Pesquisa Operacional (2024) 44: e283585 p.1-27
doi: 10.1590/0101-7438.2023.043.00283585

Printed version ISSN 0101-7438 / Online version ISSN 1678-5142
www.scielo.br/pope
ARTICLES

EFFECTS OF HOLISTIC EVALUATIONS ON FITRADEOFF METHOD USING A SIMULATION STUDY

Manoel Lucas Sousa Ribeiro[1], Tarsila Rani Soares de Vasconcelos[2], Eduarda Asfora Frej[3*] and Adiel Teixeira de Almeida[4]

Received February 21, 2024 / Accepted May 11, 2024

ABSTRACT. The Flexible and Interactive Tradeoff (FITradeoff) multicriteria method was developed under the scope of the Multiattribute Value Theory (MAVT) to elicit preferences under a partial information perspective. The flexibility of this method has been recently enhanced by the combination of two paradigms of preference modeling in the decision process, namely elicitation by decomposition (ED) and holistic evaluation (HE), improving the preferences elicitation task in a sense that fewer elicitation questions are expected to be answered by the Decision Maker (DM) to find a recommendation for the multicriteria decision making (MCDM) problem. The combination of two preference modelling paradigms in FITradeoff is an innovative characteristic on MCDM methods which deserves further analysis on its implications for the practice of decision making, not yet investigated in previous works. Considering that, this work aims to evaluate the impact of conducting HE in FITradeoff in terms of the expected reduction of the number of questions (NQ) that shall be answered to obtain the problem results. To do so, a simulation study has been carried out, in which multiple criteria problems have been generated and solved by simulating different preferences profiles. Hypotheses on the reduction of the number of questions were formulated and tested for identifying the actual effect of two different types of HE in the elicitation process: selection of the best alternative and elimination of the worst alternative considering a previously defined subset. Besides, we also analyze the relationship between the HE effects and important variables of MCDM problems such as number of criteria (NC), number of alternatives (NA), and weights distribution patterns (WD). The results obtained show that

*Corresponding author

[1]Universidade Federal de Pernambuco, Center for Decision Systems and Information Development (CDSID), Av. Acadêmico Hélio Ramos, s/n, Cidade Universitária, Recife, PE, Brazil – E-mail: ribeiro.manoelsousa@gmail.com – http://orcid.org/0000-0002-2837-9150

[2]Universidade Federal de Pernambuco, Center for Decision Systems and Information Development (CDSID), Av. Acadêmico Hélio Ramos, s/n, Cidade Universitária, Recife, PE, Brazil – E-mail: tarsila.vasconcelos@ufpe.br – http://orcid.org/0000-0002-3423-7036

[3]Universidade Federal de Pernambuco, Center for Decision Systems and Information Development (CDSID), Av. Acadêmico Hélio Ramos, s/n, Cidade Universitária, Recife, PE, Brazil – E-mail: eafrej@cdsid.org.br – http://orcid.org/0000-0001-6529-9910

[4]Universidade Federal de Pernambuco, Center for Decision Systems and Information Development (CDSID), Av. Acadêmico Hélio Ramos, s/n, Cidade Universitária, Recife, PE, Brazil – E-mail: almeida@cdsid.org.br – http://orcid.org/0000-0002-2757-1968

holistic evaluation judgments have a significant impact on the reduction of effort made by the DM in the elicitation process, in a sense that the total number of questions reduces in a significant manner, considering either selection tasks or elimination tasks in HE.

Keywords: multicriteria decision making, FITradeoff method, holistic evaluation, simulation.

1 INTRODUCTION

Multicriteria decision-making methods developed under the scope of the multi-attribute value theory (MAVT) (Keeney and Raiffa, 1976) commonly use an additive aggregation model to evaluate alternatives under a value-based model of which the parameters should be defined according to a structured process considering the Decision Maker's (DM) preferences and values structure (Belton & Stewart, 2002). Recommendations are given based on the additive aggregation of the alternatives consequences, which are evaluated against each of the problem criteria, weighted by a scaling constant vector that represents the DM's preferences. Eliciting criteria scaling constants in additive models is a challenging task, since these are critical parameters of the value-based model that should be carefully defined, considering the ranges of consequences of the alternatives in each criterion, and not only the criteria degree of importance (Keeney & Raiffa, 1976). Motivated by inherent challenges on criteria scaling constants (commonly called weights) elicitation, the FITradeoff (Flexible and Interactive Tradeoff) (De Almeida et al, 2016; De Almeida et al., 2021) was developed. This method considers partial information of the DM's preferences in the elicitation process, in the sense that less effort is required from DMs.

The method is operated by an interactive and freely available Decision Support System (DSS) with several flexibility features, which is one of the main reasons why it is widely applied for solving practical problems. FITradeoff has been used to solve several real problems involving outsourcing decisions (Rodriguez et al, 2023), urban mobility (Oliveira, Morais & Siebert, 2023), circular economy (Lugo et al, 2023), supply disruption management (Carvalho, Roselli & Figueira, 2023), prioritization of failure modes (Zanazzi, Zanazzi & Pontelli, 2023), natural gas facility location (De Lacerda et al, 2021), public bidding (De Araujo et al, 2022), supplier performance evaluation (Rodrigues et al, 2020), endemic diseases combat strategy (Dos Santos et al, 2022), energetic matrices related decisions (Fossile et al, 2020) and many others (Santos et al, 2023; Pessoa, Roselli & De Almeida, 2022; Silva & Morais, 2022; Ribeiro et al, 2021; Monte & Morais, 2019; De Macedo, Mota & Sola, 2018).

From an innovative perspective, De Almeida et al (2021) recently proposed a significant innovation in the FITradeoff decision process: the combination of two paradigms of preference modeling; elicitation by decomposition (ED) and holistic evaluations (HE). In ED, DMs answer preference questions in the consequences space, considering tradeoffs amongst criteria. In HE, DMs declare preferences among the alternatives of the problem, directly, stating which alternative is the best one among a subset (selection task) or eliminating the worst one in a subset (elimination task). De Almeida et al (2021) state that the combination of paradigms brings two main contributions to the decision process: including an alternative source of information and enhancing the efficiency of the method by reducing the number of questions (NQ) needed to solve a problem.

The combination of two types of information in the elicitation process of FITradeoff is an innovative aspect in the world of preference elicitation in MCDM methods, which brings many expectations of benefits in the decision-making process including mainly the increase in the efficiency of the method, in a sense that a solution can be achieved with less amount of information provided. However, even though previous application papers illustrate the combination of preference modeling paradigms (Czekajski et al., 2023) and its potential benefits, no previous study has made efforts to deeply investigate the actual effects of such combination in practical terms, reaching conclusions on the real effects of the incorporation of HE in the reduction of questions needed to find a solution.

Given the arguments presented by De Almeida et al (2021) to support the assertion that the holistic evaluation increases the efficiency of the FITradeoff method, and considering that it may be performed in two different ways (i.e., elimination and selection tasks), some questions remain to be clarified, such as: what is the reduction in NQ due to holistic evaluation? What is the influence of increasing the number of criteria and alternatives in NQ? Is there a significant difference between the reduction effects provided by the evaluations carried out for selection and elimination tasks? Which of these tasks contributes more to a reduction in NQ?

In order to investigate those questions, a simulation-based study has been carried out in this work. A set of research hypotheses was defined, and several simulation scenarios considering different types of weights distribution patterns, different numbers of criteria, and different numbers of alternatives were set. The findings obtained with this study allow for answering all the aforementioned questions. The analysis of multicriteria methods through simulation studies is a common practice in the literature, in a sense their performance can be actually evaluated in a more realistic manner, rather than considering single individual applications. Alfares & Duffuaa (2015) argue that the use of simulation techniques to conduct studies to evaluate the performance of decision support methods is quite flexible and reliable, since it becomes possible to generate and evaluate a significant amount of data, in addition to allowing the evaluation of several different scenarios.

Alfares & Duffuaa (2015) conducted a study to evaluate the performance of methods for generating weights based on the ranking of criteria, which in their study is carried out through simulations. In this sense, scenarios are considered in which the number of criteria and alternatives, and the distribution of weights (uniform, normal, or exponential) are varied allowing the analysis of five methods (RS, RR, ROC, GW, VSL). The result of the study shows that no generating weights method can outperform the others in all the scenarios, even though, the work provides interesting insights regarding the choice of a method based on the expected distribution of weights.

