Expert Systems With Applications 238 (2024) 121922

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

# Ordering vs. AHP. Does the intensity used in the decision support techniques compensate?

Carlos Sáenz-Royo [a,*], Francisco Chiclana [b,c], Enrique Herrera-Viedma [c]

[a] University of Zaragoza, Department of Business Organization and Management, C/Pedro Cerbuna, 12, 50009 Zaragoza, Spain
[b] Institute of Artificial Intelligence, School of Computer Science and Informatics, De Montfort University, Leicester, UK
[c] Andalusian Research Institute on Data Science and Computational Intelligence (DaSCI), University of Granada, Granada, Spain

ARTICLE INFO

*Keywords:*
AHP
IBR
Decision Support Systems
Expertise
Intensity Judgment

ABSTRACT

The manifestation of the intensity in the judgment of one alternative versus another in the peer comparison processes is a central element in some decision support techniques, such as the Analytical Hierarchy Process (AHP). However, its contribution regarding quality (expected performance) with respect to the priority vector has not been evaluated so far. Using the Intentional Bounded Rationality Methodology (IBRM), this work analyzes the gains obtained from requiring the decision-maker to report an intensity judgment in pairs (AHP) with respect to a technique that only requires expressing a preference (Ordering). The results show that when decision-makers have low levels of expertise, it is possible that a less informative and computational cheap technique (Ordering) performs better than a more informative and computational expensive one (AHP). When decision-makers have medium and high levels of expertise, AHP technique obtains modest gains with respect to the Ordering technique. This study proposes a cost-benefit analysis of decision support techniques contrasting the gains of a technique that requires more resources (AHP) against other that require less resources (Ordering). Our results can change the managing approach of the information obtained from experts' judgments.

## 1. Introduction

People's judgments can provide useful information for forecasting and decision-making. In decision analysis, especially when quick decisions are required, people are often the only available source of information regarding some important variables or relationships. Choosing between different alternatives may not be an easy task and people may not be always able to choose the best option. Since people make mistakes in their judgments, the probability of being right or wrong depends on a person's knowledge (expertise) regarding their expressed judgments (Chinchanachokchai, Thontirawong, & Chinchanachokchai, 2021; Liu, Eckert, & Earl, 2020; Nemeshaev, Barykin, & Dadteev, 2021; Sáenz-Royo, Chiclana, & Herrera-Viedma, 2023b). Decision support techniques aim to establish procedures to improve the results of human decisions by providing scientific rigor to the treatment of expressed judgments (Keeney, 1992; Moreno-Jiménez, Aguarón-Joven, Escobar-Urmeneta, & Turón-Lanuza, 1999). In short, decision support techniques evaluate people's judgments to detect and eliminate possible erroneous judgments and build a logical system of relationship between them to obtain a priorities vector (solution) that ranks the alternatives or indicates the alternative with best performance.

Most of the comparisons between decision support techniques analyze in a descriptive way the differences in their used methodologies. (Opricovic & Tzeng, 2007) empirically illustrate when the application of different techniques leads to differences in conclusions, and they establish disagreement indices. Zanakis, Solomon, Wishart, and Dublish (1998) simulate scenarios with a different number of alternatives, criteria, and distributions to analyze the levels of similarity between the compared techniques. Wallenius et al. (2008) quantify the use of each technique in the literature, and they give a ranking of the most used as a proxy variable of their "quality". Belton (1986) makes a theoretical and practical mathematical analysis of how different techniques manage experts' judgments. Triantaphyllou (2000) performs a detailed comparison work between techniques to analyze the sensitivity of the results to changes of value of the parameters. It is noticed though that none of these studies specify whether one technique performs better than another. The Intentional Bounded Rationality Methodology (IBRM) of Sáenz-Royo, Chiclana, & Herrera-Viedma, (2023a) contributes to closing this gap since it allows quantifying the expected performance of each decision support technique with respect to two explanatory variables:

* Corresponding author.
*E-mail address:* csaenz@unizar.es (C. Sáenz-Royo).

https://doi.org/10.1016/j.eswa.2023.121922
Received 12 January 2023; Received in revised form 28 September 2023; Accepted 28 September 2023
Available online 2 October 2023

the difficulty of the problem (difference between the latent performance of the alternatives) and the decision-makers' expertise.

The specialized literature includes many developed complex calculation systems to improve the result of human decisions. As an example, we can cite the development of methodologies based on fuzzy sets to enhance the handling of information of the Analytical Hierarchy Process (AHP), which is computationally more expensive and that ultimately requires a defuzzification process to obtain the final decision result.[1] This study questions whether the gains from a decision support technique are enough to compensate the time and effort invested in experts and computation. This idea is inspired by techniques, such as Best-Worst Multi-Criteria (BWM) (Rezaei, 2015), that significantly reduce the number of paired judgments required. The balance between profits and costs must remain in the minds of managers. To illustrate the cost-benefit analysis of decision support techniques, this paper compares two techniques: Ordering and AHP. The fundamental difference between these two techniques resides in how pairs of alternatives are compared: while AHP requires experts to provide judgments in the form of degree of cardinal preference of one alternative over another, the Ordering technique only requires judgments regarding which alternative of the pair being compared is the preferred one. Although the AHP technique is more informative than the Ordering technique, it is obvious that AHP requires more resources, i.e., more effort and more computational expensive, than Ordering. In addition, while the requirement of cardinal consistency of judgments in decision support techniques affects the level of AHP performance, it is not known if this will be the case with Ordering, since in the latter case the rejection of judgments requires an ordinal inconsistency. Thus, addressing this question will allow to explain a consistency part of the performance of the priority vector and, ultimately, lead to a better understanding of the design of decision support techniques.

The main contributions of the work reported herein are:

1. It presents a new approach that questions decision support techniques as a balance between cost and benefit;
2. It details a simplification of the AHP to establish a decision support technique based on simple preference judgments without cardinal intensity;
3. It evaluates whether it is possible that the information on the intensity in the judgments provides negative performances with respect to simpler decision support techniques such as those based on Ordering (preference judgments);
4. It analyzes how the incorporation of a consistency requirement affects the performance difference between Ordering and AHP (intensity);
5. It carries out a classification of Ordering errors and their probabilities;
6. It shows that AHP with the consistency requirement has better-expected performance than Ordering on most occasions, although the gains are small.

The fact that, in some cases, Ordering may outperforms AHP is a striking and counterintuitive result because, by definition, AHP is more informative than Ordering. The results show how incorporating the judgment of intensity increases the percentage of correct priority vectors but also significantly increases the probability of inconsistency and, using the illustrative example of Sáenz-Royo et al. (2023a) study, it is shown that AHP can have lower performance than Ordering. Finally, a robustness exercise is carried out with three different performances of alternative $A_2$ in the mentioned illustrative example (scenarios) to study the critical values of expertise needed to choose one or the other decision support technique.

The rest of this report is structured as follows: Section 2 describes the fundamentals of IBRM. Section 3 presents the Ordering and AHP decision support techniques, the concepts of ordinal and cardinal consistency, and the application of the judgment probabilities of the IBRM to these techniques. This section also includes a brief review of the differences between Inconsistency and Error. Section 4 studies the types of problems that can be solved with the proposed approach. Section 5 reports on the results of the Ordering and the AHP techniques when applied to an illustrative case and highlights their differences. In section 6, a robustness exercise is carried out to expand the results of the previous section for different levels of expertise of the decision-maker and for different scenarios (different performances for alternative $A_2$ of the illustrative case). Finally, the last section presents the most relevant conclusions.

## 2. Intentional bounded rationality methodology

On some occasions, decision-makers are asked about issues with no clear measurement of the solution to the problem; this hides possible decision-maker's errors. In these cases, decision support techniques are proposed to be applied with the aim to provide guarantees that this does not happen. Traditionally, the quality of decision support techniques and decision-makers have been evaluated a posteriori through individual aspects such as hesitation, interest, and consistency (Díaz, Fernández, Figueira, Navarro, & Solares, 2022; Herrera-Viedma, Herrera, Chiclana, & Luque, 2004; Sellak, Ouhbi, Frikh, & Ikken, 2019) or through group-level aspects such as consensus and collective consistency (Moral, Chiclana, Tapia, & Herrera-Viedma, 2018; Xu, Liu, Wang, & Shang, 2022). Even without knowing the true values, the decision support techniques approximate the correct solution under the hypothesis that the decision-maker may be wrong although, in the long run, they present a distribution of successes greater than of errors, that is, finally the correct option ends up being imposed. According to this premise, a higher level of consensus indicates a greater certainty about the goodness of the choice, given the greater probability of being right than wrong; and the level of coherence indicates the logical robustness of the choices made by the decision-maker or group of decision-makers.

However, on the same premise of experts being more likely to be correct than to err in their judgments, IBRM allows an a priori evaluation of the different elements of decision support techniques. IBRM builds an automaton that represents the experts' intentional bounded rationality (Sáenz-Royo, Chiclana, & Herrera-Viedma, 2022). The intentional bounded rationality links the mechanisms that govern human cognition and the expert's decisions by collecting in a functional way the conditioning factors of the human way of thinking and allowing decision-makers to make mistakes (Simon, 1947). This conceptual framework relates the complexity of the decision, understood as the difference in the latent performances of the alternatives, with the expertise of the decision-maker, obtaining the reliability of the person issuing the judgments, as an a priori variable of bounded but intentional rationality. This methodology is distinguished from the traditional ones that are limited to introducing random noises in the decisions of the decision-maker (Csaszar, 2013; Hogarth, 1975; O'Hagan et al., 2006; Ravinder, 1992; Vargas, 1982; Wallsten & Budescu, 1983).

The intentional bounded rationality proposes a logistic probability distribution, in which the probability that a decision-maker chooses

[1] Fuzzy AHP (FAHP) introduces fuzzy sets into AHP to resolve uncertainties of expert's preferences (Song, Zhu, Jia, & He, 2014). Mosadeghi, Warnken, Tomlinson and Mirfenderesk (2015) has shown the superiority of FAHP over AHP. The Intuitive Fuzzy AHP (IFAHP) technique (Xu & Liao, 2014) allows for greater precision than FAHP by simultaneously expressing affirmation, denial, and hesitation caused, in many practical situations, by insufficient decision-maker domain knowledge (Garg, 2019). However, while a total order is possible with sharp numbers, this is not the case with intuitive fuzzy numbers where only partial rankings are possible. Since there is no existing isomorphic operations or relations between intuitive fuzzy sets and crisp numbers, when the attribute values are crisp numbers and the attribute weights are intuitive fuzzy numbers, the subsequent decision-making cannot proceed without defuzzification.

alternative $A_i$ with value $V_i$ depends on the decision-maker's ability to process information and the complexity of decision (performance of the alternative studied with respect to the performance of the other alternatives). Thus, the probability $p_{\beta i}$ that the decision-maker chooses alternative $A_i$ over all the other alternatives is the below function of the relative weight of its performance value with respect to the value of the other alternatives:

$$p_{\beta i} = p(A_i) = \frac{e^{\beta \frac{V_i}{\sum_{k=1}^{n} V_k}}}{\sum_{j=1}^{n} e^{\beta \frac{V_j}{\sum_{k=1}^{n} V_k}}} = \frac{1}{1 + \sum_{j \neq i}^{n} e^{\frac{\beta (V_j - V_i)}{\sum_{k=1}^{n} V_k}}}. \quad (1)$$

The parameter $\beta$ measures the decision-maker's ability to process information; its value is limited to the problem studied and, therefore, the same decision-maker can show different capacities for different problems. When the decision-maker does not know anything about the problem, then $\beta = 0$, and all the alternatives have the same probability of being chosen regardless of the relative value of each one of them. High values of $\beta$ imply a high capacity to discern the best alternative.

