International Journal of Approximate Reasoning 161 (2023) 109004

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

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

Using scoring functions in a group decision-making procedure with heterogeneous experts and qualitative assessments

José Luis García-Lapresta [a,*], Raquel González del Pozo [b]

[a] IMUVA, PRESAD Research Group, Departamento de Economía Aplicada, Universidad de Valladolid, Spain
[b] IMUVA, PRESAD Research Group, Departamento de Economía Financiera, Actuarial y Estadística, Universidad Complutense de Madrid, Spain

ARTICLE INFO

Article history:
Received 8 February 2023
Received in revised form 10 July 2023
Accepted 27 July 2023
Available online 5 August 2023

Keywords:
Group decision-making
Qualitative scales
Ordinal proximity measures
Scores
Expertise
Aggregation

ABSTRACT

In this paper, we propose a group decision-making procedure to rank alternatives in the context of ordered qualitative scales that not necessarily are uniform (the proximities between consecutive terms of the scales can be perceived as different). The procedure manages two ordered qualitative scales. One of them is used to determine the weights of the experts according to their expertise, taking into account the assessments given by a decision-maker to the experts. And another one, that is used by the experts to assess the alternatives. In order to assign numerical scores to the linguistic terms of the ordered qualitative scales, we have introduced and analyzed some scoring functions. They are based on the concept of ordinal proximity measure that properly represents the ordinal proximities between the linguistic terms of the ordered qualitative scales.

1. Introduction

In many decision making situations characterized by subjectivity and imprecision, it seems reasonable that individuals express their judgments, preferences and opinions by means of words or linguistic expressions (see [53–55]).

In the literature, many different approaches and models have been proposed to handle linguistic information. Among the proposed procedures we can highlight: the fuzzy models associated with fuzzy sets and hesitant fuzzy linguistic information (see [34], [37], among others), linguistic procedures based on the 2-tuple model (see [38], [45], among others) and linguistic computation models (see [58], [59], among others).

Likewise, many group decision making problems involve the use of linguistic information collected by questionnaires based on ordered qualitative scales. Within this context, several authors have proposed many procedures to analyze and manage categorical and ordinal data. For example, McCullagh [39] proposes the use of cumulative models for ordinal data, Franceschini et al. [11] introduce a new dispersion measure called the ordinal range to handle ordered qualitative scales and Gadrich & Bashkansky [14] analyze the dispersion of ordinal data by means of the statistical tool: ORDANOVA (Ordinal data analysis of variation). Nonetheless, and although the proposed procedures try to respect and preserve qualitative information as much as possible, the assignment of numerical values or scores to each ordinal response category or linguistic term is the most popular method to manage linguistic information. In this way, Labovitz [33] analyzes some advantages of using scores to represent the terms of ordered qualitative scales.

* Corresponding author.
E-mail addresses: lapresta@uva.es (J.L. García-Lapresta), raqueg18@ucm.es (R. González del Pozo).

https://doi.org/10.1016/j.ijar.2023.109004

These numerical values or scores are used to collect information about a data set, describe and calculate statistical indicators and perform statistical tests. Nevertheless, this conversion of qualitative information into quantitative must be carried out in an appropriate way, since it can lead to misunderstanding and misinterpretation of the obtained results depending on the scores assigned (see [40], [11], [15], among others).

On the other hand, in those decision-making procedures that employ information from ordered qualitative scales, it is important to emphasize the role that plays the agents' perceptions about the scales, since sometimes these scales can be considered as non-uniform, in the sense that agents may perceive different proximities between consecutive terms of the scale. For instance, in the framework of health-care and medicine, the ordered qualitative scale {poor, fair, good, very good, excellent}, used by patients to evaluate self-rated health (see [10]), it could be considered as non-uniform if 'fair' is perceived closer to 'good' than to 'poor', or if 'good' is perceived closer to 'very good' than to 'fair', or if 'very good' is perceived closer to 'good' than to 'excellent', etc.

In recent years, there is a considerable body of literature that addresses the issue of handling linguistic terms coming from ordered qualitative scales which are not uniformly and symmetrically distributed. These types of scales has been analyzed by Herrera et al. [27] to propose a representation model for unbalanced linguistic information based on the concept of linguistic hierarchy and to develop some consensus models within the issues related to group decision-making (see [7], [57], among others).