Tučník & Bureš (2016) evaluated the applicability of four multi-criteria methods for future applications in large multi-agent-based simulation models. In this study, the main metric for comparing them was the time required to run each one during the simulations. Although all methods proved to be applicable, the VIKOR method stood out from the others even considering different hardware configurations.

Mehdi et al (2018) performed a simulation study to investigate the phenomenon of order reversal in TOPSIS and EDAS methods, establishing a comparative analysis between them. According to the authors, order reversal is an undesirable phenomenon that occurs from the inclusion of new characteristics to the previously solved problem, generating changes in the ranking of alternatives. The study considered multiple scenarios and three performance indicators.

The FITradeoff method was already evaluated through a simulation study by Mendes et al (2020), but only the elicitation by decomposition was considered in the analysis. The authors sought to evaluate the performance of the method when compared to the benchmarks established for the classical Tradeoff procedure (Keeney & Raiffa, 1976), which also works with elicitation by decomposition, but considering complete information from DMs, while FITradeoff requires partial information. Among the conclusions, it is possible to highlight that in 81% of the cases, the method manages to reduce the space of actions to 5 potentially optimal alternatives only by informing the ordering of the scaling constants. Another conclusion obtained by Mendes et al (2020) refer to the low influence of the increase in the number of alternatives over the number of questions asked in the method's elicitation by decomposition. In contrast, the number of criteria and the distribution of weights had much more influence on the observed performance. Mendes et al (2020) did not consider holistic evaluation in their analysis. The focus of the present paper is therefore to evaluate the actual influence of holistic evaluation judgments in the performance of the FITradeoff method, considering both selection and elimination tasks.

The contribution of the present paper therefore relies on performing a simulation study to actually investigate the effects of performing Holistic Evaluations in the FITradeoff multicriteria method. In the work of De Almeida et (2021), a combination of two types of preference modeling (elicitation by decomposition and holistic evaluations) was proposed to be part of the FITradeoff decision process, with a view to improving the decision-making process for the decision makers, with lots of potential benefits that could be achieved with such combination. Other studies (e.g. Cunha et al. (2020) and Czekajski et al. (2023)) presented single applications with the FITradeoff method illustrating the combination of preference modeling paradigms in practical MCDM problems, showing how the holistic evaluation can be conducted, from a decision-making perspective, to improve the decision process. However, all the benefits of Holistic Evaluations cited by these previous papers were verified in an empirical manner, with single application cases within a specific scenario, without a proper study focused on that, considering different and multiple scenarios. Our study here concerns a much more general analysis, considering different scenarios throughout a simulation analysis, in order to verify the actual effects of holistic evaluations in the decision process with FITradeoff, based on multiple research hypotheses. Moreover, statistical tests were conducted in order to verify the statistical significance of our findings in the simulation study.

Hence, while the focus of those previously published papers was to illustrate an application of the FITradeoff method with the combination of preference modeling types from the view of the decision maker, highlighting potential benefits for him/her, our research is much more focused on an analysis of the efficiency of the method in terms of finding a solution and the extent to which holistic evaluations contribute with that, considering two different types of HE: selection and elimination tasks. Therefore, the contribution of this work relies mainly on verifying and testing hypotheses related to the incorporation of HE judgments in FITradeoff which were only raised empirically by previous studies.

This paper is organized as follows. Section 2 summarizes relevant prior research and discusses the FITradeoff method. In Section 3, the research hypotheses of this study are formulated. Section 4 describes the simulation experiment design and application. The main results are reported in Section 5, followed by the conclusions and final remarks in Section 6.

2 BACKGROUND ON THE FITRADEOFF MULTICRITERIA METHOD

The FITradeoff (Flexible Interactive Tradeoff) proposed by De Almeida et al (2016) is an MCDM method that stands out in the context of preference elicitation because it requires partial information from the decision maker, which reduces cognitive effort. In addition, the method is flexible, in the sense that the DM chooses when to stop the process and how he wants to provide information (De Almeida, Frej & Roselli, 2021). Besides, the FITradeoff process is interactive, which allows the DM to understand the effect of each new piece of information provided on the recommendation, resulting in an increase in the DM's confidence and learning throughout the elicitation (Da Silva et al, 2021; Dutta et al, 2021).

The method is based on the axiomatic structure of the traditional Tradeoff procedure (Keeney & Raiffa, 1976), however, it incorporates the concept of flexible elicitation with partial information, so that it is not necessary to obtain indifference points to generate a recommendation, as in the classical tradeoff. FITradeoff was originally designed to solve problems in the choice problematic, which consists of selecting the best alternative given the DM's preferences, to this end, it evaluates the potential optimality of the alternatives at each iteration, classifying them into potentially optimal, dominated, or, in case a solution is found, optimal alternative (De Almeida et al, 2016). The method was also expanded to the ranking (Frej et al., 2019) and sorting (Kang et al., 2020) problematics.

The FITradeoff method is operated by means of an interactive Decision Support System (DSS), through which the DM provides preferential information by answering questions. Different types of information are gathered from the DM according to the type of question the DM wishes to answer at each interaction cycle: elicitation by decomposition question (considering elements in the consequences space) or holistic evaluation question (considering elements in the alternatives space). Those two types of evaluation are further detailed in Section 2.1.

2.1 Combination of Paradigms for Preference Elicitation in the FITradeoff Method

De Almeida, Frej & Roselli (2021) proposed a significant innovation in the FITradeoff method by combining two distinct paradigms for preference elicitation, which allows the DM to switch between two distinct elicitation procedures throughout the process. Elicitation by decomposition (ED) was the preference elicitation procedure initially included in the FITradeoff method and consists of a procedure similar to the one discussed by Keeney & Raiffa (1976), illustrated in Figure 1, in which two consequences are presented to the decision maker and he/she must choose the one that brings him/her greater satisfaction.

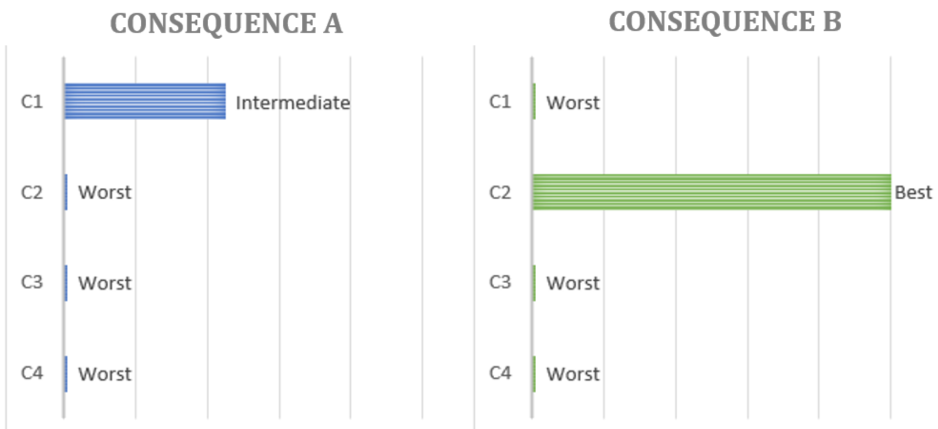

Figure 1 – Hypothetical Consequences compared in the Elicitation by Decomposition (ED).

The ED questions usually present a consequence for which there is an intermediate performance in one of the criteria and the least desirable performance for all the others, while the second consequence presents the most preferable performance in a criterion and the worst performance evaluated in all other criteria. Once the DM expresses a strict preference relation (being possible to express indifference, if they feel confident to do so), it is possible to obtain a new inequality for the model. As the scale considered is an interval one, when evaluating the overall value of the hypothetical consequences, the obtained inequality relates pairs of scaling constants. Moreover, it is worth mentioning that in this procedure the preference elicitation is conducted in the space of consequences.

Since its initial proposition, the method already allowed holistic evaluations (HE) to be carried out to finalize the problem when the DMs did not want to proceed with the elicitation of preferences because they already had a satisfactory partial result. Now, they can choose to intersperse the process between the two procedures. Few methods use holistic assessments for preference elicitation (Da Silva et al, 2021), and FITradeoff is the first one to put these two paradigms together.

In HE, the decision maker can compare real alternatives of the problem and express preference relations involving them. Since interactive elicitation procedures can contribute to the DM's learning about their own preferences (Da Silva et al, 2021; De Almeida et al., 2015), along with the procedure the DM may become confident about expressing their aspirations regarding the problem alternatives, so that kind of information can also be included in the linear programming model, shortening the process.