The intentional bounded rationality in IBRM defines the behavior of an automaton that represents a decision-maker. Starting from a situation in which the performances of the alternatives are known, the level of error of the automaton is determining by the probability of choice of each alternative according to the equation (1). The error of the automaton depends on the difference in the performance of the alternatives and the level of expertise ($\beta$). The property that the automaton shows a greater probability to choose the alternative with the highest performance guarantees that the decision support techniques add value to the search for the optimal option.

The IBRM allows a priori evaluation of the contribution to the quality of the decision of each requirement established by a decision support technique. The systematic procedure of the judgments expressed by people established by decision techniques provides a level of quality of the decision obtained (Keeney, 1992; Moreno-Jiménez et al., 1999). However, some elements of the procedures required by decision support techniques have questionable contributions (Liu, Qiu, & Zhang, 2021; Rezaei, 2015). In this sense, Sáenz-Royo et al., (2023a) have shown that the contributions of the consistency requirements can be meager. This work analyzes the quality improvement obtained when decision-makers are required to show intensities in their judgments.

## 3. Ordering vs intensity preference (AHP)

The proposed decision support techniques are based on the comparison of alternatives by pairs, assigning a ranking or a value on a scale of intensity of preference for one alternative over the other. The pairwise comparison has its origin in the psychological works of Thurstone (1927) and is used in multiple decision support techniques (Figueira, Greco, Ehrgott, & Henggeler Antunes, 2005).

### 3.1. Ordering

The Ordering decision support technique applied to a set of $n$ alternatives compared by pairs leads to the construction of a matrix of judgments $R$ of dimensions $n \times n$,

$$R = \begin{pmatrix} \widehat{r}_{11} & \cdots & \widehat{r}_{1n} \\ \vdots & \ddots & \vdots \\ \widehat{r}_{n1} & \cdots & \widehat{r}_{nn} \end{pmatrix}$$

with comparison judgment $\widehat{r}_{ij}$ being only one of two values: $\widehat{r}_{ij} = 1$ means that the decision-maker prefers alternative $A_i$ to alternative $A_j$ (therefore $\widehat{r}_{ji} = 0$), while $\widehat{r}_{ij} = 0$ in case the decision-maker prefers $A_j$ to $A_i$ (hence $\widehat{r}_{ji} = 1$). Using performance values of alternatives, it is $\widehat{r}_{ij} = 1$ when $V_i$ is higher than $V_j$.

From an expert's matrix of judgments, the Ordering technique derives a vector of priorities of alternatives by the application of a score function, $\widehat{r}_i = \sum_{j=1}^{n} \widehat{r}_{ij} | i \neq j$, that assigns each alternative the number of alternatives it is preferred to or with lower performance values. Thus the Ordering technique priority vector of matrix $R$ is as $(\widehat{r}_1, \widehat{r}_2, \cdots, \widehat{r}_n)$, and selects the alternative with highest score value as the best one. Notice that the coherent reciprocity property of the ordinal preference judgments, $\widehat{r}_{ij} = 1 \Leftrightarrow \widehat{r}_{ji} = 0$, implies $\sum_{i=1}^{n} \widehat{r}_i = n(n-1)/2$.

The Ordering method makes it easy to reduce the problems of uncertainty and ambiguity of language used that come with expressing intensity (see discussion of these problems in Cai, Lin, Han, Liu, & Zhang, 2017). This study does not address the problem of the subjectivity of judgments, since our approach is epistemological (the solution is unique), but it is true that the intensities can be interpreted differently by different experts.

### 3.2. IBRM in Ordering

The IBRM utilizes an automaton that represents a rationally bounded but intentional decision-maker which provides a priori the probabilities of each paired judgment. For the automaton, as per (1), each pairwise comparison $(A_i, A_j)$ will have associated the following probability $p_{\beta ij}$ of choosing alternative $A_i$ versus choosing alternative $A_j$ :

$$p_{\beta ij} = \frac{e^{\beta \frac{V_i}{V_j}}}{e^{\beta \frac{V_i}{V_j}} + e^{\beta \frac{V_j}{V_i}}} = \frac{1}{e^{\beta \left( \frac{V_j}{V_i} - \frac{V_i}{V_j} \right)} + 1} \quad (2)$$

As elaborated above, the ordinal decision support technique allow to derive the priority vector (ranking) of the alternatives. Since the automaton probabilities verify the reciprocity property, $p_{\beta ji} = 1 - p_{ij}$, it is easy to see that each ordinal paired judgment of preference $\widehat{r}_{ij}$ has the corresponding automaton probability $p_{\beta ij}$. Thus, the IBRM allows to compute the probability of a priority vector. Consequently, the probability that the Ordering decision support technique chooses alternative $A_i$ as the best alternative, denoted by $p_{\beta i}$, can be defined as the sum of probabilities of all those priority vectors (rankings) that propose such alternative as the best.

The IBRM establishes that the automaton can manifest different judgments with different probabilities according to (2), which when applied into the Ordering technique leads to the building of a probability matrix to indicates the probability with which the automaton will choose each alternative in each pairwise comparison. When the alternatives are ranked by score/performance, this $(n \times n)$ matrix $P$ shows the probability of missing below the main diagonal and that of hitting above. The main diagonal would represent the probability of choosing each alternative with respect to itself and therefore no value is assigned to it.

$$P_\beta = \begin{pmatrix} - & p_{\beta 12} & \cdots & p_{\beta 1n} \\ p_{\beta 21} & - & \cdots & p_{\beta 2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{\beta n1} & p_{\beta n2} & \cdots & - \end{pmatrix}$$

The IBRM requires listing all the possible combinations of judgments that the automaton can manifest and their probabilities. For each ordinal judgment, the decision-maker must evaluate if $V_i > V_j$ ($\widehat{r}_{ij} = 1$) or $V_j > V_i$ ($\widehat{r}_{ij} = 0$). Since the number of ordinal reciprocal judgments a decision-maker must do is $n(n-1)/2$, the possible combinations of preference judgments that the decision-maker can do are $VR_2^{\frac{n(n-1)}{2}} = 2^{\frac{n^2-n}{2}}$. Each of these combinations generates a priority vector and its probability of being chosen by the automaton. Priority vectors with the same ranking of alternatives are grouped. The probability of a ranking is obtained as the sum of the probabilities of all the priority vectors in their corresponding group and the rankings are classified according to the impact of the error they carry.

The probability of choosing the correct ranking is analyzed and the (expected) performance of the decision support techniques can be calculated. This methodology allows evaluating the decision support technique a priori by incorporating the possibility that the automaton is wrong. The results obtained will depend on how the decision support technique handles the errors of the automaton.

### 3.3. Ordinal consistency

One way to improve the performance of a decision support technique is to establish a requirement for consistency in accepting the judgments of a decision-maker. A complete set of paired judgments can generate inconsistent information, mainly because it is inherent in human nature to be wrong sometimes.

The ordinal consistency of the decision-maker is defined as the property that verifies: if $V_i$ is greater than $V_j$ and $V_j$ is greater than $V_k$, then $V_i$ is greater than $V_k$. Therefore, ordinal inconsistency happens with a set of circular preference judgments (Kendall & Smith, 1940) by the decision-maker that verifies: $V_i > V_j$; $V_j > V_k$; $V_k > V_i$. This ability of decision-makers to compare alternatives and to maintain three-level transitivity is the first assumption made by Luce and Raiffa (1957). Gass (1998) determines that when all the decisions of the decision-maker are transitive, it is true that $\frac{n(n-1)(n-2)}{6} = \frac{1}{2}\sum_{i=1}^{n}\widehat{r}_i(\widehat{r}_i - 1)$, since the matrix $R$ of ranked alternatives has ones above the main diagonal and zeros below it. From here, the number of judgments that break transitivity is $c = \frac{n(n-1)(n-2)}{6} - \frac{1}{2}\sum_{i=1}^{n}\widehat{r}_i(\widehat{r}_i - 1)$.

To measure ordinal inconsistency as a relationship between the number of intransitive judgments over the total of possible transitive judgments, Kendall and Smith (1940) proposed a coefficient that they called the consistency coefficient ($Ke$):

$$Ke = \begin{cases} 1 - \dfrac{24c}{n(n^2-1)} & \text{when } n \text{ is even} \\ 1 - \dfrac{24c}{n(n^2-4)} & \text{when } n \text{ is odd} \end{cases}.$$

When $Ke = 1$, it is $c = 0$ and therefore there is no intransitivity; while the higher the value of $c$, the closer $Ke$ is to zero, the worst possible intransitivity situation. Harary and Moser (1966) showed that the maximum proportion of intransitivities over the total of possible relationships of the judgments shown by a decision-maker is around ¼, even for large values of $n$. Thus, it seems reasonable to require the decision-maker not to present any intransitivity in their judgments.

In relation to the IBRM, when the automaton has very high levels of expertise ($\beta$), it hardly makes mistakes in its judgments and the probability of an inconsistency appearing is very low. In the cases where the automaton does not show any inconsistency and the alternative $A_i$ is preferred in all paired judgments, the score assigned by the automaton will be $\widehat{r}_i = n-1 > \widehat{r}_j \forall j \neq i$, and $p_{\beta i}$ will be

$$p_{\beta i} = \prod_{j \neq i}^{n} p_{\beta ij} = \prod_{j \neq i}^{n} \frac{1}{e^{\beta\left(\frac{V_j}{V_i} - \frac{V_i}{V_j}\right)} + 1} \tag{3}$$

### 3.4. Intensity in AHP

The judgment of intensity refers to the valuation by the decision-maker of how many times (the performance of) one alternative is greater than (the performance of) another alternative (Fichtner, 1986). The intensity judgment adds information to an ordinal preference or judgment and it is a central element of decision support techniques such as the AHP. From the intensity judgments manifested by the decision-maker, AHP builds a preference comparison matrix that is solved algebraically to derive the priority vector of alternatives $(w_1, w_2, \cdots, w_n)$ (Saaty, 1980).