In the context of non-uniform ordered qualitative scales, García-Lapresta & Pérez-Román [21] introduce the notion of ordinal proximity measure to represent the information about how agents perceive the proximities between the terms of the scales by means of ordinal degrees of proximity. Ordinal proximity measures have been also implemented in some decision-making procedures to evaluate a set of alternatives by means of non-necessarily uniform ordered qualitative scales (see [23], [16], [24], [20], [19]).

Since, as previously mentioned, sometimes the meaning or the interpretation of linguistic terms may lead to some ordered qualitative scales considered as non-uniform, it may seem unreasonable to assign equidistant numerical values to consecutive linguistic terms (see, for instance, [33], [8]). For this reason, in this paper we propose and analyze several scoring functions which can be applied to decision-making situations in which the linguistic assessments are given through non-uniform ordered qualitative scales. To determine the scores of the linguistic terms, the proposed scoring functions are based on the notion of ordinal proximity measure introduced by García-Lapresta & Pérez-Román [21].

Although in previous papers (some of them mentioned above) we have used ordinal proximity measures to devise and implement several decision-making procedures in a pure ordinal fashion, we realize that they require a good knowledge of the notions and techniques developed in those papers.

In the present contribution, we propose a cardinal approach based on the ordinal perceptions on the proximities between the terms of the scales by means of appropriate scoring functions. The reason is to facilitate the users that are not familiar with ordinal proximity measures the implementation of more simple decision-making procedures, taking into account the subjective perceptions about the closeness between the terms of the scales they use in their questionnaires.

Once these perceptions are known, the proposed scoring functions assign a numerical value to each term of the scale, preserving the ordinal information as much as possible, and then the usual statistical procedures can be used.

The main novelty of this proposal is that the numerical values assigned to the terms of the scales are not arbitrary, but based on ordinal perceptions.

Taking into account these scoring functions, we propose a group decision-making procedure where a group of experts evaluates a set of alternatives through an ordered qualitative scale with the purpose of generate a ranking on the set of alternatives. The procedure manages two ordered qualitative scales: one of them is applied to determine the experts' weights and the other one is used by the experts to evaluate the alternatives.

In the framework of group-decision making, much of the literature considers that the weight associated with each expert is determined by means of hierarchical analysis or mathematical techniques such as the AHP procedure or fuzzy methods (see [44], [5], [36], among others).

Assigning different weights to the experts' opinions is usual when the decision-maker is aware of the experts' heterogeneity, because they have different knowledge and expertise on the issue and the alternatives under evaluation.

In our proposal the importance of the experts is established taking into account the qualitative assessments given by a decision-maker to each expert.[1]

Due to the fact that the group decision-making procedure uses two ordered qualitative scales (which can be considered as non-uniform), each scale is equipped with a metrizable ordinal proximity measure (see [17]) that collects the perceptions about the scale and assigns a numerical value to each linguistic term by means of a scoring function. Finally, the proposed procedure aggregates the weights associated with each expert and the scores assigned to each alternative in order to rank the alternatives according to their global score.

The rest of the paper is organized as follows. Section 2 recalls some important notions related to ordinal proximity measures. Section 3 presents several scoring functions which are applied in the proposed decision-making procedure. Section 4 discusses the main features, pros and cons of our proposal. Section 5 introduces the group decision-making procedure.

[1] This idea was already considered by [12, Subsect. 3.3], where one of the paradigms to determine the experts' importance is based on ordinal proximity measures.

Section 6 includes a illustrative case study. Section 7 shows how to extend the procedure of Section 5 to multiple criteria. Finally, Section 8 contains some concluding remarks.

2. Preliminaries

Let us consider an ordered qualitative scale (**OQS**) $\mathcal{L} = \{l_1, \dots, l_g\}$, with $g \geq 2$ and $l_1 \prec \cdots \prec l_g$.

To measure the ordinal proximities between linguistic terms of an OQS we use the notion of ordinal proximity measure (**OPM**) introduced by García-Lapresta & Pérez-Román [21].

An OPM is a mapping that assigns an ordinal degree of proximity to each pair of linguistic terms of an OQS $\mathcal{L}$. The ordinal degrees of proximity belong to a linear order $\Delta = \{\delta_1, \dots, \delta_h\}$, with $\delta_1 \succ \cdots \succ \delta_h$. It is important noticing that the members of Δ are not numbers, but ordinal degrees: δ_1 represents the maximum proximity, δ_2 the second degree of proximity, etc., and δ_h the minimum proximity.

