwise Tallying $\succeq^{C3}$ is the closest approximation of how participants balanced the pros and cons in our experiment. The performance of Levelwise Tallying is not, however, perfect. In two situations in particular, Levelwise Tallying predicted a preference that was shared by less than 40% of participants. In situation 3, the choice was to get a rare stamp in exchange for another rare stamp, or to get nothing and give away nothing. Levelwise Tallying predicts indifference between the two options, but participants either found the options incomparable or preferred the second option. In situation 11, the choice was to give away one common stamp and get nothing in return, or to get a rare stamp in exchange for another rare stamp. Levelwise Tallying predicts a preference for the second option, but participants typically preferred the first option.

There are several possible explanations for these discrepancies between the prediction of Levelwise Tallying and participants' actual choices, e.g. the possibility of an endowment effect or a negativity bias. As these explanations point to possible modifications of the heuristic, we discuss them in the *Perspectives* section below, where we also discuss the relation of Take the Best and Cumulative Prospect Theory to the Levelwise Tallying heuristic.

6. Perspectives

In this article, we have provided a formal framework for the study of decisions based on balancing the pros and cons of two options. We allowed the pros and cons to be of different importance, but we did not require this importance to be precisely quantified. We presented eight possible heuristics for making such decisions, and tested the predictions of these heuristics against the choices made by human participants on a selection of 33 situations. However intuitively appealing they might have seemed, most of our heuristics did not fare well in this test. Only Levelwise Tallying did show good accuracy in predicting participants' preferences.

Moreover, close examination of the data showed no major shortcoming of the Levelwise Tallying heuristic, bar one: participants disliked options featuring one strong pro and one strong con, to which they preferred the null option (no pro, no con) or even an option with a weak con an no pro. We consider two ways of accounting for these preferences: a narrow account, based on the endowment effect; and a broader account, based on the negativity bias.

6.1. Endowment effect or Negativity bias?

In our experiment, pros and cons were always related to the trading of goods (namely, stamps). A pro always came down to getting some good, and a con always came down to giving away some good. Now, it is a well-known result (known as the *endowment effect*) that people see more value in a good they own, than in a similar good when they are looking to acquire it (**?**; **?**)—and more generally, that the value function for losses is steeper than the value function for gains (**?**).

Thus, it could be that participants were reluctant to trade one rare stamp for another because they valued their own rare stamp more than the one they would get, and that they were even ready to give away one common stamp to avoid that exchange. Note, however, that this explanation only points at some idiosyncracy of our material. Pros and cons are not always about goods one will get or give away, as illustrated in our Emma example: Movie 1 is by Emma's favorite director, but it has attracted terrible critics. Nothing in these two arguments relate to material goods. Thus, if what we have observed was only a manifestation of the endowment effect, we should not expect Emma to give precedence to the con over the pro and to stay at home. We would rather expect Emma to be indifferent between staying at home and watching the movie.

However, we believe that there might be more than just an endowment effect in the greater weight some participants placed on strong cons, compared to strong pros. Cacioppo and colleagues (**?**; **?**; **?**) postulated the existence of two motivational systems running in parallel, one for the processing of negative affects, and another for the processing of positive affects. Most importantly, they have suggested that the positive motivational system is characterized by a *positivity offset*, whilst the negative motivational system is characterized by a *negativity bias*.

The positivity offset represents the tendency of the positive motivational system to respond more than the negative motivational system for low levels of evaluative input (say, for weak arguments). In contrast, the negativity bias represents the tendency of the negative system to respond more than the positive system for high levels of evaluative input (say, for strong arguments). From that perspective, we should expect strong cons to weight slightly heavier than strong pros, but also, and conversely, weak pros to weight slightly heavier than weak cons. Our data are certainly consistent with the first prediction, but are silent about the second one, as none of our situations pitted weak pros against weak cons. However, data reported in **?** (**?**) would appear to be consistent with both predictions. Whether pros and cons decisions are generally subject to both a positivity offset and a negativity bias is an important question for future research—if they are, the question will arise of how to formalize these two features, as it is not quite clear right now how the Levelwise Tallying heuristic might accommodate them.

6.2. Levelwise Tallying, Take the Best, and Cumulative Prospect Theory

In this final section, we briefly discuss how *Cumulative Prospect Theory* (**?**) and the *Take the Best* heuristic (**?**) can accommodate pros and cons decisions, and the relation of these two approaches to Levelwise Tallying.