Let $M = (\widehat{a}_{ij})$ be the preference comparison matrix; with $\widehat{a}_{ij}$ being the decision maker's times the performance of alternative $A_i$, $V_i$, is higher than the performance of alternative $A_j$, $V_j$.

$$M = \begin{pmatrix} \widehat{a}_{11} & \cdots & \widehat{a}_{1n} \\ \vdots & \ddots & \vdots \\ \widehat{a}_{n1} & \cdots & \widehat{a}_{nn} \end{pmatrix}$$

As with ordinal judgments, the coherence reciprocity property is required from the decision-maker: $\widehat{a}_{ij} = 1/\widehat{a}_{ji}$. If $\widehat{a}_{ij} > 1$ then the decision-maker states that $V_i > V_j$ and therefore $A_i$ is preferred to $A_j$. From the reciprocal preference comparison matrix $M = (\widehat{a}_{ij})_{n\times n}$, a priority vector $\widehat{w}$, verifying $\widehat{a}_{ij} = \widehat{w}_i/\widehat{w}_j$ and $\sum_i^n \widehat{w}_i = 1$, is derived. The priority vector obtained by the AHP method represents the solution to the decision problem and takes the form $M\widehat{w} = n\widehat{w}$.

If there are no errors and the decision-maker is totally rational ($\beta = \infty$ in (2)), then $\widehat{a}_{ij} = \frac{V_i}{V_j}$. In this case, $M$ has rank 1, and equation $(M - nI)\widehat{w} = 0$ has a solution if and only if $n$ is an eigenvalue ($\lambda_i$) of $M$. Given that $M$ has rank 1, all the eigenvalues are equal to zero ($\lambda_i = 0$) except one $\lambda_{max} = \sum_i^n \lambda_i = tr(M) = n$. Although any column of $M$ is a solution to the system, the normalized solution is unique, specifically what has been called the priority vector $\widehat{w}$ where $\widehat{w}_i/\widehat{w}_j = V_i/V_j$.

When intensity evaluates tangible quantifiable criteria, its value is obtained directly from measured information, for example, distances (in meters), or times (in minutes), etc. Many times, the judgment of a decision-maker is required but a measure is not available to evaluate the intensity because it is complex information or qualitative criteria not amenable to be quantified. In these cases, $\widehat{a}_{ij}$ is provided using the numerical Table 1 given by Saaty (1977) with 17 verbal preference judgments, which has been widely used (Belton & Gear, 1983; Dyer, 1990; Ishizaka & Labib, 2011). The critical points between intervals represent the point from which the decision-maker changes from manifesting one verbal intensity to another. If a decision-maker were able to cardinally know the intensity of her preferences, then the critical points are the value from which the decision-maker changes section because the table has a discrete character. If the decision-maker considers that her/his preference $A_i$ over $A_j$ is 7.51, her/his preference is represented by an 8 (verbally: Between absolute and demonstrated importance $A_i$ over $A_j$) while if it is 7.49 it is represented by 7 (verbally: Demonstrated importance $A_i$ over $A_j$).

**Table 1**
Saaty (1977) preference scales equivalences between verbal judgments and cardinal valuation.

| PREFERENCE SCALES | |
|---|---|
| **The verbal expression judgments of the preference** | **Numerical expression** |
| Absolute importance $A_i$ over $A_j$ | 9 |
| Between absolute and demonstrated importance $A_i$ over $A_j$ | 8 |
| Demonstrated importance $A_i$ over $A_j$ | 7 |
| Between demonstrated and essential or strong importance $A_i$ over $A_j$ | 6 |
| Essential or strong importance $A_i$ over $A_j$ | 5 |
| Between essential or strong and weak importance $A_i$ over $A_j$ | 4 |
| Weak importance $A_i$ over $A_j$ | 3 |
| Between weak and equal importance $A_i$ over $A_j$ | 2 |
| Equal importance $A_i$ over $A_j$ and $A_j$ over $A_i$ | 1 |
| Between weak and equal importance $A_j$ over $A_i$ | 1/2 |
| Weak importance $A_j$ over $A_i$ | 1/3 |
| Between essential or strong and weak importance $A_j$ over $A_i$ | 1/4 |
| Essential or strong importance $A_j$ over $A_i$ | 1/5 |
| Between demonstrated and essential or strong importance $A_j$ over $A_i$ | 1/6 |
| Demonstrated importance $A_j$ over $A_i$ | 1/7 |
| Between absolute and demonstrated importance $A_j$ over $A_i$ | 1/8 |
| Absolute importance $A_j$ over $A_i$ | 1/9 |

Sáenz-Royo et al. (2023a) proposed a simplified Table 2 of it with four sections without the possibility of indifference, ($A_i$ is highly preferred to $A_j$, $A_i$ is preferred to $A_j$, $A_j$ is preferred to $A_i$, $A_j$ is highly preferred to $A_i$).

In this simplification, there are three critical points ($Z$) which are: $18/4$ above this intensity the decision-maker states that $A_i$ is "extremely preferred"; 1 from this intensity and down to $4/18$, the decision-maker shows $A_i$ "preferred" and below it shows $A_j$ "preferred" until $4/18$; and below $4/18$ it manifests $A_j$ "extremely preferred".

### 3.5. IBRM in AHP

The AHP technique requires the decision-maker to establish the intensity of preference for each paired judgment. The IBRM must not only provide the probability that the automaton manifests in favor of one or the other alternative, but it must also provide the probability that the automaton will manifest each of the established discrete intensity ranges. The simplified scale proposed by Sáenz-Royo et al. (2023a) provides 3 critical points ($Z$) that allow obtaining the probabilities with which an automaton with $\beta$ expertise will manifest each of the intensities. The calculation of these intensities requires an intermediate step, obtaining the probability that the automaton manifests an intensity greater than $Z$ with respect to the true intensity $V_i/V_j$

$$p_{\beta ij}(Z) = \frac{1}{e^{\beta\left(\left(Z-\frac{1}{Z}\right)-\left(\frac{V_i}{V_j}-\frac{V_j}{V_i}\right)\right)}+1} \tag{4}$$

Equation (2) is a particular case of (4) when $Z = 1$, that is when what is interesting to know is the probability that the automaton judges that $V_i > V_j$. The probability function representing the automaton is a logistic function whose mean is at the true value of intensity ($V_i/V_j$) and whose standard deviation depends on $\beta$.

The AHP also requires knowing which intensity range the decision-maker opts for. Let $k$ be the intensity valuation of a discrete interval of preference in which the decision-maker judgments that $V_i$ is $k$ times $V_j$ measured discretely ($\widehat{a}_{ij,k} = k \in \{1,\cdots,K\}$). To quantify the probabilities of each intensity interval that determines the IBRM ($p_{\beta ij,k}$), it is necessary to make the difference between the probability that the automaton considers that the intensity is greater than the lower critical point of interval $k$ ($Z_{k,l}$) and the probability that it considers the intensity to be greater than the upper critical point of interval $k$ ($Z_{k,h}$), that is, $p_{\beta ij}(\widehat{a}_{ij,k}) = p_{\beta ij}(Z_{k,l}) - p_{\beta ij}(Z_{k,h})$. Using (4), we have:

$$p_{ij}(k) = \frac{1}{e^{\beta\left(\left(Z_{k,l}-\frac{1}{Z_{k,l}}\right)-\left(\frac{V_i}{V_j}-\frac{V_j}{V_i}\right)\right)}+1} - \frac{1}{e^{\beta\left(\left(Z_{k,h}-\frac{1}{Z_{k,h}}\right)-\left(\frac{V_i}{V_j}-\frac{V_j}{V_i}\right)\right)}+1} \tag{5}$$

### 3.6. Consistency in AHP

When the matrix $M$ is cardinally consistent (transitive) then $\widehat{a}_{ik} = \widehat{a}_{ij}\widehat{a}_{jk}\left(\frac{V_i}{V_k} = \frac{V_i}{V_j}\frac{V_j}{V_k}\right)|\forall i \neq j \neq k$. A matrix $M$ is totally consistent if and only if $\lambda_{max} = n$; while if $\lambda_{max} > n$, then $M$ is not consistent. Saaty (1980) defines the consistency index of $M$ as.

**Table 2**
Simplification preference scales proposed by Sáenz-Royo et al. (2023a) between verbal judgments and cardinal assessment.

| PREFERENCE SCALES | |
|---|---|
| **The verbal expression judgments of the preference** | **Numerical expression** |
| $A_i$ is extremely preferable to $A_j$ | 25/4 |
| $A_i$ is preferable to $A_j$ | 10/4 |
| $A_j$ is preferable to $A_i$ | 4/10 |
| $A_j$ is extremely preferable to $A_i$ | 4/25 |

$CI_M = \frac{\lambda_{max}-n}{n-1}$, where $\lambda_{max} = \sum_{i=1}^{n} \widehat{a}_{ij}\frac{\widehat{w}_j}{\widehat{w}_i}$ indicates the cardinal difference between the decision-maker's judgment and the inverse of the priority estimate. This formulation indicates that the closer each paired judgment is to its estimate, the closer $\lambda_{max}$ will be to $n$ and $CI_M$ to 0. Saaty (1980) introduces the consistency ratio ($CR$) as the quotient between $CI_M$ and the random consistency index ($RI$), i.e., the mean of the $CI$ for a set of randomly generated matrices of dimension $n$,

$$CR_M = \frac{CI_M}{RI}$$

Saaty (1980) established that for the judgments of a decision-maker to be acceptable $CR_M$ should be less than 0.1.

### 3.7. Error and inconsistency

Rationality of the decision-makers in both Ordering and AHP implies consistency, however, the fact that the decision support technique gives a consistent solution does not imply that it is free of error (Liang, Brunelli, & Rezaei, 2020; Sáenz-Royo et al., 2023a; Sugden, 1985; Temesi, 2011). Consistency measures the precision of the decision-maker's judgment (coherence of the judgments) but does not measure the accuracy (proximity to the optimal solution). Two types of consistency are differentiated: ordinal consistency that manifests transitivity in the ordinal paired judgments manifested by the decision-maker, and cardinal consistency that manifests cardinal coherence in the intensities manifested by the decision-maker. Both cardinal and ordinal consistency can occur at a high level and yet present a wrong ranking, "this happens when the starting hypotheses are false but the logical structures of the relationship between the judgments are correct" (Sáenz-Royo et al., 2023a).

Cardinal consistency is more demanding than ordinal consistency since it requires that the intensity relationships of the paired comparisons be related, regardless of whether the decision-maker chooses an appropriate ranking (Rezaei, 2015). The usual measures of inconsistency try to measure a priori, from the judgments made, the level of rationality of the decision-maker (Grzybowski & Starczewski, 2020).

Transitivity has been linked to consistency as a fundamental tool to avoid misleading solutions. Consistency of the judgments in the Ordering must satisfy the following transitivity property: if $\widehat{r}_{ij} = 1$ and $\widehat{r}_{jk} = 1$ then $\widehat{r}_{ik} = 1$. Consistency of the judgments in the AHP must satisfy the following transitivity property: $\widehat{a}_{ik} = \widehat{a}_{ij}\widehat{a}_{jk}$. Consistency requires that all the judgments show a unique ranking of the values of the alternatives, and proportionality of the direct and indirect judgments in the cardinality; this property has been central in the development of the literature of pairs comparison, becoming the main indicator of the goodness of the consensus and the coherence of each expert (Herrera-Viedma et al., 2004; Saaty, 1980; Triantaphyllou, 2000). When the number of experts is large, AHP becomes impractical as it can lead to a high degree of inconsistency.