Definition 1. ([21]) An *ordinal proximity measure on $\mathcal{L}$ with values in Δ* is a mapping $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$, where $\pi(l_r, l_s) = \pi_{rs}$ represents the degree of proximity between l_r and l_s, satisfying the following conditions:

1. *Exhaustiveness*: For every $\delta \in \Delta$, there exist $l_r, l_s \in \mathcal{L}$ such that $\delta = \pi_{rs}$.
2. *Symmetry*: $\pi_{sr} = \pi_{rs}$, for all $r, s \in \{1, \dots, g\}$.
3. *Maximum proximity*: $\pi_{rs} = \delta_1 \Leftrightarrow r = s$, for all $r, s \in \{1, \dots, g\}$.
4. *Monotonicity*: $\pi_{rs} \succ \pi_{rt}$ and $\pi_{st} \succ \pi_{rt}$, for all $r, s, t \in \{1, \dots, g\}$ such that $r < s < t$.

Every OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$ can be represented by a $g \times g$ symmetric matrix with coefficients in Δ, where the elements in the main diagonal are $\pi_{rr} = \delta_1$, $r = 1, \dots, g$:

$$\begin{pmatrix} \pi_{11} & \cdots & \pi_{1s} & \cdots & \pi_{1g} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \pi_{r1} & \cdots & \pi_{rs} & \cdots & \pi_{rg} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \pi_{g1} & \cdots & \pi_{gs} & \cdots & \pi_{gg} \end{pmatrix}.$$

The above matrix is called *proximity matrix associated with π*.

In García-Lapresta et al. [17] is introduced a prominent class of OPMs: the metrizable OPMs which are based on linear metrics on OQSs.

Definition 2. ([17]) A *linear metric* on $\mathcal{L}$ is a mapping $d : \mathcal{L} \times \mathcal{L} \longrightarrow \mathbb{R}$ satisfying the following conditions for all $r, s, t \in \{1, \dots, g\}$:

1. *Positiveness*: $d(l_r, l_s) \geq 0$.
2. *Identity of indiscernibles*: $d(l_r, l_s) = 0 \Leftrightarrow l_r = l_s$.
3. *Symmetry*: $d(l_s, l_r) = d(l_r, l_s)$.
4. *Linearity*: $d(l_r, l_t) = d(l_r, l_s) + d(l_s, l_t)$ whenever $r < s < t$.

Definition 3. ([17]) An OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$ is *metrizable* if there exists a linear metric $d : \mathcal{L} \times \mathcal{L} \longrightarrow \mathbb{R}$ such that $\pi_{rs} \succ \pi_{tu} \Leftrightarrow d(l_r, l_s) < d(l_t, l_u)$, for all $r, s, t, u \in \{1, \dots, g\}$. We say that π is generated by d.

We note that for $g = 2$ the only OPM that exists is metrizable; for $g = 3$ there are 3 OPMs, and all of them are metrizable; for $g = 4$ there are 51 OPMs, but only 25 of them are metrizable (see [17]). In turn, for $g = 5$ there are 716 metrizable OPMs; and for $g = 6$ there are 18,262 metrizable OPMs.

The metrizable OPMS for $g = 2, 3, 4$ are collected in the Annex. The subindices of the matrices A's correspond to the subindices of the δ's which appear just over the main diagonal, $\pi_{12}, \pi_{23}, \dots, \pi_{(g-1)\,g}$.

Definition 4. ([17]) An OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$ is *uniform* if $\pi_{r\,(r+1)} = \pi_{s\,(s+1)}$ for all $r, s \in \{1, \dots, g-1\}$, and *totally uniform* if $\pi_{r\,(r+t)} = \pi_{s\,(s+t)}$ for all $r, s, t \in \{1, \dots, g-1\}$ such that $r + t \leq g$ and $s + t \leq g$.

We note that if $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$ is a uniform metrizable OPM on $\mathcal{L}$, then it is also totally uniform (see [17, Prop. 3]). Additionally, for each OQS $\mathcal{L}$, there exists one and only one totally uniform OPM on $\mathcal{L}$ (see [17, Remark 6]).

Remark 1. If $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$ is the totally uniform OPM on $\mathcal{L}$, then $h = g$ and $\pi_{r\,(r+s)} = \delta_{s+1}$ for all $r, s \in \{1, \dots, g-1\}$ such that $r + s \leq g$. In particular, we have $\pi_{r\,(r+1)} = \delta_2$, $\pi_{r1} = \delta_r$ and $\pi_{rg} = \delta_{g-r+1}$, for every $r \in \{1, \dots, g-1\}$.