Cumulative Prospect Theory (CPT) assumes that potential gains and losses are measured by means of two capacities σ^+ and σ^-, respectively. The greater is $\sigma^+(U^+)$, the more appealing is the positive side of option U (the potential gains); and the greater is $\sigma^-(U^-)$, the more repulsive is the negative side of U (the potential losses). The *net predisposition* for an option U can then be computed as the difference $\sigma^+(U^+) - \sigma^-(U^-)$. Note that this computation supposes a quantitative assessment of how important the pros and cons of option U are. However, it can be shown that Levelwise Tallying provides a qualitative counterpart to CPT. Indeed, as argued in **?** (**?**), for all U, V of 2^X, there are two capacities σ^+ and σ^- such that:

$$U \succeq^{\mathrm{C3}} V \iff \sigma^+(U^+) - \sigma^-(U^-) \geq \sigma^+(V^+) - \sigma^-(V^-).$$

This should not come as a surprise. Indeed, comparing the net predispositions for U and V is equivalent to comparing $\sigma^+(U^+) + \sigma^-(V^-)$ to $\sigma^+(V^+) + \sigma^-(U^-)$. Changing $+$ into max, we get the $\succeq^{\mathrm{F2}}$ Focus heuristic, of which Levelwise Tallying is a refinement. In other words, Levelwise Tallying is a refinement of what was already a qualitative counterpart to CPT.

In addition to providing a qualitative counterpart to CPT, Levelwise Tallying is a generalization of the oft-studied (if controversial) Take the Best heuristic (TTB). This heuristic was developed for situations of paired comparisons, where two options are compared based on the values they take on a series of binary cues $c_1, c_2, \ldots, c_n$. TTB requires a strict ordering of the cues by validity. Then, applying TTB amounts to considering the cues in decreasing order of validity, and to stopping as soon as one cue discriminates between the two options.

A large body of analytical and empirical research has investigated the performance of this heuristic (**?**; **?**; **?**; **?**), as well as its descriptive validity (**?**; **?**; **?**; **?**; **?**). It has generally been shown that TTB fares quite well compared to sophisticated regression models, and that its use by decision makers is influenced (in particular) by the cost of obtaining cues.

How would TTB apply to pros and cons decisions? Suppose that the importance of each argument can be assessed with such precision that no two arguments share the same level of importance. Of two pros, one is always more compelling than the other; of two cons, one is always more repulsive than the other; and, perhaps less plausibly, of one pro and one con, one is always more attractive than the other is repulsive, or vice-versa. Under this assumption, for all α in $[0_L, 1_L]$, one and only one of the following is true: (1) $|U_\alpha| = |V_\alpha| = 0$, (2) $|U_\alpha| = 0$ and $|V_\alpha| = 1$, or (3) $|U_\alpha| = 1$ and $|V_\alpha| = 0$. Now we can frame options U and V in such a way that TTB will be applicable, by giving them a value on a series of strictly ordered "cues." $\forall \alpha \in [0_L, 1_L]$,

$$c_\alpha(U) = \begin{vmatrix} 1 & \text{if} \quad |U^+_\alpha| = 1 \text{ or } |V^-_\alpha| = 1 \\ 0 & \quad \text{otherwise} \end{vmatrix}$$

In such a situation, TTB makes exactly the same choices as Levelwise Tallying.[2] However, as soon as the granularity of argument importance gets coarser, chances are that several arguments will share the same (highest) level of importance. When this happens, TTB cannot make any decision anymore—but Levelwise Tallying still can, with commendable descriptive validity. In that sense, Levelwise Tallying (and, to some extent, the other heuristics we have considered) is a natural generalization of TTB to cues of coarser granularity.

In conclusion, Levelwise Tallying is a generalization of both TTB and CPT. It generalizes the former to cues of coarse granularity, and the latter to qualitative pros and cons. At one end of the spectrum (when arguments can be totally ordered by rank of qualitative importance), Levelwise Tallying turns into TTB; at the other end (when the importance of arguments can be quantitatively assessed), Levelwise Tallying turns into CPT; in between, Levelwise Tallying provides the missing link between Take the Best and Cumulative Prospect Theory.

[2] In fact, it makes the same predictions as all the heuristics in Section **??**, bar simple Tallying C1.

References

Amgoud, L., Bonnefon, J. F. and Prade, H.: (2005), An argumentation-based approach for multiple criteria decision. In: L. Godo (ed.): *Proceedings of the European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty—Lecture Notes in Artificial Intelligence 3571*: pp. 269–280, Springer Verlag, Berlin.