If the decision-maker is the only one who can know the values assigned to the alternatives, there could be invisible errors because it is not possible to contrast these values with reality, and the inconsistency only shows the existence of errors when the judgments show the incoherence of the decision-maker (lack of precision when comparing elements separately). However, the internal consistency of the decision-maker may be hiding systematic errors that cannot be detected by either cardinal or ordinal consistency. Therefore, the fact that a matrix $R$ or/and a matrix $M$ are consistent does not ensure that said matrix is errors free. The IBRM makes it possible to assess a priori the error percentages of the ordinal and cardinal inconsistency of the automaton's judgments.

## 4. The ∊-type and δ-type problems

The IBRM studies all the possible paired judgments that an automaton can manifest by assigning a probability to each one of them and obtaining the expected performance of the vector of priorities provided by a decision support technique. Sáenz-Royo et al. (2023a) raises two types of problems that we will call δ-type and ∊-type:

- δ-type problems are those in which the whole ranking provided by the priority vector is relevant when the alternatives studied are not mutually exclusive, but at least one alternative is to be excluded; in this case, the expected performance of the complete priority vector must be evaluated;
- ∊-type problems are those in which only the alternative considered the best is relevant because the alternatives are mutually exclusive and only one can be chosen; in these cases the errors in the ranking of the alternatives not chosen are irrelevant and the expected performance of the best alternative must be evaluated.

According to the IBRM, the following steps are carried out to evaluate a decision support technique:

1) All the possible paired judgments that the automaton can manifest are listed (in our proposal in the Ordering and in the AHP) and the probability of each of these judgments occurring is calculated.
2) All possible combinations of preference or intensity are obtained from comparisons between alternatives (set of judgments shown by one decision-maker).
3) For each combination of preference or intensity, the priority vector of the decision support technique is obtained. Each of these priority vectors has an assigned probability obtained from the product of the probabilities of the paired (preference or intensity) judgments that constitute it.
4) In δ-type problems, the priority vectors that propose the same ranking of alternatives are grouped by adding their probabilities; in ∊-type problems all priority vectors that have the same first alternative are grouped by adding their probabilities. At the end of this process, a probability is available for each priority vector in the δ-problems and a probability for each alternative in the ∊-type problems.
5) Multiplying the probabilities by the performance of the selected alternatives and adding the products, the expected performance of the decision support technique is obtained. In δ-type problems, the probability of each vector will be multiplied by the sum of the performances of the alternatives not discarded, while in ∊-type problems each probability is simply multiplied by the performance of the selected alternative.

In the ∊-type problem with three alternatives, the expected performance is defined as the probability of choosing alternative $A_1$ ($p_1$) multiplied by its latent performance $V_1$, plus the probability of choosing alternative $A_2$ ($p_2$) multiplied by its latent performance $V_2$, plus the probability that alternative $A_3$ ($p_3$) multiplied by its latent performance $V_3$, plus the mean performance multiplied by the probability that indecision is shown between any of the three alternatives ($p_{1=2=3}$).

$$E(\in - \textit{Ordering}\ \text{without}\ \textit{consistency}) = p_1V_1 + p_2V_2 + p_3V_3 + p_{1=2=3}\left(\frac{V_1+V_2+V_3}{3}\right)$$

Here $p_1$ is the sum of all the probabilities of the rankings that have alternative $A_1$ as the best alternative ($p_1 = p(O_{123}) + p(O_{132})$), and $p_2$ and $p_3$ are calculated similarly.

The probabilities $p_i$ will be different in the Ordering technique than in the AHP.

In the δ-type problem with three alternatives, we are interested in the total ranking of alternatives provided by each technique. In this type of problem, it is only possible to discard the alternative ranked last. If the three alternatives were eligible, no analysis would be necessary and if two alternatives had to be ruled out, we would be faced with a choice of the best (∊-type problem).

The expected performance in the δ-type problem with three alternatives is computed as the performance when alternative $A_3$ is discarded ($V_1 + V_2$) multiplied by the sum of the probabilities of the rankings in which alternative $A_3$ is in last position ($p(O_{123}) + p(O_{213})$), plus the performance obtained when alternative $A_2$ is discarded ($V_1 + V_3$) multiplied by the sum of the probabilities of the rankings in which alternative $A_2$ is in last position ($p(O_{132}) + p(O_{312})$), plus the performance obtained when alternative $A_1$ is discarded ($V_2 + V_3$) multiplied by the sum of the probabilities of the rankings in which alternative $A_1$ is in last position ($p(O_{231}) + p(O_{321})$), plus two times (it is chosen two alternatives) the mean performance multiplied by the probability of the technique showing indecision between any of the three alternatives ($p_{1=2=3}$).

$$E(\delta - \textit{Ordering}\ \text{without}\ \textit{consistency}) = (p(O_{123}) + p(O_{213}))(V_1+V_2) + (p(O_{132}) + p(O_{312}))(V_1+V_3) + (p(O_{231}) + p(O_{321}))(V_2+V_3) + p_{1=2=3}\frac{2(V_1+V_2+V_3)}{3}$$

The expected performance of the δ-type problem for the general case of $n$ alternatives will depend on the number of alternatives that are not eligible.

## 5. Ordering vs AHP. The value of intensity through an illustrative case

IBRM is used to obtain the expected performances of the "total priority vector" (δ-type) and the "choice of the first alternative" (∊-type) by comparing the two decision support techniques: Ordering vs AHP. This comparison is made in both cases by establishing a requirement for consistency and without requiring it. The difference between these two techniques is that the AHP requires the decision-maker to manifest intensities, and this has consequences both in the consistency restrictions and in obtaining the vector of priorities, therefore, the difference in performance will indicate the relative value of said information. To illustrate this, the case proposed by Sáenz-Royo et al. (2023a) is used: "An automaton with limited rationality and a reliability marked by (1) with $\beta = 1$ (by simplicity), and three alternatives whose latent performances are $V_1 = 62.5$, $V_2 = 25$ and $V_3 = 10$.".

### 5.1. Ordering

According to the Ordering-based paired comparison decision support technique, a probability matrix $P$, with the alternatives ranked from highest to lowest performance, that indicates the probability with which the automaton (decision-maker with intentional bounded rationality) will choose each alternative in each pairwise comparison is constructed. Thus, the probability of correctly preferring alternative $A_1$ over $A_2$ in the comparison of $V_1$ and $V_2$ will be $p_{12} = \frac{1}{e^{\left(\frac{V_j}{V_i}-\frac{V_i}{V_j}\right)}+1} = 0.8909$, therefore, the probability of incorrectly choosing the alternative $A_2$ over $A_1$ will be $p_{21} = 1 - p_{12} = 0.1091$. Similarly, the rest of probability values are computed, resulting in

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} - & 0.8909 & 0.9977 \\ 0.1091 & - & 0.8909 \\ 0.0023 & 0.1091 & - \end{pmatrix}$$

From this matrix it is possible to obtain the probabilities of the automaton expresses a concrete preference judgment.

- The only correct priority vector is the one that shows that $\hat{r}_1 > \hat{r}_2 > \hat{r}_3$, this vector is generated when the combination of judgments of the decision-maker is alternative $A_1$ is preferred to alternative $A_2$ ($\hat{r}_{12} = 1$), alternative $A_1$ is preferred to alternative $A_3$ ($\hat{r}_{13} = 1$) and alternative $A_2$ is preferred to alternative $A_3$ ($\hat{r}_{23} = 1$); this ranking is called *Error Free* and is denoted as the ranking ($O_{123}$) (Sáenz-Royo et al., 2023a).
- When the components of the priority vector satisfy that $\hat{r}_1 > \hat{r}_3 > \hat{r}_2$, this technique provides an erroneous ranking that chooses the best alternative correctly ($A_1$) and the combination of judgments of the decision-maker is $\hat{r}_{12} = 1$; $\hat{r}_{13} = 1$; $\hat{r}_{23} = 0$; this ranking is called *Right-Soft* and is denoted as $O_{132}$.
- Priority vector in which $\hat{r}_2 > \hat{r}_1 > \hat{r}_3$ (*Medium-Soft Error*) $O_{213}$: this technique chooses $A_2$ as the best alternative, and the decision-maker is wrong only in a pairwise judgment, the combination is $\hat{r}_{12} = 0$; $\hat{r}_{13} = 1$; $\hat{r}_{23} = 1$.
- Priority vector in which $\hat{r}_2 > \hat{r}_3 > \hat{r}_1$ (*Medium-Hard Error*) $O_{231}$: this technique chooses $A_2$ as the best alternative and the decision-maker is wrong in two judgments; she/he is always wrong when she/he is involved in judgment $A_1$ (he undervalues it), $\hat{r}_{12} = 0$; $\hat{r}_{13} = 0$; $\hat{r}_{23} = 1$.
- Priority vector in which $\hat{r}_3 > \hat{r}_1 > \hat{r}_2$ (*Extreme-Soft Error*) $O_{312}$: the technique opts for $A_3$ as the best alternative. The decision-maker is wrong in all pairwise judgments involving $A_3$ (she/he overvalues it), $\hat{r}_{12} = 1$; $\hat{r}_{13} = 0$; $\hat{r}_{23} = 0$.
- Priority vector in which $\hat{r}_3 > \hat{r}_2 > \hat{r}_1$ (*Extreme-Hard Error*) $O_{321}$: the technique opts for $A_3$ as the best alternative and the decision-maker shows error in all her/his judgments by pairs, $\hat{r}_{12} = 0$; $\hat{r}_{13} = 0$; $\hat{r}_{23} = 0$.
- Priority vector in which $\hat{r}_1 = \hat{r}_2 = \hat{r}_3$ (*Total error*) $O_{1=2=3}$: The technique shows the three alternatives as equivalent. The decision-maker shows two possible combinations of judgments: $\hat{r}_{12} = 1$; $\hat{r}_{13} = 0$; $\hat{r}_{23} = 1$ or $\hat{r}_{12} = 0$; $\hat{r}_{13} = 1$; $\hat{r}_{23} = 0$, manifesting a circular preference system.