As for the HE to be carried out, FITradeoff presents four distinct types of visualization (figure 2 presents examples of bar charts and spider graphs). The DM shall then choose a set of alternatives that they wish to evaluate holistically and the visualization kind that makes them most confident to express their preferences.

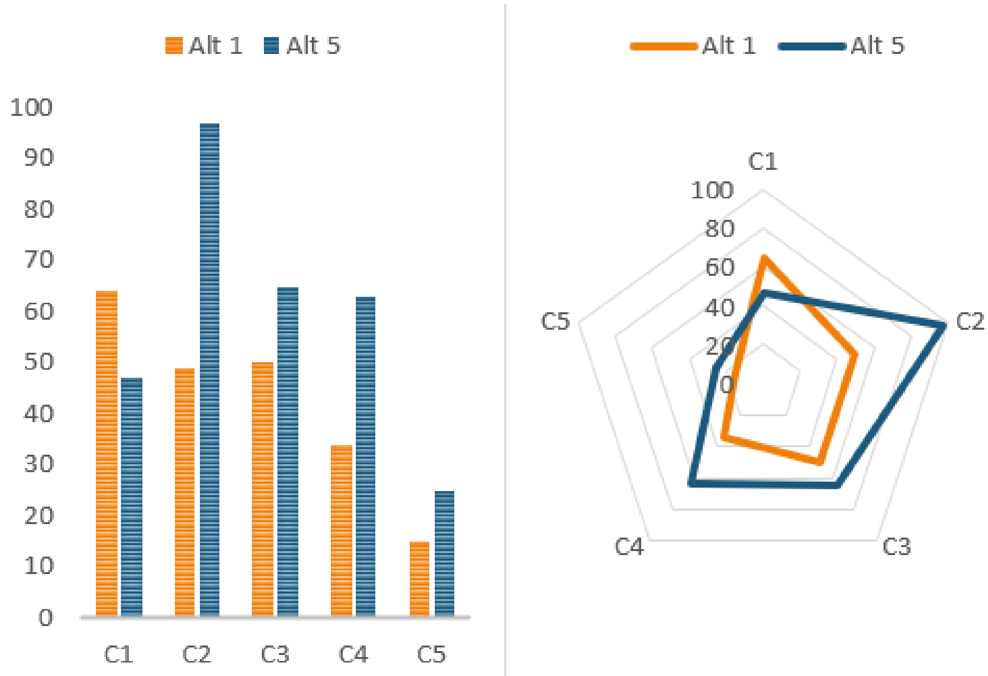

Figure 2 – Examples of Visualizations for Holistic Evaluation.

In the choice problematic, holistic evaluation can be conducted through two distinct procedures, namely selection and elimination. The selection task can be performed when, upon evaluating the subset of alternatives, the DM feels confident to point out the best alternative in the evaluated subset. The elimination procedure, on the other hand, can be performed when the DM feels more confident to point out the worst alternative in a subset.

As mentioned, HE in the choice problematic can be performed by considering subsets containing as many alternatives as the DM wishes to evaluate. In all cases, expressions of inequalities on global values of the compared alternatives will be included in the linear programming model, thus reducing the space of feasible weights. It should be considered that when performing a holistic evaluation in the choice problematic, a reduction in the number of POAs will necessarily be observed. With HE, preference elicitation is conducted in the space of actions of the problem.

As FITradeoff does not require the definition of indifferences, it is not possible to obtain a system of equations, as proposed by Keeney & Raiffa (1976), but rather a system of inequalities containing infinite possible solutions. Thus, there will be numerous additive aggregation functions that satisfy the system, so as new information is incorporated into the model, it is possible to reduce the set of feasible solutions. At each interaction cycle, the information given by the DM is converted into an inequality and incorporated as constraints of a linear programming problem (LPP) model, that runs to verify what are the Potentially Optimal Alternatives (POAs) for the problem (De Almeida et al., 2016; De Almeida et al., 2021). For each potentially optimal alternative (POA) found in the previous cycle, the LPP model runs again with the new set of constraints, which was updated with the additional information obtained, to evaluate whether it remains potentially optimal. If so, the alternative is kept for the following interactions, otherwise, it is eliminated from the set. The solution of the problem is therefore obtained when only one (optimal) alternative remains, or the set of remaining alternatives is considered equivalent according to the DM's preferences. The FITradeoff LPP model is as follows (De Almeida et al., 2016; De Almeida et al., 2021):

$$\max \quad V(A_i) = \sum_{j=1}^{m} k_j v_j(x_{ij}) \tag{1}$$

s/t

$$k_1 \geq k_2 \geq \ldots \geq k_m \tag{2}$$

$$k_j v_j\left(x_j^{'}\right) > k_{j+1} \tag{3}$$

$$k_j v_j\left(x_j^{''}\right) < k_{j+1} \tag{4}$$

$$\sum_{j=1}^{m} k_j v_j(x_{ij}) \geq \sum_{j=1}^{m} k_j v_j(x_{tj}) \quad i \neq t \tag{5}$$

$$\sum_{j=1}^{m} k_j v_j(x_{zj}) > \sum_{j=1}^{m} k_j v_j(x_{lj}) \quad z \neq l \tag{6}$$

$$k_j \geq 0 \tag{7}$$

$$\sum_{j=1}^{m} k_j = 1 \tag{8}$$

$$j = 1, \ldots, m \qquad i = 1, \ldots, n \quad t = 1, \ldots, n \tag{9}$$

Equation (1) represents the objective function to be maximized, which consists of the overall value of alternative A_i, V(A_i), computed by an additive aggregation sum of the value of the consequence (x_{ij}) of alternative A_i in criterion j, $v_j(x_{ij})$, considering a 0-1 normalized interval scale, weighted by criteria weights k_j. The inequality in (2) includes the ordering of the scaling constants in the model, information is provided by the DM at the beginning of the preference elicitation process. Inequalities (3) and (4) represent the information obtained when the DM answers elicitation by decomposition questions, relating a pair of scaling constants. The set of inequalities represented in (5) configures the constraints of potential optimality of alternative A_i, in which a vector of weights that makes the global value of alternative A_i greater than or equal to the global value of all other alternatives in the set is sought. The inequalities presented in (6) are obtained through HE, which represents an alternative source of information for the model and relates to the global values of the alternatives evaluated holistically, leading to the elimination of one or more of them from the POA set. Finally, expression (7) concerns the domain of the decision variables, equation (8) ensures the normalization of the weights, while expression (9) presents the possible values for the variable indices.

The decision variables of the LPP model (1-9) are the criteria scaling constants. If there is at least one vector of weights $[k_1, k_2, \dots k_m]$ that satisfies all constraints (2-9), then alternative A_i is considered a potentially optimal alternative for the problem. Otherwise, it is considered dominated and is eliminated from the process.

De Almeida, Frej & Roselli (2021) point out that HE presents two main contributions, providing an additional source of information for the model, which increases the flexibility of the method, and accelerating the preference elicitation process, since new constraints involving all scaling constants simultaneously are included in the model. Furthermore, HE can be performed to provide additional information to the model, or even to finalize the elicitation process.

Having said that, the present work looks to investigate whether HE has a significant contribution to FITradeoff efficiency improvement, which is here measured by the number of questions (NQ) that a DM must answer until converging to the problem solution. Attention should be given to NQ once the DMs in general do not dispose of much time, besides, stating many questions may lead to inconsistencies and tiredness along the process, which the methods usually try to avoid. Based on the discussion made by the authors in their proposition, it was possible to speculate on the HE mechanism, allowing for stating a few hypotheses, which are introduced in the next section.

3 RESEARCH HYPOTHESES

In this topic, the research hypotheses will be developed considering aspects from the literature as well as observations obtained from a descriptive analysis of the data gathered.

3.1 The effect of performing a HE right after scaling constants ordering step

As mentioned in the previous topic, De Almeida, Frej & Roselli (2021) argue that performing HE in the FITradeoff method allows for introducing additional information into the decision model. In contrast to the information obtained through the ED, the HE includes preferences relations involving real alternatives of the problem, which result in inequalities relating all the scaling constants simultaneously, which is stronger information, in a sense that affects the space of weights compared to an inequality that related only two scaling constants, as those obtained from ED. Even though previous studies (Cunha et al. (2020); Czekajski et al. (2023)) have shown that, for certain specific application cases, holistic evaluations can reduce the total number of questions made for the DM, this should be formally tested considering a significant number of scenarios.