3. Scoring functions

The assignment of numerical scores to the linguistic terms of the scales is one of the most popular method to manage data coming from OQSs. Nevertheless, the main problem of this method is the subjectivity in which numerical values represent qualitative information (see [46]).

On the other hand, if the linguistic terms of an OQS are labeled with numbers, individuals' perceptions about the proximities between the terms can be influenced, even misleaded, by this twofold semantics. If these numerical values are consecutive integer numbers, then individuals can believe that the terms of the scale are equispaced (see [11]).

In addition, they may cause a misinterpretation of the obtained results depending on how the linguistic terms are coded (see [47], [11], among others).

Some of these drawbacks are presented and discussed by Franceschini et al. [11] through an example in which two different numerical conversions are applied to the OQS {reject, poor quality, medium quality, good quality, excellent quality}: {1, 2, 3, 4, 5} and {1, 3, 9, 27, 81}.

To overcome these problems, this section is addressed to determine numerical scores by means of several scoring functions based on metrizable OPMs. These scoring functions consider how individuals perceive the proximities between linguistic terms and generate scores accordingly.

To do that, first we introduce the notion of scoring function on an OQS. It assigns a score to each linguistic term of an OQS satisfying two simple conditions: the higher term, the higher score; and in totally uniform OPMs the scores are equidistant.

Definition 5. Given an OQS $\mathcal{L} = \{l_1, \ldots, l_g\}$, a *scoring function on* $\mathcal{L}$ is a function $S : \mathcal{L} \longrightarrow \mathbb{R}$ satisfying the following conditions for all $r, s \in \{1, \ldots, g\}$:

1. $S(l_r) < S(l_s) \;\Leftrightarrow\; r < s$.
2. If π is the totally uniform OPM on $\mathcal{L}$, then there exists $d > 0$ such that $S(l_r) = S(l_1) + (r-1) \cdot d$.

Remark 2. Many organizations and data providers implicitly aggregate qualitative information by means of grouping the two highest linguistic terms of the scales. This is the case of some reports published by EUROSTAT, the statistical office of the European Union,[2] the Asian Barometer Survey, institution that collects public opinion across Asia on issues such as political values, democracy, and governance,[3] the Arab Barometer, the central resource for quantitative research on the Middle East,[4] the Sociological Research Center of Spain (CIS in Spanish)[5] or the Pew Research Center of United States,[6] as well as some companies specialized in market and consumer data such as Statista,[7] Nielsen[8] or IPSOS,[9] among many others. It is important to note that this practice does not fulfill the first condition of scoring functions: it is equivalent to allocate a null score to the terms $l_1, \ldots, l_{g-2}$ and one point to the terms l_{g-1} and l_g.

We now introduce two normalizations of scoring functions that are also scoring functions (see [32] for different normalization procedures).

Proposition 1. *If $\mathcal{L} = \{l_1, \ldots, l_g\}$ is an OQS and $S : \mathcal{L} \longrightarrow \mathbb{R}$ is a scoring function on $\mathcal{L}$ such that $S(l_g) > 0$, then the function $S' : \mathcal{L} \longrightarrow \mathbb{R}$ defined as*

$$S'(l_r) = \frac{S(l_r)}{S(l_g)}$$

is also a scoring function on $\mathcal{L}$.

Proof. Since $S(l_g) > 0$, we have $S'(l_r) < S'(l_s) \;\Leftrightarrow\; S(l_r) < S(l_s) \;\Leftrightarrow\; r < s$. Thus, the first condition is satisfied.

Let π be the totally uniform OPM on $\mathcal{L}$. Since $S(l_r) = S(l_1) + (r-1) \cdot d$ for some $d > 0$, we have

$$S'(l_r) = \frac{S(l_1) + (r-1) \cdot d}{S(l_g)} = \frac{S(l_1)}{S(l_g)} + (r-1) \cdot \frac{d}{S(l_g)} = S'(l_1) + (r-1) \cdot \frac{d}{S(l_g)}.$$

Hence, the two conditions of scoring functions are satisfied. □

2 www.ec.europa.eu/eurostat.
3 www.asianbarometer.org.
4 www.arabbarometer.org.
5 http://www.cis.es.
6 www.pewresearch.org.
7 www.statista.com.
8 www.www.nielsen.com.
9 www.ipsos.com.

Note that $S'(l_g) = 1$ and $S'(l_r) > 0 \Leftrightarrow S(l_r) > 0$, for every $r \in \{1, \dots, g\}$.