A decision-maker can provide the below eight possible comparison matrices. Each matrix is marked with superscripts to indicate the type of error they generate. Each comparison matrix will have a probability of occurrence according to the level of reliability of the automaton. With the alternatives ranked by performance, the different possible matrices are:

$$R^{EF} = \begin{pmatrix} - & 1 & 1 \\ 0 & - & 1 \\ 0 & 0 & - \end{pmatrix} \qquad R^{RS} = \begin{pmatrix} - & 1 & 1 \\ 0 & - & 0 \\ 0 & 1 & - \end{pmatrix}$$

$$R^{MSE} = \begin{pmatrix} - & 0 & 1 \\ 1 & - & 1 \\ 0 & 0 & - \end{pmatrix} \qquad R^{MHE} = \begin{pmatrix} - & 0 & 0 \\ 1 & - & 1 \\ 1 & 0 & - \end{pmatrix}$$

$$R^{ESE} = \begin{pmatrix} - & 1 & 0 \\ 0 & - & 0 \\ 1 & 1 & - \end{pmatrix} \qquad R^{EHE} = \begin{pmatrix} - & 0 & 0 \\ 1 & - & 0 \\ 1 & 1 & - \end{pmatrix}$$

$$R^{TE1} = \begin{pmatrix} - & 1 & 0 \\ 0 & - & 1 \\ 1 & 0 & - \end{pmatrix} \qquad R^{TE2} = \begin{pmatrix} - & 0 & 1 \\ 1 & - & 0 \\ 0 & 1 & - \end{pmatrix}$$

Each matrix, and therefore each priority vector, is assigned a probability, which is denoted by $p(R) = p_{ij}p_{jk}p_{ik}$. The probabilities of each ranking obtained for the illustrated case are:

$$p(R^{EF}) = p(O_{123}) = p_{12}p_{23}p_{13} = 0.7919 \qquad p(R^{RS}) = p(O_{132}) = 0.0970$$
$$p(R^{MSE}) = p(O_{213}) = 0.0970 \qquad p(R^{MHE}) = p(O_{231}) = 0.0002$$
$$p(R^{ESE}) = p(O_{312}) = 0.0002 \qquad p(R^{EHE}) = p(O_{321}) = 0.0000$$
$$p(R^{TE1}) = 0.0018 \qquad p(R^{TE2}) = 0.0119$$
$$p(R^{TE}) = p(O_{1=2=3}) = p(R^{TE1}) + p(R^{TE2}) = 0.0137$$

In the case of three alternatives, only two matrices have the same ranking Total Error, and therefore their probabilities must be added, $p(O_{1=2=3}) = 0.0137$.

In the $\delta$-type problem, the entire priority vector is important, and the probability of each ranking is obtained by adding the probability of all the priority vectors that assume the same ranking. The expected performance is

$$\begin{aligned} E(\delta - Ordering \text{ without } consistency) &= (p(O_{123}) + p(O_{213}))(V_1 + V_2) \\ &+ (p(O_{132}) + p(O_{312}))(V_1 + V_3) \\ &+ (p(O_{231}) + p(O_{321}))(V_2 + V_3) \\ &+ p(O_{1=2=3})\frac{2(V_1 + V_2 + V_3)}{3} \\ &= 85.7216 \end{aligned}$$

In this case, the expected performance of a decision-maker who does not make mistakes is $V_1 + V_2 = 62.5 + 25 = 87.5$ and, therefore, the technique achieves 97.97% of the best possible performance.

In the $\epsilon$-type problem, a second grouping must be performed once the probabilities of each ranking have been obtained, adding the probabilities of all the rankings that choose the same alternative as the best option, $p_i = \sum_{j\neq i}^{n}\sum_{k\neq i\neq j}^{n} p(O_{ijk})$. The number of rankings that select each alternative is $n!/n$, in the example they are:

$$p_1 = p(O_{123}) + p(O_{132}) = 0.8889; \; p_2 = 0.0972; \; p_3 = 0.0002; \; p_{1=2=3} = 0.0137$$

For the ranking, the probability that the priority vector shows $A_1$ as the best alternative is 0.8889, that it chooses $A_2$ as the best alternative is 0.0972 and that it chooses $A_3$ as the best alternative is 0.0002, while the probability of the technique shows that the three alternatives are equivalent is 0.0137. Once the probabilities are obtained, the expected performance of the $\epsilon$-problem is calculated:

$$\begin{aligned} E(\epsilon - Ordering \text{ without } consistency) &= p_1V_1 + p_2V_2 + p_3V_3 \\ &+ p_{1=2=3}\left(\frac{V_1 + V_2 + V_3}{3}\right) \\ &= 0.8889 \cdot 62.5 + 0.0972 \cdot 25 \\ &+ 0.0002 \cdot 10 \\ &+ 0.0137 \cdot \left(\frac{65.5 + 25 + 10}{3}\right) \\ &= 58.4322 \end{aligned}$$

The expected performance of a decision-maker who does not commit errors is $V_2 = 62.5$, and, therefore, the technique obtains 93.49% of the best possible performance.

Ordering can add the requirement to show ordinal consistency in judgments combination to validate the priority vector. It is established as inconsistency when the judgments by pairs shown by the decision-maker are not coherent in the sense that they do not meet the required transitivity. For the case of three alternatives, the pairwise judgment combinations that give rise to inconsistency are $\widehat{r}_{12} = 1$; $\widehat{r}_{13} = 0$; $\widehat{r}_{23} = 1$ and $\widehat{r}_{12} = 0$; $\widehat{r}_{13} = 1$; $\widehat{r}_{23} = 0$, that match matrices $R^{TE1}$ and $R^{TE2}$. The Ordering with consistency requirement rejects 2 of the 8 possible matrices due to inconsistency. Therefore, the priority vectors that show the three alternatives as equivalent are considered unacceptable since they come from a judgments combination that does not present the required levels of coherence.

When the alternatives are performance ranked, in the judgment matrices acceptable by the decision support technique it is observed that as the error increases, the matrices show a greater number of ones in the lower part of the matrix. If the performance of each alternative is unknown, any of the error matrices is indistinguishable from the Error-Free matrix since orderings with a manifestation of consistent preferences have traditionally been classified as correct. Only the IBRM allows evaluating the decision support technique a priori.

The probability that the decision-maker shows a consistent judgment matrix is $0.9863(= 1 - 0.0137)$, and therefore, after removing inconsistencies 80.29% of the supported priority vectors are Error-Free, without this probability the errors are undetectable. In $\delta$-type problems when consistency is required in the decisions, and 1.37% of the judgments issued by the automaton are eliminated, the expected performance is:

$$\begin{aligned} E(\delta - \textit{Ordering} \text{ with } \textit{consistency}) &= (p(O_{123}) + p(O_{213}))(V_1 + V_2) \\ &+ (p(O_{132}) + p(O_{312}))(V_1 + V_3) \\ &+ (p(O_{231}) + p(O_{321}))(V_2 + V_3) \\ &= 86.3002 \end{aligned}$$

This supposes a gain over not requiring consistency of 0.67% and a percentage of 98.63% on the maximum possible performance.

Once the inconsistencies have been eliminated, 90.12% of the time this technique accepts a ranking that puts alternative $A_1$ first, 9.85% choose alternative $A_2$ as the best, and 0.03% choose alternative $A_3$. The expected performance of the $\in$-type problems requiring consistency is

$$E(\in - \textit{Ordering} \text{ with } \textit{consistency}) = p_1V_1 + p_2V_2 + p_3V_3 = 58.7916$$

which represents a 0.62% gain over not requiring consistency, obtaining 94.07% of the best possible performance.

### 5.2. AHP

According to the decision support technique based on the intensity of the pairwise comparison, the AHP, as many matrices can be constructed as there are combinations of intensities the decision-maker can manifest.

In our case, three alternatives with four possible intensities, this makes a total of 64 ($4^3$) matrices, we call them $M(1), \cdots, M(64)$, while in Ordering there were 8 ($2^3$) (detailed in Section 5.1 depending on the type of error they represent: $R^{EF}$, $R^{RS}$, $\cdots$). Each of these matrices is associated with a probability of occurrence provided by IBR. Thus, grouping all those combinations of intensities (matrices) when applying the AHP produce the same priority vector $A_1 \succ A_2 \succ A_3$, that is, $w_1 > w_2 > w_3$, it is $p(O_{123}) = 0.8019$. Some examples of these matrices are:

$$M(1) = \begin{pmatrix} - & \frac{25}{4} & \frac{25}{4} \\ \frac{4}{25} & - & \frac{25}{4} \\ \frac{4}{25} & \frac{4}{25} & - \end{pmatrix}$$; priority vector associated by AHP: $\widehat{w}(M(1)) = (\widehat{w}_1, \widehat{w}_2, \widehat{w}_3) = (0.6880, 0.2397, 0.0723)$. Its consistency ratio is $CR(M(1)) = 0.3539$. The probability that the automaton with expertise $\beta = 1$ shows a preference matrix equal to $M(1)$ according to IBR is $p(M(1)) = p_{12}(25/4)^*p_{13}(25/4)^*p_{23}(25/4) = 0.0089$.

$$M(2) = \begin{pmatrix} - & \frac{25}{4} & \frac{25}{4} \\ \frac{4}{25} & - & \frac{10}{4} \\ \frac{4}{25} & \frac{4}{10} & - \end{pmatrix} \widehat{w}(M(2)) = (0.7385, 0.1694, 0.0920)$$

$$CR(M(2)) = 0.0829$$

$$p(M(2)) = p_{12}(25/4)^*p_{13}(25/4)^*p_{23}(10/4) = 0.0690$$

Notice that the preference matrix $M(1)$ would be rejected with the consistency criterion while $M(2)$ would not.

…

Grouping the probabilities of all those priority vectors that provide the same ranking, the following values are obtained:

$$p(O_{123}) = 0.8019;\ p(O_{132}) = 0.0973;$$
$$p(O_{213}) = 0.0973;\ p(O_{231}) = 0.0002;$$

$$p(O_{312}) = 0.0002;\ p(O_{321}) = 0.0000;$$
$$p(O_{1=2=3}) = 0.0030$$

As can be seen, all the rankings but one (Total Error) increase their probability with respect to the Ordering technique. Although the probability of obtaining an Error-Free ranking is higher than that provided by the Ordering technique, so are the probabilities of making errors, except for Total Error. This does not ensure that the results of the AHP are always superior to those of the Ordering, despite being a more informative technique.

In $\delta$-type problems, the expected performance of AHP is

$$\begin{aligned} E(\delta - \textit{AHP} \text{ without } \textit{consistency}) &= (p(O_{123}) + p(O_{213}))(V_1 + V_2) \\ &+ (p(O_{132}) + p(O_{312}))(V_1 + V_3) \\ &+ (p(O_{231}) + p(O_{321}))(V_2 + V_3) \\ &+ p(O_{1=2=3})\frac{2(V_1 + V_2 + V_3)}{3} \\ &= 85.9540 \end{aligned}$$

The technique achieves 98.23% of the maximum possible performance and, therefore, its contribution regarding the Ordering without consistency is 0.27%.

As with the Ordering, the expected performance of the $\in$-type problem is obtained. All the rankings are grouped according to the alternative with the highest valuation ($p_1 = 0.8992$; $p_2 = 0.0976$; $p_3 = 0.0003$; $p_{1=2=3} = 0.0030$) and subsequently, the expected performance of the AHP technique is obtained without consistency requirement:

$$\begin{aligned} E(\in - \textit{AHP} \text{ without } \textit{consistency}) &= p_1V_1 + p_2V_2 + p_3V_3 \\ &+ p_{1=2=3}\left(\frac{V_1 + V_2 + V_3}{3}\right) \\ &= 58.7366 \end{aligned}$$

This expectation is higher (0.3044) than the obtained by the Ordering technique, that is, 0.52%. The technique presents 93.98% of the maximum possible performance.