Benferhat, S., Cayrol, C., Dubois, D. , J. Lang, and Prade, H.: (1993), Inconsistency Management and Prioritized Syntax-Based Entailment. In: R. Bajcsy (ed.): *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*: pp. 640–645, Morgan Kaufmann, San Mateo, CA.

Benferhat, S., Dubois, D., Kaci, S., and Prade, H.: (2006), Bipolar possibility theory in preference modeling: Representation, fusion and optimal solutions. *Information Fusion* **7**(1): 135–150.

Bröder, A.: (2000), 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**(5): 1332–1346.

Bröder, A. and Schiffer, S.: (2003), "Take The Best" versus simultaneous feature matching: Probabilistic inferences from memory and effects of representation format. *Journal of Experimental Psychology: General* **132**(2): 277–293.

Cacioppo, J. T. and Berntson, G. G.: (1994), Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin* **115**(3): 401–423.

Cacioppo, J. T., Gardner, W. L., and Berntson, G. G.: (1999), The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology* **76**(5): 839–855.

Dawes, R. M. and Corrigan, B.: (1974), Linear models in decision making. *Psychological Bulletin* **81**(2): 95–106.

Dubois, D. and Fargier, H.: (2005), On the qualitative comparison of sets of positive and negative affects. In: L. Godo (ed.): *Proceedings of the European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty—Lecture Notes in Artificial Intelligence 3571*: pp. 305–316, Springer-Verlag, Berlin.

Dubois, D. and Fargier, H.: (2006), Qualitative decision making with bipolar information. In: P. Doherty, J. Mylopoulous, and C. Welty (eds.): *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning*: pp. 175–186, AAAI Press, Menlo Park, CA.

Dubois, D. and Prade, H.: (2004), Possibilistic logic: a retrospective and prospective view. *Fuzzy Sets and Systems* **144**(1): 3–23.

Einhorn, H. J. and Hogarth, R. M.: (1975), Unit weighting schemes for decision making. *Organizational Behavior and Human Decision Processes* **13**(2): 171–192.

Fargier, H., Dubois, D., and Prade, H.: (1996), Refinements of the maximin approach to decision-making in fuzzy environment. *Fuzzy Sets and Systems* **81**(1): 103–122.

Gigerenzer, G. and Goldstein, D. G.: (1996), Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review* **103**(4): 650–669.

Hogarth, R. M. and Karelaia, N.: (2005), Simple models for multi-attribute choice with many alternatives: When it does and does not pay to face trade-offs with binary attributes. *Management Science* **51**(12): 1860–1872.

Ito, T. A. and Cacioppo, J. T.: (2005), Variations on a human universal: Individual differences in positivity offset and negativity bias. *Cognition and Emotion* **19**(1): 1–26.

Johnson, E. J., Hershey, J., Meszaros, J., and Kunreuther, H.: (1993), Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty* **7**(1): 35–51.

Kahneman, D., Knetsch, J. L., and Thaler, R. H.: (1991), The Endowment Effect, Loss Aversion, and Status Quo Bias. *Journal of Economic Perspectives* **5**(1): 193–206.

Kahneman, D. and Tversky, A.: (1979), Prospect theory: An analysis of decision under risk. *Econometrica* **47**(2): 263–291.

Katsikopoulos, K. V. and Martignon, L.: (2006), Naïve heuristics for paired comparisons: Some results on their relative accuracy. *Journal of Mathematical Psychology* **50**: 488–494.

Martignon, L. and Hoffrage, U.: (2002), Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision* **52**(1): 29–71.

Newell, B. R. and Shanks, D. R.: (2003), Take-the-best or look at the rest? Factors influencing 'one-reason' decision making. *Journal of Experimental Psychology: Learning. Memory, and Cognition* **29**(1): 53–65.

Newell, B. R., Weston, N. J., and Shanks, D. R.: (2003), Empirical tests of a fast and frugal heuristic: Not everyone "takes-the-best". *Organizational Behavior and Human Decision Processes* **91**(1): 82–96.

Rieskamp, J. and Hoffrage, U.: (1999), When do people use simple heuristics and how can we tell?. In: G. Gigerenzer, P. M. Todd, and the ABC group (eds.): *Simple heuristics that make us smart*: pp. 141–167, Oxford University Press, Oxford.

Tversky, A., and Kahneman, D.: (1992), Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* **5**(4): 297–323.

Wald, A.: (1971), Statistical Decision Functions. Wiley, New York. (Original work published 1950.)