By performing a holistic assessment, a subset of POAs containing at least one alternative is necessarily excluded from the problem, and depending on how close the evaluated alternatives are from one another, this information may accelerate the process, making it possible even to finalize it.

To illustrate it, one may think of a situation in which two very close alternatives are evaluated, in this case, the weight vectors space that support the preference of one over the other may be rather small when compared to the whole search space, so providing this information to the model, may eliminate a huge fraction of the space, leading to a more rapid convergence to the problem solution.

On the other hand, as the questions conducted in the HE are not affected by the holistic assessment, if the DM is consistent with his preferences, it is not possible to observe an increase in the number of questions made. So, either the problem is finalized before stating all the questions that would be made given DM's preferences, or the method makes them all. Based on that, it is possible to state the first research hypothesis:

- ***Research Hypothesis 1.*** *Performing holistic evaluations significantly reduces the number of questions necessary to find a solution in FITradeoff.*

3.2 The effect of each of the HE's procedures

Neuroscience studies show that there is a difference, from the cognitive effort point of view, when performing each of the procedures (Pessoa, Roselli & De Almeida, 2021). These same studies show that although there is no difference in the success rate when performing each of the procedures, holistic elimination performs better in terms of DM's response time and accuracy of responses, which in turn is defined as the time until the DM gives a correct answer (that is consistent with preferences).

It is also known that from the point of view of the linear programming problem and the effect of the holistic evaluation on the set of POAs, each procedure generates different constraints and consequently is expected to present different effects.

Suppose that in holistic evaluation n alternatives are compared and evaluated with respect to the m criteria of the problem. If the selection procedure is performed and alternative *i* is chosen as the best one from the subgroup, *n-1* constraints will be included in the LPP, leading to the elimination of *n-1* alternatives from the POA subset.

On the other hand, if the elimination procedure is performed and alternative *j* is said to be the least preferable of the subgroup, *n-1* constraints will also be included in the LPP, but in this case, only alternative *j* will be eliminated for sure from the subset of POAs. Therefore, the following research hypothesis is stated:

- ***Research Hypothesis 2.*** *The selection procedure has a greater effect on the reduction of the NQ than the elimination procedure.*

3.3 The effect of the number of alternatives over the reduction of NQ

Mendes et al (2020) showed that the increase in the number of alternatives in the problem had little effect on the performance of the FITradeoff method, considering only ED. For the case of HE, on the other hand, one may expect a difference in performance, since the number of alternatives influences the number of potentially optimal alternatives (POA), which consequently limits the number of alternatives that can be evaluated holistically.

To better understand that situation one may consider firstly that the number of POAs is directly related to the number of alternatives evaluated in the problem, in addition, if a wider subset of alternatives is considered potentially optimal, it is possible that although they are not initially dominated, there may be many of them that are optimal in regions of the space that are far distant from DM's true aspirations, so, it may be easy to DM to point it out and eliminate them, however in terms of converging to the region in which their preferences are satisfied, this kind of information may have little relevance. Having said that, it is possible to state the following research hypothesis:

- ***Research Hypothesis 3.*** *The increase in the number of alternatives (NA) decreases the reduction effect of NQ provided by the HE.*

3.4 The effect of the number of criteria over the reduction of the NQ

As the number of criteria increases, the number of non-dominated alternatives (solutions) also does so (Deb & Himanshu, 2014). From the literature, it is also known that the number of questions necessary to obtain the most desirable solution also increases. Mendes *et al* (2020) also showed that in the FITradeoff context.

Roselli & De Almeida (2021), on the other hand, argued that the DM should not perform the HE in situations for which there are many criteria due to the cognitive effort necessary to do so, which may lead to errors along the process.

Following the literature information, one may think that in a m-dimensional space, as m increases, more information is needed to delimitate the space region that really satisfies the DM's aspirations, so it is naturally expected that the HE is likely to have a smaller effect in terms of the reduction of questions made. So, the following research hypothesis may be stated:

- ***Research Hypothesis 4.*** *The increase in the number of criteria (NC) decreases the reduction effect of NQ provided by the HE.*

3.5 The effect of the weight's distribution over the reduction of the NQ

The weights or scaling constants are the model parameters that represent the DM's aspirations. As presented by De Almeida *et al* (2016), depending on the distribution of the weights the FITradeoff method follows a distinct heuristic in the ED. Mendes *et al* (2020) showed that more questions are performed when a more skewed distribution is observed.

According to the study developed by Roselli & De Almeida (2021), the distribution of weights plays a significant role in the chance of success when performing the holistic evaluation. So, it is also expected to play a significant role in the HE effects.

Considering all that, one may think that, in the limit, an extremely more skewed distribution would lead some criteria weights to be close to zero, which would mean that some criteria are not influencing much the decision and would be like the situation with a smaller subset of objectives. In that case, as the heuristic applied in the ED always asks questions in the middle of the interval, in cases where the distribution is extremely skewed, it may take some rounds until it captures such kind of information.

HE, in contrast, would much more rapidly depict that, once it considers all the scaling constants simultaneously, hence, it may lead to a faster reduction of the feasible space of solutions. Based on that, the following research hypothesis is drawn.

- ***Research Hypothesis 5.*** *As the weight distribution (WD) becomes more skewed, the HE effect is amplified.*

4 SIMULATION STUDY: DESIGN OF THE EXPERIMENT

In this topic, relevant aspects concerning the simulation experiment performed will be detailed. In Table 1, the main variables analyzed in this study are presented and described.

Table 1 – Variables under analysis.

Variable	Definition
RP	Percentage of reduction of the number of questions needed to solve a problem after performing the HE. $RP = \frac{NQ-NQHE}{NQ}$
Sel	Refers to HE of selection
Eli	Refers to HE of elimination
RP_Sel	Percentage of reduction of the number of questions needed to solve a problem after performing the HE of selection
RP_eli	Percentage of reduction of the number of questions needed to solve a problem after performing the HE of elimination
WD	Ratio between the biggest and smallest scaling constants
POA	Number of potentially optimal alternatives
NC	Number of criteria
NA	Number of alternatives
NQ	Number of questions answered in the elicitation by decomposition
NQHE	Number of questions answered in the elicitation by decomposition having performed an HE

4.1 Data gathering and experiment design

A computational simulation was conducted to gather the data to be analyzed. In the simulations, multicriteria problems were randomly generated. The parameterization of the experiment was defined as follows:

a) Number of alternatives and criteria: Set C = $\{3, 4, 5, 7, 10, 15\}$ was defined for the number of criteria, and set A = $\{5, 10, 15, 20, 30, 50, 70\}$ for the number alternatives.

   Speaking about the criteria set, a number of criteria smaller than three, besides being uncommon in the literature, could represent a problem that is too simple, on the other hand, as the number of criteria becomes much larger, the possibility of observing preferential dependences or even redundancies between the criteria and/or the objectives measured by them also increases, which according to Keeney (1992) should be avoided.

   In terms of the number of alternatives, sets involving a very large number of alternatives are likely to include dominated alternatives, which would be eliminated in the first LPP in the FITradeoff, becoming, in this way, irrelevant for this study.

Table 2 – Scenarios under analysis.

		NC					
		3	**4**	**5**	**7**	**10**	**15**
	5			1000	1000	1000	1000
	10			1000	1000	1000	1000
	15			1000	1000	1000	1000
NA	**20**	1000	1000	1000	1000	1000	1000
	30	1000	1000	1000	1000	1000	1000
	50	1000	1000	1000			
	70	1000	1000	1000			

b) Weights patterns and distribution: As for considering different distributions of weights, it was identified four patterns in literature:

   (a) Distribution 1: Equal weights: $\frac{k_1}{k_m} = 1$

   (b) Distribution 2: Random and similar weights: $\frac{k_1}{k_m} = 1,5$

   (c) Distribution 3: Random and distinct weights: $\frac{k_1}{k_m} = 4$

   (d) Distribution 4: Random and markedly distinct weights: $\frac{k_1}{k_m} = 10$

   k_j is the scale constant of criterion j, $j = 1$ is the index of the criterion with the greatest scale constant value while $j = m$ is the index of the criterion with the smaller scale constant value, and m denotes the number of criteria of the problem.

c) Simulated scenarios: Table 2 shows the scenarios considered for the simulation, highlighting the number of instances solved for each scenario. It is worth noting that each instance in each of the scenarios was solved considering each of the weight standards, thus, the scenarios generated by the combination of number of criteria, number of alternatives, and weight standards are equivalent to 120, as 1000 instances were created for each scenario, there are a total of 120,000 instances created, which was considered adequate for the present study taking into account the trade-off between the number of samples and the time required to generate and solve them.

d) Simulation procedure: To perform the simulation, a system was developed using Pascal-Object language, besides MySQL as a database. Initially, the system generates the consequence matrix randomly considering a uniform continuous distribution and the setup relative to the scenario being considered (number of criteria x number of alternatives x weights distribution).