The AHP technique establishes a criterion for a priority vector to be admitted: $CR < 0.1$. This is due to the requirement of coherence in the intensity ratios, that is, if the decision-maker's judgment states that $V_1$ is 2.5 times $V_2$ and states that $V_2$ is 2.5 times $V_3$, then $V_1$ must be 6.25 times greater than $V_3$. This criterion makes the intensity-based decision support technique (AHP) to reject more matrices with errors than those rejected by the Ordering technique, but it also rejects Error-Free matrices.

The probability and *CR* of each of the 64 possible priority vectors have been calculated. Grouping by ranking (with $CR < 0.1$) according to the type of error made by the automaton, the results are that the probability that the priority vector is rejected is 0.1949, much higher than the 0.0137 of the Ordering technique. Among the admitted priority vectors, the probability distribution is:

$$p(O_{123}) = 0.9433;\ p(O_{132}) = 0.0281;$$
$$p(O_{213}) = 0.0281;\ p(O_{231}) = 0.0002;$$

$$p(O_{312}) = 0.0002;\ p(O_{321}) = 0.0000$$

All priority vectors that show equality between the alternatives are rejected by the consistency ratio. The introduction of the consistency requirement restriction increases the probability of Error-Free ranking and reduces the probability of errors at the cost of rejecting close to 20% of the judgments issued by the automaton.

In $\delta$-type problems, the expected performance of the AHP requiring consistency is

$$\begin{aligned} E(\delta - AHP \text{ with } consistency) &= (p(O_{123}) + p(O_{213}))(V_1 + V_2) + (p(O_{132}) \\ &\quad + p(O_{312}))(V_1 + V_3) + (p(O_{231}) \\ &\quad + p(O_{321}))(V_2 + V_3) \\ &= 87.0605 \end{aligned}$$

The technique achieves 99.50% of the maximum possible performance and, therefore, its contribution with respect to the Ordering with consistency is 0.88%.

To study the problem of ∊-type, the rankings are grouped to obtain the probability with which each alternative is chosen. Among the vectors established as acceptable by AHP, the probability that this decision support technique chooses the alternative $A_1$ is 0.9714, that it chooses $A_2$ is 0.0284 and that it chooses $A_3$ is 0.0003. These probabilities are better than those obtained by the Ordering technique, and they also suppose an increase in the probability of choosing alternative $A_1$, with respect to the same technique without the requirement of consistency.

The expected performance of AHP demanding consistency is

$$E(\epsilon - AHP \text{ with } consistency) = p_1V_1 + p_2V_2 + p_3V_3 = 61.4225$$

which represents 4.57% gain over not requiring consistency, 98.28% of the maximum performance and a gain over the Ordering with consistency of 4.48%.

In summary, it can be seen that for the used case study, the differences between methods are meager, the most relevant is the difference in performance of 5.12% in the ∊-type problems between AHP with consistency and Ordering without consistency, for which it is was necessary to request intensity judgments from the experts and eliminate 20% of the judgments due to cardinal inconsistency.

## 6. Robustness analysis of the ∊-type problem

Choo and Wedley (2004) and Lin (2007) linked the efficiency of the decision support technique to the characteristics of the paired comparison matrix. However, the IBRM forces to study all possible paired comparison matrices, transferring the problem to the characteristics of the automaton (decision-maker expertise) and the difficulty of the problem (differences between performances) as was demanded by Tsoukiàs (2008). Studying different levels of expertise (characteristics of the decision-maker) in different scenarios (difficulties) is a solid way of evaluating decision support techniques. To complete the case analysis developed in the previous section, we proceed to obtain the expected performance of the ∊-type problem for different levels of expertise and for different performance distributions of the alternatives. IRBM makes relative comparisons, therefore it is perfectly valid for any reference system that is established, and its relative results do not change. In fact, the change of value of $V_2$ modifies the probability of judgment favorable or unfavorable to alternative $A_2$ in all its comparisons, which allows to analyze the corresponding variation in the distribution of error probability and the distributions of the expected performances.

Three scenarios are recreated establishing that alternative $A_2$ has a performance: close to alternative $A_3$ ($V_2 = 15$); central between alternative $A_1$ and $A_3$ ($V_2 = 25$, as in the case study); and close to alternative $A_1$ ($V_2 = 55$). For each of these scenarios, levels of expertise of the automaton ranging from 0.05 to 1.65 were simulated. These levels were chosen since, based on them, the performances of the different decision support techniques compared converged. The techniques can hardly help in the decision when the decision-maker's capacity to judge is very high (high expertise), she/he reaches the best solution with any technique and when her/his judgment capacity is very low and she/he values all the alternatives equally.

In all scenarios, the technique that shows the worst results is Ordering without consistency. However, the technique that shows the best-expected performance at low levels of expertise of the automaton (less than 0.25) is Ordering with consistency, while for levels of expertise of the automaton greater than or equal to 0.25 the best technique is AHP with consistency. This result is maintained in the three scenarios studied. This is due to the fact that when the automaton shows a low level of expertise (0.15), although the AHP technique rejects a higher proportion of judgments than Ordering (the values in the scenario with $V_2 = 25$ are 70.28% in the AHP compared to 22.27% of the Ordering), the probability that the AHP achieves for judgments Error-Free (0.3154) plus Right-Soft Error (0.2138) is 0.5291, lower than that obtained by the Ordering (EF: 0.3069 and RS: 0.2240) 0.5308. In AHP, 70.28% of the total of possible priority vectors are rejected due to inconsistency, of which 17.63% are Error-Free judgments, 14.07% are Right-Soft Error, 14.07% are Medium-Soft Error, 5.91% are Medium-Hard Error, 5.91% are Extreme-Soft Error, 4.67% are Extreme-Hard Error, and the rest of rejections (8.02%) are the priority vectors that presented equality in the evaluation of the alternatives. Therefore, in this case, the introduction of the consistency requirement in AHP provides lower gains than in Ordering due to the high percentage of Right-Soft Error rejected due to inconsistency. This result shows that in certain circumstances the requirement of the decision-maker to show intensity does not outperform a technique based on Ordering.

The change in scenario (change in the performance of alternative $A_2$) implies important changes in the range of expected performance of all the decision support techniques, despite the fact that the maximum achievable value remains unchanged at 62.5 being an ∊-type problem. The highest performance range occurs when $V_2 = 15$ (see Table 3). In this scenario, the automaton has difficulties distinguishing between alternative $A_1$ and alternative $A_2$, but it is easier to recognize the best alternative than in the other scenarios (see Fig. 1). The fact of having more alternatives with low performances means that with low levels of expertise of the automaton the expected performance is lower than in the rest of the scenarios, while when the expertise of the automaton is high it is easier to recognize the best alternative than in the rest of the scenarios; so its expected performance is higher in all techniques. The lowest range occurs when $V_2 = 55$ for two reasons: when the expertise of the automaton is low, the expected performance of the automaton is high because there are more alternatives with high performances; and when the expertise of the automaton is high, the expected performances of the techniques are relatively low because it is difficult to distinguish whether alternative $A_1$ or $A_2$ is the best (see Fig. 3).

The greatest differences between the expected performances of the decision support techniques occur in scenario $V_2 = 25$ where AHP with consistency presents higher expected performances than the rest of the techniques (see Fig. 2). The greatest difference between techniques (from the one that gives the highest performance to the one that gives the least) in this scenario ranges from 1.7% to 8.30%. The scenario

**Table 3**
Expected performance of decision support techniques in the tree scenarios for different values of $\beta$.

| $\beta$ | $V_2$ | | 15 | | |
|---|---|---|---|---|---|
| | ϵ-AHP without C. | ϵ-Ordering without C. | ϵ-AHP with C. | ϵ-Ordering without C. | (Max-Min)/Min |
| 0.05 | 32.6019 | 32.4530 | 32.8089 | 33.5254 | 3.30% |
| 0.15 | 40.0952 | 39.2093 | 41.6644 | 41.9959 | 7.11% |
| 0.25 | 46.8347 | 45.4028 | 49.5239 | 48.7793 | 9.08% |
| 0.35 | 51.9653 | 50.4408 | 54.8736 | 53.4977 | 8.79% |
| 0.45 | 55.5352 | 54.2117 | 58.1507 | 56.5839 | 7.27% |
| 0.55 | 57.9137 | 56.8894 | 60.0609 | 58.5755 | 5.57% |
| 0.65 | 59.4723 | 58.7323 | 61.1434 | 59.8706 | 4.11% |
| 0.75 | 60.4910 | 59.9790 | 61.7480 | 60.7238 | 2.95% |
| 0.85 | 61.1595 | 60.8147 | 62.0834 | 61.2930 | 2.09% |
| 0.95 | 61.6008 | 61.3729 | 62.2691 | 61.6765 | 1.46% |
| 1.05 | 61.8941 | 61.7453 | 62.3719 | 61.9366 | 1.01% |
| 1.15 | 62.0902 | 61.9939 | 62.4288 | 62.1140 | 0.70% |
| 1.25 | 62.2220 | 62.1601 | 62.4604 | 62.2352 | 0.48% |
| 1.35 | 62.3110 | 62.2714 | 62.4779 | 62.3184 | 0.33% |
| 1.45 | 62.3713 | 62.3460 | 62.4877 | 62.3754 | 0.23% |
| 1.55 | 62.4123 | 62.3962 | 62.4931 | 62.4145 | 0.16% |
| 1.65 | 62.4402 | 62.4300 | 62.4962 | 62.4414 | 0.11% |

| $\beta$ | $V_2$ | | 25 | | |
|---|---|---|---|---|---|
| | ϵ-AHP without C. | ϵ-Ordering without C. | ϵ-AHP with C. | ϵ-Ordering without C. | (Max-Min)/Min |
| 0.05 | 35.3164 | 35.1914 | 35.5089 | 36.0730 | 2.50% |
| 0.15 | 41.0783 | 40.3312 | 42.4118 | 42.5751 | 5.56% |
| 0.25 | 45.9726 | 44.7244 | 48.1936 | 47.4829 | 7.76% |
| 0.35 | 49.6220 | 48.2098 | 52.2093 | 50.8174 | 8.30% |
| 0.45 | 52.2278 | 50.9001 | 55.0502 | 53.0594 | 8.15% |
| 0.55 | 54.1141 | 52.9899 | 57.1378 | 54.6527 | 7.83% |
| 0.65 | 55.5405 | 54.6484 | 58.6811 | 55.8755 | 7.38% |
| 0.75 | 56.6724 | 55.9941 | 59.8099 | 56.8756 | 6.81% |
| 0.85 | 57.6045 | 57.1042 | 60.6233 | 57.7259 | 6.16% |
| 0.95 | 58.3894 | 58.0286 | 61.2014 | 58.4612 | 5.47% |
| 1.05 | 59.0572 | 58.8013 | 61.6073 | 59.0992 | 4.77% |
| 1.15 | 59.6264 | 59.4474 | 61.8896 | 59.6508 | 4.11% |
| 1.25 | 60.1104 | 59.9866 | 62.0844 | 60.1245 | 3.50% |
| 1.35 | 60.5201 | 60.4353 | 62.2180 | 60.5282 | 2.95% |
| 1.45 | 60.8651 | 60.8074 | 62.3092 | 60.8697 | 2.47% |
| 1.55 | 61.1540 | 61.1150 | 62.3712 | 61.1566 | 2.06% |
| 1.65 | 61.3947 | 61.3684 | 62.4131 | 61.3962 | 1.70% |

| $\beta$ | $V_2$ | | 55 | | |
|---|---|---|---|---|---|
| | ϵ-AHP without C. | ϵ-Ordering without C. | ϵ-AHP with C. | ϵ-Ordering without C. | (Max-Min)/Min |
| 0.05 | 45.9835 | 45.8219 | 46.2384 | 46.8998 | 2.35% |
| 0.15 | 51.9772 | 51.1117 | 53.0914 | 53.3909 | 4.46% |
| 0.25 | 55.7069 | 54.5381 | 56.7719 | 56.7667 | 4.10% |
| 0.35 | 57.5570 | 56.5428 | 58.2545 | 58.1973 | 3.03% |
| 0.45 | 58.3909 | 57.6670 | 58.8498 | 58.7461 | 2.05% |
| 0.55 | 58.7698 | 58.2989 | 59.1117 | 58.9588 | 1.39% |
| 0.65 | 58.9575 | 58.6650 | 59.2476 | 59.0555 | 0.99% |
| 0.75 | 59.0652 | 58.8879 | 59.3349 | 59.1150 | 0.76% |
| 0.85 | 59.1388 | 59.0328 | 59.4015 | 59.1636 | 0.62% |
| 0.95 | 59.1973 | 59.1345 | 59.4576 | 59.2096 | 0.55% |
| 1.05 | 59.2493 | 59.2123 | 59.5069 | 59.2552 | 0.50% |
| 1.15 | 59.2983 | 59.2766 | 59.5515 | 59.3011 | 0.46% |
| 1.25 | 59.3459 | 59.3332 | 59.5924 | 59.3472 | 0.44% |
| 1.35 | 59.3928 | 59.3854 | 59.6304 | 59.3934 | 0.41% |
| 1.45 | 59.4393 | 59.4351 | 59.6664 | 59.4396 | 0.39% |
| 1.55 | 59.4856 | 59.4831 | 59.7010 | 59.4857 | 0.37% |
| 1.65 | 59.5317 | 59.5302 | 59.7345 | 59.5317 | 0.34% |

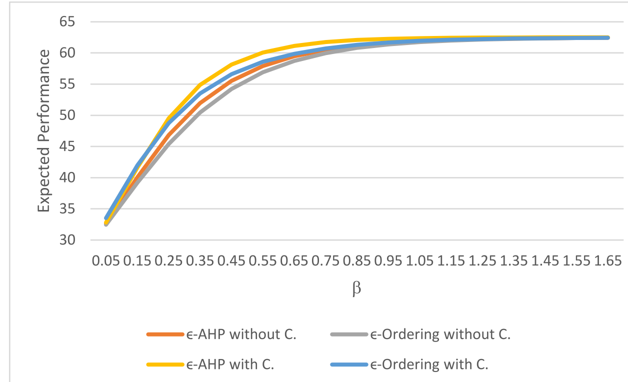


**Fig. 1.** Expected performance of decision support techniques in the scenario where $V_1 = 62.5; V_2 = 15; V_3 = 10$ for different values of $\beta$.

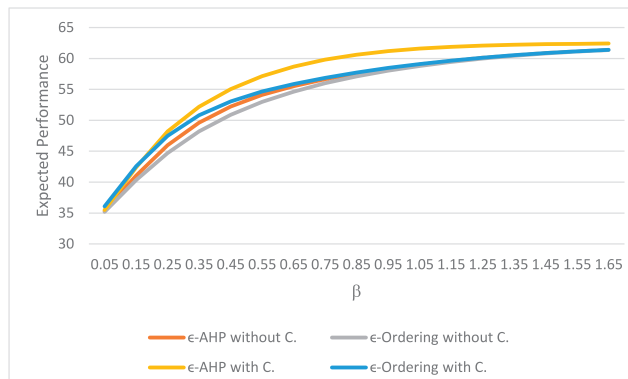


**Fig. 2.** Expected performance of decision support techniques in the scenario where $V_1 = 62.5; V_2 = 25; V_3 = 10$ for different values of $\beta$.

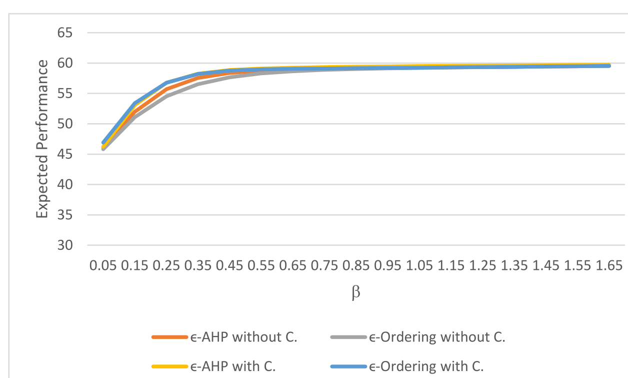


**Fig. 3.** Expected performance of decision support techniques in the scenario where $V_1 = 62.5; V_2 = 55; V_3 = 10$ for different values of $\beta$.

where the smallest differences occur is $V_2 = 55$, these are between 0.34% and 4.46%. In the three scenarios, the greatest differences in performance between techniques occur for expertise levels of the automaton of 0.25 –0.35. In the $V_2 = 55$ scenario, the differences are highly concentrated in low expertise of the automaton (differences greater than 2% are in the expertise interval [0.05 –0.45]), that is, in this scenario, the cumulative performance of the Ordering with consistency is superior to AHP with consistency up to the expertise of the automaton of 0.75. While in the scenario $V_2 = 25$ the interval of expertise of the automaton with differences greater than 2% is [0.05 –1.55] and in the scenario with $V_2 = 15$ this interval is [0.05 –0.85]. Finally, it should be noted that the maximum difference occurs between the AHP technique with consistency and Ordering without consistency in the $V_2 = 15$ scenario with 0.25 expertise while their lowest difference occurs in the $V_2 = 15$ scenario with 1.65 expertise of the automaton.

The differences in the performance of the decision support techniques depend on the difficulty of the problem as something exogenous and the expertise level of the decision-maker. To evaluate each problem, it must be considered that in order to obtain the maximum gains exposed, the decision support technique must consume resources in two ways: the first must require the decision-maker to show intensities, and the second refers to the rejection of a high number of combinations of judgments due to cardinal inconsistency.

This section suggests the possibility of analyzing resilience when other reliability measures (consensus level, group and individual consistency requirements, hesitation, etc.) are added, increasing the complexity of decision support techniques (see measures of resilience for algorithms Han, Liu and Zhang, 2016). IBRM can be used to develop a measure of resilience for decision support techniques, which seems to be a promising line of future research.

## 7. Conclusions

Human judgment is a valuable decision instrument in many fields. In this study, the problems have been classified according to the need to choose an alternative or obtain a complete ranking ($\epsilon$- and $\delta$-type problems). Decision support techniques try to analyze the logical relationships that are deduced from the judgments shown by the decision-maker to detect possible errors and reject the judgments that do not meet certain properties. It has been analyzed whether a decision support technique that is based on intensities such as AHP (systems that collect a higher level of information from judgments) always presents better results than a technique based on preference Ordering. To analyze this problem, the IBRM is used as a rationality framework where it is possible to evaluate any decision support technique a priori. In this conceptual framework, the decision-maker can be wrong or right. Section 3 details the Ordering technique and how the IBRM is applied to it, detailing possible errors, and relating them to the consistency criterion.

The results show how it is possible for a technique based on intensities (with more information) to present worse results than a simpler one based solely on preferences. This is because one of the most used properties to detect erroneous judgments is inconsistency. Sáenz-Royo et al. (2023a) showed the drawbacks that this measure has in AHP, fundamentally that a set of Error-Free judgments implies consistency in the judgments, but consistency in the judgments does not imply that they are Error-Free. This study has illustrated this concept for Ordering and by means of the IRBM, the probabilities of committing different types of errors in Ordering have been detailed to compare them with those committed in intensities. Finally, the relationship between the errors with the difficulty of the problem (distribution of the performance of the alternatives) and the expertise of the decision-maker has been revealed, detailing how these aspects affect the performance of the different decision support techniques.

The analysis shows that Ordering with consistency presents better results than the other techniques in the three scenarios when the levels of expertise of the descision-maker are low (less than 0.25). For expertise level greater than or equal to 0.25, AHP with consistency is better than the rest of the techniques. For very low and very high values of expertise, the performances of all the techniques converge and the simplicity of the technique to be used should be the determining factor. The AHP with consistency requires a higher level of resources than the rest of the techniques in two senses: on the one hand, it requires the provision of intensities in judgment preferences while, on the other hand, the requirement of cardinal consistency translates into a significant percentage of rejected judgments (may be higher than 70%, as shown section 6 for $\beta = 0.15$ in the scenario $V_2 = 25$,) because small errors of appreciation in the judgments can lead to rejection due to cardinal consistency, whether it comes from a correct or incorrect ranking (in the example mentioned in this paragraph, out of every 70 rejected judgments, 18 presented an Error-Free ranking and 14 had a Right-Soft Error). The Ordering technique without consistency is the one that shows the worst results in all scenarios and with all levels of expertise, but it is the one that consumes the least resources since it does not require the decision-maker to show intensities, nor does it eliminate any judgment and the performance differences with the AHP without consistency are meager.

The robustness Section shows how the introduction of consistency and intensity restriction presents a maximum gain of 9.08%. The maximum difference between Ordering with consistency and AHP with consistency is 5.16% in the $V_2 = 25$ scenario; in the $V_2 = 15$ scenario this difference is always less than 2.8%; while in the $V_2 = 55$ scenario, it is always less than 0.5%. Ultimately, the results show how incorporating intensity shows meager performances in most situations and does not guarantee a better result. A preference-based comparison system is faster and less demanding for experts than one based on intensity. This statement should open the debate on what criteria to use for selecting decision support techniques.

## Funding

This work was supported by the Spanish Ministerio de Economía y Competitividad, [ECO2017-86305-C4-3-R], the Diputación General de Aragón (DGA) and the European Social Fund [CREVALOR], the CUD (UZCUD2017-SOC-04) and by the Spanish State Research Agency under Project PID2019-103880RB-I00/AEI/https://doi.org/10.13039/501, 100,011,033 and PID2020-113338RB-I00.

## Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Carlos Saenz Royo reports administrative support was provided by Spanish Ministerio de Economía y Competitividad. Carlos Saenz Royo reports administrative support was provided by Diputación General de Aragon. Carlos Saenz Royo reports administrative support was provided by Defense University Centre Zaragoza. Enrique Herrera-Viedma reports administrative support and travel were provided by Spanish State Agency of Research. Carlos Saenz Royo reports administrative support was provided by Spanish State Agency of Research.