   For this step, a parameterization was performed that defines a probability of 80% for natural scale criteria and 20% for constructed scale, in addition, the direction of preferences as well as the number of levels of the constructed scale are defined randomly. This parameterization was defined since it is recommended to use natural scale criteria whenever possible since they present a higher level of information, so scales should be created only when it is not possible to use a natural criterion (De Almeida, 2013).

   The system applies linear value functions for all the criteria and generates random weights considering each distribution, then it runs the first LPP obtaining the non-dominated alternatives. In the sequence, the system generates all the possible combinations of the POAs. For each of them, the system simulates the answers of a DM based on the weights vector, making it possible to solve the problem three times.

   The first time each problem is solved, no HE is performed, the idea is to identify the number of questions necessary to solve it using only the ED (in this study, not performing HE means that no treatment was applied). In sequence, the problem is solved a second time, in which the HE procedure of selection is applied (which consists of the applications of the 'Sel' treatment) only once right after ranking the scaling constants. Finally, the problem is solved once more, this time the applied procedure is the elimination (which consists of the applications of the 'Elim' treatment).

   As mentioned, for each subset of POAS (obtained by the combination of them), it is stored the number of questions necessary to solve the problem in each of the three rounds.

e) By applying the aforementioned procedure, a set with N = 8.746.803 (eight million, seven hundred, forty-six thousand, eight hundred and three) observations were stored concerning each round. In this way, each HE procedure was analyzed by comparing the treatment group (the group of N observations in which the procedures were separately applied) against the control group (formed by the N observations that had not received any treatment).

Figure 3 summarizes the whole simulation process carried out for each simulated scenario defined.

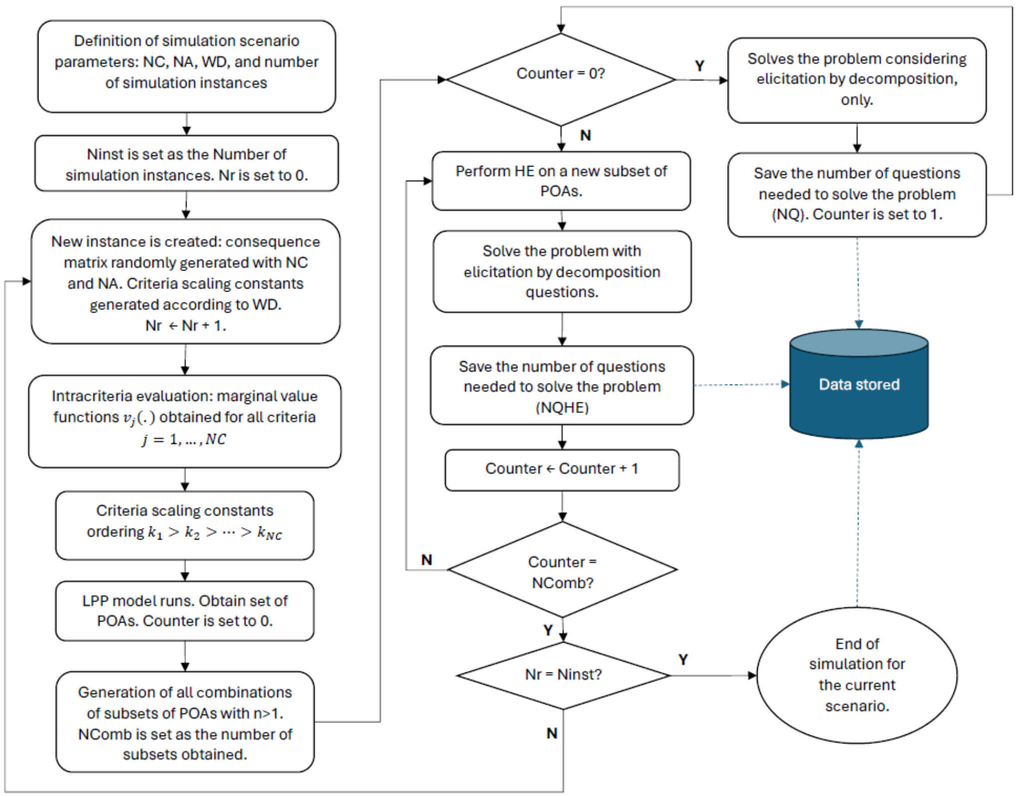

Figure 3 – Simulation process.

5 RESULTS AND DISCUSSION

According to the simulation study described in Section 4, the research hypotheses raised in Section 3 could be deeply investigated. The results obtained are described and discussed in the following subsections.

5.1 The effect of performing a HE right after scaling constants ordering step

Considering all the above-mentioned designed scenarios, after the simulations performed, it was found that performing a holistic evaluation right after the criteria scaling constants ordering step brings an average reduction of 25.81% in the total number of questions, considering both selection and elimination tasks. I.e., the average value obtained for variable RP considering all simulation scenarios and both selection and elimination tasks was 25.81%.

Considering the first research hypothesis raised in this study, the following alternative hypotheses were formulated: i) NQ is higher than NQHE for selection; ii) NQ is higher than NQHE for elimination. Considering the large amount of data obtained with the simulations carried out, a normal distribution was assumed, and t-student tests were performed for comparison of means since population variances are unknown (Seward and Doane, 2014). Each combination of the number of criteria, number of alternatives, and distribution of weights determined a scenario to be evaluated. So, a total of 120 scenarios were considered. For all of them, the p-values were approximately equal to 0. This indicates that, in fact, NQ is higher than NQHE and also higher than NQHE for elimination.

It means that, when performing a holistic evaluation – either for selection or elimination purposes –, the reduction in the total number of questions answered by the DM in the elicitation process is reduced in a statistically significant manner. This finding corroborated the empirical ideas of De Almeida et al (2021) and Czekajski et al (2023), which raised the reduction in the number of questions as one of the main benefits of the incorporation of holistic evaluations in the decision process.

5.2 The effect of each of the HE's procedures

Considering all simulation scenarios, the results of this study have shown that a more significant reduction in the number of questions is observed when performing the selection task, rather than elimination task. The average value of reduction in the number of questions was found to be 34.15% for selection task, and 17.46% for elimination task.

To analyze the effect of HE in both selection and elimination tasks, in accordance with Research Hypothesis 2, the following alternative hypothesis was formulated: NQHE for elimination is higher than NQHE for selection. As in the previous analysis, t-Student tests were used for the 120 scenarios analyzed and, for all of them the p-values were approximately equal to 0. This indicates that NQHE for selection is lower than NQHE for elimination. I.e., it is possible to infer that the selection procedure performed through the HE has on average a greater effect in terms of reducing the number of questions asked when compared to the elimination procedure.

The impact of the reduction in the number of questions was determined through the variable RP for selection and elimination. Figure 4 illustrates the behavior of RP for each scenario, considering the four distributions of weights tested in our simulation study. The horizontal axis indicated the scenarios of 'number of criteria x number of alternatives'. For example, 3C_20A indicates that the RP has been calculated for the combination of 3 criteria and 20 alternatives.

By observing Figure 4, the graphics present a clearly descending pattern, which indicates that there is a trend of decreasing in RP when raising the number of criteria, which was already expected. The peaks in the graphics indicate also a seasonal pattern, which is due to the fact that the number of alternatives changes in a cyclic manner along the horizontal axis. It is also possible to see that the reduction percentage for the selection task is higher than the reduction percentage for the elimination task for all scenarios and weights distributions. The spread is

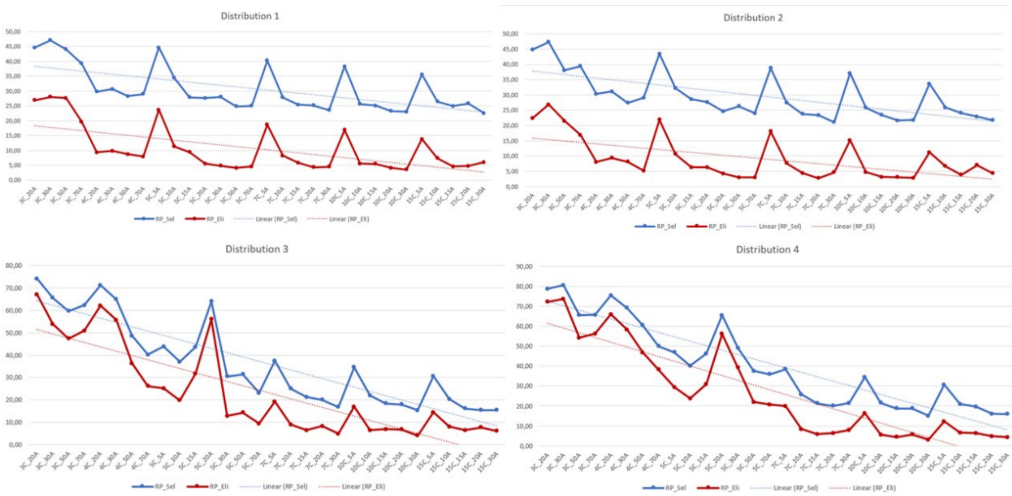

Figure 4 – RPs of each distribution.

higher for distributions 1 and 2, which indicates that, when the weights values are closer to each other, the effect on the reduction of number of questions for the selection task is more significant than for the elimination task. However, it should also be observed that, for distributions 3 and 4 (which indicate a more skewed pattern for the criteria weights), the percentage of reduction in the number of questions is greater than for distributions 1 and 2 (even though the spread between the two different tasks effect is less significative), which also corroborates with research hypothesis 5.

5.3 The effect of the number of alternatives over the reduction of the NQ

Research Hypothesis 3 stated that increasing the number of alternatives being evaluated would lead to a reduction in the effect caused by both the procedures of the HE. To verify this, the Pearson correlation coefficient (ρ) between NA and RP was calculated for both selection and elimination procedures (Table 3). The Pearson correlation statistical test was also applied to test the statistical significance of the correlation, as shown in Table 4.

Table 3 – Pearson's correlation coefficients.

	ρ(NA X RP_SEL)	ρ(NA X RP_ELI)
Distribution 1	-0.051	-0.010
Distribution 2	-0.042	-0.068
Distribution 3	0.210	0.227
Distribution 4	0.322	0.350

Table 4 – Correlation tests.

	NA X RP_SEL	NA X RP_ELI
Distribution 1	t = -0.271	t = -0.055
	p-value = 0.394	p-value = 0.478
Distribution 2	t = -0.221	t = -0.360
	p-value = 0.413	p-value = 0.361
Distribution 3	t = 1.139	t = 1.232
	p-value = 0.868	p-value = 0.886
Distribution 4	t = 1.800	t = 1.978
	p-value = 0.959	p-value = 0.971

For weight distributions 1 and 2, the Pearson correlation coefficient values show that there is a weak negative correlation between the variables. For weight distributions 3 and 4 there is a weak positive correlation between the variables. Since the values obtained for ρ are all below 0.4, the correlations between these variables are very weak, in the sense that we do not expect the correlation to be significant. The results of the Pearson correlation test shown in Table 4 corroborate that, as we can see when analyzing high p-values.

In this sense, we conclude that there is no correlation between the number of alternatives and the reduction in the number of questions, and thus Research Hypothesis 3 is not valid. The graphics in Figure 4 somehow corroborate this finding, since cyclic patterns are observed when changing the number of alternatives, without a clear trend.

As discussed in topic 3.3, as more alternatives are considered in the decision process, the more complex the problem becomes. The alternatives that are not eliminated by the information of the scaling constant ranking may be optimal for different regions from the space. Evaluating holistically a subset of alternatives that are potentially optimal in regions distant from the one that holds the weight vector which would more adequately model the DM's aspirations would add less relevant information to the model, once those regions could be rapidly eliminated from the feasible space after a few questions in the beginning of ED.

5.4 The effect of the number of criteria over the reduction of the NQ

Research Hypothesis 4 stated that increasing the number of criteria of the MCDM problem would lead to a reduction in the effect caused by both the procedures of the HE. To verify this effect, Pearson's correlation coefficient (ρ) between NC and RP was calculated for both selection and elimination procedures (Table 5). The Pearson correlation statistical test was also applied to test the statistical significance of the correlation, as shown in Table 6.

Table 5 – Pearson's correlation coefficients.

	ρ**(NC X RP_SEL)**	ρ**(NC X RP_ELI)**
Distribution 1	-0.436	-0.415
Distribution 2	-0.466	-0.387
Distribution 3	-0.714	-0.662
Distribution 4	-0.767	-0.746

Table 6 – Correlation tests.

	NC X RP_SEL	**NC X RP_ELI**
Distribution 1	t = -2.563	t = -2.414
	p-value = 0.008	p-value = 0.011
Distribution 2	t = -2.790	t = -2.223
	p-value = 0.005	p-value = 0.017
Distribution 3	t = -5.398	t = -4.670
	p-value = 0.000	p-value = 0.000
Distribution 4	t = -6.330	t = -5.931
	p-value = 0.000	p-value = 0.000

The results in Tables 5 and 6 show a negative correlation between the number of criteria and the reduction in the number of questions, which was expected according to Research Hypothesis 4. The correlation is stronger for distributions 3 and 4, showing that when the weights pattern is more skewed, the effect on the number of criteria in RP becomes more evident. However, the p-values in Table 6 indicate that, even for distributions 1 and 2, the correlation between NC and RP is still significant, when considering a significance level of 2%, for instance.

This result is aligned with expectations once as the number of criteria increases, the dimensionality of the space where the weight vectors are searched also becomes greater, resulting in more possible solutions, and requiring more information to limit the region that contains the DM's real aspirations.

5.5 The effect of the weight distribution over the reduction of the NQ

To verify Research Hypothesis 5, we can go back to the graphics in Figure 4. The weight distribution seems to have a significant effect on the reduction of the number of questions promoted by the HE. As previously mentioned, we can see that for distributions 3 and 4 (more skewed pattern), the percentage of reduction in the number of questions is greater than for distributions 1 and 2 (even though the spread between the two different tasks effect is less significative), which corroborates with research hypothesis 5. As previously discussed, the distribution of the weights is here evaluated based on the ratio between the highest and smallest scaling constants, so the results show that as this ratio becomes greater, the HE effects also increase. Both HE procedures are affected in a similar way by variations in the weight distribution.

This finding is rather relevant to practitioners once the performance of the FITradeoff elicitation procedure was found to be less efficient as the weight distribution becomes more skewed (Mendes *et al*, 2020), in this sense, the HE may be used as a complementary tool that enhances the method efficiency in these cases.

5.6 Discussion of results

Based on the results obtained through the simulation analysis, it was possible to see that the combination of preference modeling paradigms in FITradeoff does indeed enhance the method efficiency, in the sense that the total number of questions needed to find a recommendation is reduced when introducing holistic evaluation information into the model, as we could see by confirming Research Hypothesis 1. In addition, by confirming Research Hypothesis 2, it was possible to conclude the effect on this reduction is even more significant when an HE selection task is performed, when compared to the elimination task. However, the contribution of the HE to the reduction in the number of questions also depends on other variables, such as number of criteria and weights distribution pattern, as could be analyzed by confirming Research Hypotheses 4 and 5. Research Hypothesis 3 was not validated, in the sense that the number of alternatives of the MCDM problem does not have an influence on the effects of HE in the total number of questions.

These results also make it possible to draw some recommendations for the practice of decision-making with FITradeoff, especially for the analyst, when guiding the DM in the decision process. The importance of the role of the analyst in this process is emphasized by De Almeida et al (2021), in a sense that the question-answering process faced by the DM throughout the DSS with the possibility of switching between preference modeling paradigms should always be guided by an analyst with a good background on the method and its features.

Although the HE is recommended in several scenarios, the analyst could incentivize the DM, especially in cases where it is expected to observe a more skewed weight distribution. The FITradeoff method's first question compares the bigger and the smaller scaling constants, in case the DM prefers A, it is known that the ratio between them equals at least 2 (De Almeida *et al.*, 2016). In that sense, a possibility that arises would be displaying the information on the maximum ratio between them during the elicitation procedure. Another possibility would be enhancing heuristics to better explore this aspect.

It is expected that as the number of alternatives evaluated in the HE increases, the complexity and effort necessary to conduct such an analysis also increases, in this way, the analysts should consider that increasing the number of alternatives evaluated while performing the elimination procedure will have minor effects. So, they could advise the DM to look for a smaller group of alternatives that could provide more relevant information to the model (i.e., a group of alternatives that seems to be more in line with DM's aspirations, rather than a group of several ones that are too easily evaluated) or recommending the use of the selection procedure.