## Data availability

No data was used for the research described in the article.

## References

Belton, V. (1986). A comparison of the analytic hierarchy process and a simple multi-attribute value function. *European Journal of Operational Research, 26*(1), 7–21. https://doi.org/10.1016/0377-2217(86)90155-4

Belton, V., & Gear, T. (1983). On a short-coming of Saaty's method of analytic hierarchies. *Omega, 11*(3), 228–230. https://doi.org/10.1016/0305-0483(83)90047-6

Cai, M., Lin, Y., Han, B., Liu, C., & Zhang, W. (2017). On a simple and efficient approach to probability distribution function aggregation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 47*(9), 2444–2453. https://doi.org/10.1109/TSMC.2016.2531647

Chinchanachokchai, S., Thontirawong, P., & Chinchanachokchai, P. (2021). A tale of two recommender systems: The moderating role of consumer expertise on artificial intelligence based product recommendations. *Journal of Retailing and Consumer Services, 61*, Article 102528. https://doi.org/10.1016/j.jretconser.2021.102528

Choo, E. U., & Wedley, W. C. (2004). A common framework for deriving preference values from pairwise comparison matrices. *Computers & Operations Research, 31*(6), 893–908. https://doi.org/10.1016/S0305-0548(03)00042-X

Csaszar, F. A. (2013). An efficient frontier in organization design: Organizational structure as a determinant of exploration and exploitation. *Organization Science, 24*(4), 1083–1101. https://doi.org/10.1287/orsc.1120.0784

Díaz, R., Fernández, E., Figueira, J.-R., Navarro, J., & Solares, E. (2022). A new hierarchical multiple criteria ordered clustering approach as a complementary tool for sorting and ranking problems. *Omega*. , Article 102820. https://doi.org/10.1016/j.omega.2022.102820

Dyer, J. S. (1990). Remarks on the analytic hierarchy process. *Management Science, 36*(3), 249–258. https://doi.org/10.1287/mnsc.36.3.249

Fichtner, J. (1986). On deriving priority vectors from matrices of pairwise comparisons. *Socio-Economic Planning Sciences, 20*(6), 341–345. https://doi.org/10.1016/0038-0121(86)90045-5

Figueira, J. R., Greco, S., Ehrgott, M., & Henggeler Antunes, C. (Eds.). (2005). *Multiple criteria decision analysis: State of the art surveys*. New York, NY: Springer.

Garg, H. (2019). Generalized Intuitionistic Fuzzy Entropy-Based Approach for Solving Multi-attribute Decision-Making Problems with Unknown Attribute Weights. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, *89* (1), 129–139. 10.1007/s40010-017-0395-0.

Gass, S. I. (1998). Tournaments, transitivity and pairwise comparison matrices. *Journal of the Operational Research Society, 49*(6), 616–624. https://doi.org/10.1057/palgrave.jors.2600572

Grzybowski, A. Z., & Starczewski, T. (2020). New look at the inconsistency analysis in the pairwise-comparisons-based prioritization problems. *Expert Systems with Applications, 159*, Article 113549. https://doi.org/10.1016/j.eswa.2020.113549

Han, B., Liu, C. L., & Zhang, W. J. (2016). A method to measure the resilience of algorithm for operation management. *IFAC-PapersOnLine, 49*(12), 1442–1447.

Harary, F., & Moser, L. (1966). The Theory of Round Robin Tournaments. *The American Mathematical Monthly, 73*(3), 231–246. https://doi.org/10.1080/00029890.1966.11970749

Herrera-Viedma, E., Herrera, F., Chiclana, F., & Luque, M. (2004). Some issues on consistency of fuzzy preference relations. *European Journal of Operational Research, 154*(1), 98–109. https://doi.org/10.1016/S0377-2217(02)00725-7

Hogarth, R. M. (1975). Cognitive Processes and the Assessment of Subjective Probability Distributions. *Journal of the American Statistical Association, 70*(350), 271–289. https://doi.org/10.1080/01621459.1975.10479858

Ishizaka, A., & Labib, A. (2011). Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2011.04.143

Keeney, R. L. (1992). *Value-focused thinking: A path to creative decisionmaking*. Cambridge, Mass.: Harvard Univ. Press.

Kendall, M. G., & Smith, B. B. (1940). On the method of paired comparisons. *Biometrika, 31*(3/4), 324–345. https://doi.org/10.2307/2332613

Liang, F., Brunelli, M., & Rezaei, J. (2020). Consistency issues in the best worst method: Measurements and thresholds. *Omega, 96*, Article 102175. https://doi.org/10.1016/j.omega.2019.102175

Lin, C.-C. (2007). A revised framework for deriving preference values from pairwise comparison matrices. *European Journal of Operational Research, 176*(2), 1145–1150. https://doi.org/10.1016/j.ejor.2005.09.022

Liu, F., Qiu, M.-Y., & Zhang, W.-G. (2021). An uncertainty-induced axiomatic foundation of the analytic hierarchy process and its implication. *Expert Systems with Applications, 183*, Article 115427. https://doi.org/10.1016/j.eswa.2021.115427

Liu, Y., Eckert, C. M., & Earl, C. (2020). A review of fuzzy AHP methods for decision-making with subjective judgements. *Expert Systems with Applications, 161*, Article 113738. https://doi.org/10.1016/j.eswa.2020.113738

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York: Dover Publications.

del Moral, M. J., Chiclana, F., Tapia, J. M., & Herrera-Viedma, E. (2018). A comparative study on consensus measures in group decision making. *International Journal of Intelligent Systems, 33*(8), 1624–1638. https://doi.org/10.1002/int.21954

Moreno-Jiménez, J. M., Aguarón-Joven, J., Escobar-Urmeneta, M. T., & Turón-Lanuza, A. (1999). Multicriteria procedural rationality on SISDEMA. *European Journal of Operational Research, 119*(2), 388–403. https://doi.org/10.1016/S0377-2217(99)00141-1

Mosadeghi, R., Warnken, J., Tomlinson, R., & Mirfenderesk, H. (2015). Comparison of Fuzzy-AHP and AHP in a spatial multi-criteria decision making model for urban land-use planning. *Computers, Environment and Urban Systems, 49*, 54–65. https://doi.org/10.1016/j.compenvurbsys.2014.10.001

Nemeshaev, S., Barykin, L., & Dadteev, K. (2021). Selection of experts for scientific and technical expertise based on semantic search. *Procedia Computer Science, 190*, 643–646. https://doi.org/10.1016/j.procs.2021.06.102

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., … Rakow, T. (2006). *Uncertain judgements: Eliciting experts' probabilities*. John Wiley & Sons.

Opricovic, S., & Tzeng, G.-H. (2007). Extended VIKOR method in comparison with outranking methods. *European Journal of Operational Research, 178*(2), 514–529. https://doi.org/10.1016/j.ejor.2006.01.020

Ravinder, H. V. (1992). Random error in holistic evaluations and additive decompositions of multiattribute utility—An empirical comparison. *Journal of Behavioral Decision Making, 5*(3), 155–167. https://doi.org/10.1002/bdm.3960050302

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega, 53*, 49–57. https://doi.org/10.1016/j.omega.2014.11.009

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology, 15*(3), 234–281. https://doi.org/10.1016/0022-2496(77)90033-5

Saaty, T. L. (1980). *The analytic hierarchy process: Planning, priority setting, resource allocation*. New York; London: McGraw-Hill International Book Co.

Sáenz-Royo, C., Chiclana, F., & Herrera-Viedma, E. (2022). Functional representation of the intentional bounded rationality of decision-makers: A laboratory to study the decisions a priori. *Mathematics, 10*(5), 739. https://doi.org/10.3390/math10050739

Sáenz-Royo, C., Chiclana, F., & Herrera-Viedma, E. (2023a). Intentional bounded rationality methodology to assess the quality of decision-making approaches with latent alternative performances. *Information Fusion, 89*, 254–266. https://doi.org/10.1016/j.inffus.2022.08.019

Sáenz-Royo, C., Chiclana, F., & Herrera-Viedma, E. (2023b). Steering committee management. Expertise, diversity, and decision-making structures. *Information Fusion, 101888*.

Sellak, H., Ouhbi, B., Frikh, B., & Ikken, B. (2019). Expertise-based consensus building for MCGDM with hesitant fuzzy linguistic information. *Information Fusion, 50*, 54–70. https://doi.org/10.1016/j.inffus.2018.10.003

Simon, H. A. (1947). *Administrative behavior*. New York: Free Press.

Song, Z., Zhu, H., Jia, G., & He, C. (2014). Comprehensive evaluation on self-ignition risks of coal stockpiles using fuzzy AHP approaches. *Journal of Loss Prevention in the Process Industries, 32*, 78–94. https://doi.org/10.1016/j.jlp.2014.08.002

Sugden, R. (1985). Why be consistent? A critical analysis of consistency requirements in choice theory. *Economica, 52*(206), 167–183. https://doi.org/10.2307/2554418

Temesi, J. (2011). Pairwise comparison matrices and the error-free property of the decision maker. *Central European Journal of Operations Research, 19*(2), 239–249. https://doi.org/10.1007/s10100-010-0145-8

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34*(4), 273–286. https://doi.org/10.1037/h0070288

Triantaphyllou, E. (2000). *Multi-criteria decision making methods: A comparative study*. Boston, MA: Springer, US. Retrieved from http://www.springerlink.com/content/978-1-4757-3157-6.

Tsoukiàs, A. (2008). From decision theory to decision aiding methodology. *European Journal of Operational Research, 187*(1), 138–161. https://doi.org/10.1016/j.ejor.2007.02.039

Vargas, L. G. (1982). Reciprocal matrices with random coefficients. *Mathematical Modelling, 3*(1), 69–81. https://doi.org/10.1016/0270-0255(82)90013-6

Wallenius, J., Dyer, J. S., Fishburn, P. C., Steuer, R. E., Zionts, S., & Deb, K. (2008). Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. *Management Science, 54*(7), 1336–1349. https://doi.org/10.1287/mnsc.1070.0838

Wallsten, T. S., & Budescu, D. V. (1983). State of the art—Encoding subjective probabilities: A psychological and psychometric review. *Management Science, 29*(2), 151–173. https://doi.org/10.1287/mnsc.29.2.151

Xu, Y., Liu, S., Wang, J., & Shang, X. (2022). Consensus checking and improving methods for AHP with q-rung dual hesitant fuzzy preference relations. *Expert Systems with Applications, 208*, Article 117902. https://doi.org/10.1016/j.eswa.2022.117902

Xu, Z., & Liao, H. (2014). Intuitionistic fuzzy analytic hierarchy process. *IEEE Transactions on Fuzzy Systems, 22*(4), 749–761. https://doi.org/10.1109/TFUZZ.2013.2272585

Zanakis, S. H., Solomon, A., Wishart, N., & Dublish, S. (1998). Multi-attribute decision making: A simulation comparison of select methods. *European Journal of Operational Research, 107*(3), 507–529. https://doi.org/10.1016/S0377-2217(97)00147-1