Performing HE in scenarios with many criteria, as suggested by different authors, should be a tough task, in these cases, if the DM feels confident to do so, the model presented in this work may support in the evaluation of the expected benefit of doing so. As the ratio between the scaling constants is not known, this information may be the estimate used in the FITradeoff LPP.

Moreover, the analyst may encourage the DM to perform the selection procedure especially when there are fewer POAs, given that it can more significantly shorten the process than elimination would do. Finally, based on the results, it would be recommended to perform a holistic elimination pre-analysis as soon as the elicitation begins once it could be an opportunity to introduce additional information to the model.

It is worth mentioning that when an alternative is said to be the worst from a subset, all the alternatives remaining in the analyzed group have their overall values raised to some point that is greater than the overall value of the eliminated one, on the other hand, if that alternative is eliminated by the answers given in the elicitation by decomposition, not necessarily all the alternatives from the group is supposed to dominate the eliminated one for all the feasible weight vectors. Considering the information that the DM can more rapidly give a correct answer in the elimination procedure (Pessoa, Roselli & De Almeida, 2022) that the selection one and that the dimensionality of POAs set do not significantly affect elimination treatment, performing it as the problem is initiated, can be a great opportunity to include relevant information into the model and reduce the number of questions necessary to solve it.

6 CONCLUSIONS

This work aimed at verifying the effectiveness of the Holistic Evaluation in terms of diminishing the number of questions that decision-makers are asked to answer to solve choice problems in the FITradeoff method. To investigate the effect promoted by the two procedures of HE, a simulation-based study was conducted, in which decision problems were randomly generated under several conditions (number of criteria and alternatives) and solved by simulating the behaviour and preferences of different DMs.

A very representative set of observations could be obtained, and then used to test empirically the built research hypotheses. Based on the results obtained, it was possible to conclude that performing HEs does indeed accelerate the obtention of the results in FITradeoff. Both selection and elimination procedures have a significant potential to reduce NQ, making the former even more powerful. We could see that the selection procedure, on average, tends to contribute more to speeding the convergence of the results. Moreover, it could be seen that other variables such as number of criteria and distribution of the weights have an influence on the magnitude of the effects of holistic evaluations in the reduction of the number of questions.

With the results obtained in this paper, the empirical assumptions raised by De Almeida et al (2021) on the benefits of the combination of preference modeling paradigms in the FITradeoff process were confirmed. Indeed, we could see an increase in the efficiency of the method in terms of less effort required to find a solution, when combining the information obtained in the elicitation by decomposition with holistic judgments. Either in selection or elimination tasks, the information obtained by holistic evaluation is quite strong in the sense that the inequalities incorporated into the model involve the scaling constants of all criteria at the same time; hence, the modification in the space of weights is expected to be significant. Future works should deeply analyze analytical aspects of the space of weights and how it is affected by the information obtained through ED and HE.

The implications of the results obtained for the practice of decision-making with the FITradeoff were also discussed in this paper. The role of the analyst is fundamental to guide the decision process, to advise the DMs on how to take advantage of the flexibility features of the method, including the possibility of switching between the preference modeling paradigms according to the DM's wishes and cognitive style. The benefits of HE confirmed through the research hypotheses tested in this paper enhance even more the potential of the practical applicability of the method, since decision-makers are always looking for ways to save time and effort when making decisions.

It is worth mentioning that the findings of this study are only applicable to the FITradeoff for choice problematic. The effect of performing HEs in different decision problematics such as ranking, sorting, and portfolio should be explored in future studies, with a proper adaptation of the simulation design and variables investigated. In addition, this study considered only the effect of performing one holistic evaluation at the very beginning of the elicitation procedure.

In future research, it would be interesting to conduct studies where multiple HEs could be conducted along the elicitation, so it would be possible to determine the effect of performing multiple assessments. Even though De Almeida et al (2021) raise the possibility to switch between elicitation by decomposition and holistic evaluations, and possibly conducting multiple holistic evaluations during the process, the present study did not address how such multiple evaluations may affect the efficiency of the method; instead, we have analysed a single holistic judgment made right after the ranking of criteria weights step. Hence, future studies should also investigate that.

Also, it would be relevant to consider a measure of distance between the alternatives that could represent how difficult would it be to perform such an analysis and could be used to weigh the effect of the treatment. In addition, investigating contributions of other factors in the analysis, beyond the DSS solely, such as the role of the analyst, is a good perspective for future studies.

Acknowledgements

The authors are grateful to CNPq (Brazilian Research Council) (grant numbers 312695/2020-9 and 311886/2022-1), FACEPE (Foundation for Science and Technology of the State of Pernambuco) and CAPES (the Brazilian Coordination Unit for Improving the Qualifications of Higher Education Faculty) (grant number 88887.603838/2021-00) for the partial financial support to this research.

References

BELTON V & STEWART T. 2002. *Multiple criteria decision analysis: An integrated approach*. Berlin: Springer.

CAMILO DGG, SOUZA RP, FRAZÃO TC & COSTA JUNIOR JF. 2020-02-21. Multi-criteria analysis in the health area: selection of the most appropriate triage system for the emergency care units in natal. In: *Bmc Medical Informatics And Decision Making*, vol. 20. Media LLC: Springer Science and Business. Available at: http://dx.doi.org/10.1186/s12911-020-1054-y.

CARVALHO RCLC, ROSELLI LRP & FIGUEIRA JR. 2023. Assigning priorities for raw material of a large pet food producer in the context of supply disruption. *Pesquisa Operacional*, **43**: 1–15. Available at: http://dx.doi.org/10.1590/0101-7438.2023.043spe1.00263605.

CONSUL PC. 1989. *Generalized Poisson Distribution: Properties and Applications. Statistics, text book and monograms*. New York: Marcel Dekker.

CORREIA LMAM, SILVA JMN, LEITE WKS, LUCAS REC & COLAÇO GA. 2021. A multicriteria decision model to rank workstations in a footwear industry based on a FITradeoff-ranking method for ergonomics interventions. *Operational Research*, p. 3335–3371. Available at: http://dx.doi.org/10.1007/s12351-021-00671-0.

CUNHA CPCB, MOTA DE MIRANDA MC, DE ALMEIDA AT, FREJ EA & ROSELLI LRP. 2020. Applying the FITradeoff method for aiding prioritization of special operations of Brazilian federal police. In: *Innovation for Systems Information and Decision: Second International Meeting, INSID 2020*, vol. Proceedings 2. p. 110–125. Recife, Brazil: Springer International Publishing.

CZEKAJSKI M, WACHOWICZ T & FREJ EA. 2023. Exploring the combination of holistic evaluation and elicitation by decomposition in FITRadeoff: prioritizing cultural tourism products in Poland. *Pesquisa Operacional*, **43**. Available at: http://dx.doi.org/10.1590/0101-7438.2023.043spe1.00263454.

DA SILVA LBL, FREJ EA, DE ALMEIDA AT, FERREIRA RJP & MORAIS DC. 2022. A review of partial information in additive multicriteria methods. *IMA Journal of Management Mathematics*, .

DE ALMEIDA AT. 2013. *Processo de Decisão nas Organizações: Construindo Modelos de Decisão Multicritério*. São Paulo: Atlas.

DE ALMEIDA AT, ALMEIDA JA, COSTA APCS & ALMEIDA-FILHO AT. 2016. A new method for elicitation of criteria weights in additive models: Flexible and interactive tradeoff. *European Journal of Operational Research*, **250**: 179–191.

DE ALMEIDA AT, CAVALCANTE CAV, ALENCAR MH, FERREIRA RJP, DE ALMEIDA-FILHO AT & GARCEZ TV. 2015. *Multicriteria and Multiobjective Models for Risk, Reliability and*

Maintenance Decision Analysis. vol. 231 of *International Series in Operations Research & Management Science*. New York: Springer.

De Almeida AT, Frej EA & Roselli LRP. 2021. Combining holistic and decomposition paradigms in preference modeling with the flexibility of FITradeoff. *Central European Journal of Operations Research*, **29**: 7–47.

De Almeida AT, Frej EA, Roselli LRP & Costa APCS. 2023. A summary on FITradeoff method with methodological and practical developments and future perspectives. *Pesquisa Operacional*, **43**.

De Araujo YPG, Frej E, Roselli LR & Machado AJ. 2022. Uso do método FITradeoff para seleção de fornecedores em um processo licitatório com base na modalidade pregão. In: *Anais Do SimpÓsio Brasileiro De Pesquisa Operacional, 2022, Juiz de Fora. Anais eletrônicos*. Campinas.

De Macedo PP, Mota CMM & Sola AVH. 2018. Meeting the Brazilian Energy Efficiency Law: a flexible and interactive multicriteria proposal to replace non-efficient motors. *Sustainable Cities And Society*, p. 822–832. Available at: http://dx.doi.org/10.1016/j.scs.2018.06.020.

Deb K & Jain H. 2014. An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: solving problems with box constraints. In: *IEEE Transactions On Evolutionary Computation*, vol. 18. p. 577–601. ago: Institute of Electrical and Electronics Engineers (IEEE). Available at: http://dx.doi.org/10.1109/tevc.2013.2281535.

Dell'Ovo M, Oppio A & Capolongo S. 2020. Approaching the Location of Healthcare Facilities: how to model the decision problem. Decision Support System For The Location Of Healthcare Facilities. Available at: http://dx.doi.org/10.1007/978-3-030-50173-0_3.

Dos Santos AL, Santos AFA, Assis AG, Costa Júnior JF & Souza RP. 2022-05-10. Model to support intervention prioritization for the control of Aedes aegypti in Brazil: a multi-criteria approach. *Bmc Public Health*, **22**: 932. Available at: http://dx.doi.org/10.1186/s12889-022-13006-1.

Dutta B, Dao SD, Martínez L & Goh M. 2021. An evolutionary strategic weight manipulation approach for multi-attribute decision making: topsis method. *International Journal Of Approximate Reasoning*, p. 64–83.

Ferrari S & Cribari-Neto F. 2004. Beta Regression for Modelling Rates and Proportions. *Journal Of Applied Statistics*, p. 799–815. Available at: http://dx.doi.org/10.1080/0266476042000214501.

Fossile DK, Frej EA, Costa SEG, Lima EP & Almeida AT. 2020-07. Selecting the most viable renewable energy source for Brazilian ports using the FITradeoff method. *Journal of Cleaner Production*, p. 121107. Available at: http://dx.doi.org/10.1016/j.jclepro.2020.121107.

FREJ EA, DE ALMEIDA AT & COSTA APCS. 2019. Using data visualization for ranking alternatives with partial information and interactive tradeoff elicitation. *Operational Research*, **39**.

FREJ EA, EKEL P & DE ALMEIDA AT. 2021-02. A benefit-to-cost ratio based approach for portfolio selection under multiple criteria with incomplete preference information. *Information Sciences*, p. 487–498. Available at: http://dx.doi.org/10.1016/j.ins.2020.08.119.

HINES WW & MONTGOMERY DC. 1990. *Probability and Statistics in Engineering and Management Science*. New York, Wiley.

HWANG CL & YOON K. 1981. *Multiple attribute decision making: methods and applications*. Heidelberg: Springer.

KANG THA, FREJ EA & DE ALMEIDA AT. 2020. Flexible and interactive tradeoff elicitation for multicriteria sorting problems. *Asia-Pacific Journal Of Operational Research*, p. 2050020. 17 set.

KEENEY RL. 1992. *Value-focused thinking: A Path to Creative Decisionmaking*. Harvard University Press.

KEENEY RL & RAIFFA H. 1976. *Decision analysis with multiple conflicting objectives*. New York: Wiley & Sons.

LIU D, LI T & LIANG D. 2020. An integrated approach towards modeling ranked weights. *Computers & Industrial Engineering*, p. 106629–106645.

LUGO SDR, DU B, ALMEIDA JA & NISHINO N. 2023. A circular food economy multicriteria decision problem based on the FITradeoff method. *Pesquisa Operacional*, **43**. Available at: http://dx.doi.org/10.1590/0101-7438.2023.043spe1.00263528.

MARQUES AC, FREJ EA & DE ALMEIDA AT. 2022. Multicriteria decision support for project portfolio selection with the FITradeoff method. *set*, **111**: 102661–102673.

MENDES JAJ, FREJ EA, DE ALMEIDA AT & ALMEIDA JA. 2020. Evaluation of flexible and interactive tradeoff method based on numerical simulation experiments. *Pesquisa Operacional*, **40**: 1–25.

MONTE MB & MORAIS DC. 2019-11. A Decision Model for Identifying and Solving Problems in an Urban Water Supply System. *Water Resources Management*, **33**: 4835–4848. Available at: http://dx.doi.org/10.1007/s11269-019-02401-w.

OLIVEIRA RSM, MORAIS DC & SIEBERT J. 2023. Developing and evaluating new alternatives for urban mobility: a case study of a Brazilian city. *Pesquisa Operacional*, **43**: 1–15. Available at: http://dx.doi.org/10.1590/0101-7438.2023.043spe1.00263440.

PESSOA MEBT, ROSELLI LRP & DE ALMEIDA AT. 2021. Neuroscience Experiment to improve the new feature of the FITradeoff Method. In: *ICDSST, Decision Support Systems, Analytics and Technologies in Response to Global Crisis Management*.

PESSOA MEBT, ROSELLI LRP & DE ALMEIDA AT. 2022. Using the FITradeoff Decision Support System to Support a Brazilian Compliance Organization Program. In: *Information Systems Frontiers*. p. 1–15. Media LLC: Springer Science and Business. Available at: http://dx.doi.org/10.1007/s10796-022-10290-5.

RIBEIRO MLS, ROSELLI LRP, FREJ EA, DE ALMEIDA AT & MORAIS DC. 2021. Using the FITradeoff method to solve a shopping mall location problem in the northeastern countryside of Brazil. *Control And Cybernetics*, p. 109–126. Available at: http://dx.doi.org/10.2478/candc-2021-0007.

RODRIGUES LVS, CASADO RSGR, CARVALHO EN, SILVA MM & SILVA LC. 2020. Using FITradeoff in a ranking problem for supplier selection under TBL performance evaluation: an application in the textile sector. *Production*, Available at: http://dx.doi.org/10.1590/0103-6513.20190032.

RODRÍGUEZ JMM, FREJ EA, KANG THA & DE ALMEIDA AT. 2023. Outsourcing laboratory services from a Colombian agricultural research company using the FITradeoff method under multiple stakeholders analysis. *Pesquisa Operacional*, **43**. Available at: http://dx.doi.org/10.1590/0101-7438.2023.043spe1.00258518.

ROSELLI LRP & DE ALMEIDA AT. 2021. The use of the success-based decision rule to support the holistic evaluation process in FITradeoff. *International Transactions in Operational Research*, .

SANTO PPPE, FREJ EA & ALMEIDA ATD. 2021. Improving the Elicitation Process for Intracriterion Evaluation in the FITradeoff Method. In: *Innovation for Systems Information and Decision. Lecture Notes in Business Information Processing*, vol. 435. Cham: Springer.

SANTOS AG, PESSÔA LAM, MOTA CMM & FREJ EA. 2023. A FITradeoff-based approach for strategic decisions on military budget. *Pesquisa Operacional*, **43**. Available at: http://dx.doi.org/10.1590/0101-7438.2023.043spe1.00262789.

SEWARD LE & DOANE DP. 2014. *Estatística Aplicada à Administração e Economia-4*. AMGH Editora.

SILVA WDO & MORAIS DC. 2022. Impacts and insights of circular business models' outsourcing decisions on textile and fashion waste management: a multi-criteria decision model for sorting circular strategies. *Journal Of Cleaner Production*, p. 133551.

SMITHSON M & VERKUILEN J. 2006-03. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables. *Psychological Methods*, p. 54–71. Available at: http://dx.doi.org/10.1037/1082-989x.11.1.54. american Psychological Association (APA).

WEBER M. 1987. Decision making with incomplete information. *European Journal of Operational Research*, **28**: 44–57.

ZANAZZI JF, ZANAZZI JL & PONTELLI D. 2023. Prioritization of improvement actions in industrial production: application of the FITradeoff method to order improvement actions identified through the failures modes and effects analysis (fmea). *Pesquisa Operacional*, **43**. Available at: http://dx.doi.org/10.1590/0101-7438.2023.043spe1.00263696.

How to cite

RIBEIRO MLS, VASCONCELOS TRS, FREJ EA & DE ALMEIDA AT. 2024. Effects of holistic evaluations on FITradeoff method using a simulation study. *Pesquisa Operacional*, **44**: e283585. doi: 10.1590/0101-7438.2023.043.00283585.