Proposition 2. *If $\mathcal{L} = \{l_1, \dots, l_g\}$ is an OQS and $S : \mathcal{L} \longrightarrow \mathbb{R}$ is a scoring function on $\mathcal{L}$, then the function $S'' : \mathcal{L} \longrightarrow \mathbb{R}$ defined as*

$$S''(l_r) = \frac{S(l_r) - S(l_1)}{S(l_g) - S(l_1)}$$

is also a scoring function on $\mathcal{L}$.

Proof. Since $S(l_g) > S(l_1)$, we have

$$S''(l_r) < S''(l_s) \Leftrightarrow S(l_r) - S(l_1) < S(l_s) - S(l_1) \Leftrightarrow S(l_r) < S(l_s) \Leftrightarrow r < s.$$

Thus, the first condition is satisfied.

Let π be the totally uniform OPM on $\mathcal{L}$. Since $S(l_r) = S(l_1) + (r-1) \cdot d$ for some $d > 0$, we have

$$S''(l_r) = \frac{S(l_1) + (r-1) \cdot d - S(l_1)}{S(l_1) + (g-1) \cdot d - S(l_1)} = \frac{r-1}{g-1} = (r-1) \cdot \frac{1}{g-1}.$$

Hence, the two conditions of scoring functions are satisfied. □

Note that $S''(l_1) = 0$, $S''(l_g) = 1$ and, consequently, $S''(l_r) \in [0, 1]$ for every $r \in \{1, \dots, g\}$.

We now justify that any convex combination of two scoring functions on an OQS is also a scoring function.

Proposition 3. *If $\mathcal{L} = \{l_1, \dots, l_g\}$ is an OQS and $S_1 : \mathcal{L} \longrightarrow \mathbb{R}$ and $S_2 : \mathcal{L} \longrightarrow \mathbb{R}$ are two scoring functions on $\mathcal{L}$, then the function $S^\lambda : \mathcal{L} \longrightarrow \mathbb{R}$ defined as*

$$S^\lambda(l_r) = \lambda \cdot S_1(l_r) + (1-\lambda) \cdot S_2(l_r)$$

is also a scoring function on $\mathcal{L}$ for every $\lambda \in [0, 1]$.

Proof. We now prove the first condition.

$\Rightarrow$) Suppose $S^\lambda(l_r) < S^\lambda(l_s)$ and $r \geq s$.

If $r = s$, then $S^\lambda(l_r) = S^\lambda(l_s)$, contrary to the hypothesis.

If $r > s$, then $S_1(l_r) > S_1(l_s)$ and $S_2(l_r) > S_2(l_s)$. Hence, for every $\lambda \in (0, 1)$ we have $\lambda \cdot S_1(l_r) > \lambda \cdot S_1(l_s)$ and $(1-\lambda) \cdot S_2(l_r) > (1-\lambda) \cdot S_2(l_s)$. Thus, we have

$$S^\lambda(l_r) = \lambda \cdot S_1(l_r) + (1-\lambda) \cdot S_2(l_r) > \lambda \cdot S_1(l_s) + (1-\lambda) \cdot S_2(l_s) = S^\lambda(l_s),$$

contrary to the hypothesis.

If $\lambda = 0$, then $S^\lambda(l_r) = S_2(l_r) > S_2(l_s) = S^\lambda(l_s)$, contrary to the hypothesis.

If $\lambda = 1$, then $S^\lambda(l_r) = S_1(l_r) > S_1(l_s) = S^\lambda(l_s)$, contrary to the hypothesis.

$\Leftarrow$) If $r < s$, then $S_1(l_r) < S_1(l_s)$ and $S_2(l_r) < S_2(l_s)$. Hence, for every $\lambda \in (0, 1)$ we have $\lambda \cdot S_1(l_r) < \lambda \cdot S_1(l_s)$ and $(1-\lambda) \cdot S_2(l_r) < (1-\lambda) \cdot S_2(l_s)$. Thus, we have

$$S^\lambda(l_r) = \lambda \cdot S_1(l_r) + (1-\lambda) \cdot S_2(l_r) < \lambda \cdot S_1(l_s) + (1-\lambda) \cdot S_2(l_s) = S^\lambda(l_s).$$

If $\lambda = 0$, then $S^\lambda(l_r) = S_2(l_r) < S_2(l_s) = S^\lambda(l_s)$.

If $\lambda = 1$, then $S^\lambda(l_r) = S_1(l_r) < S_1(l_s) = S^\lambda(l_s)$.

Thus, the first condition is satisfied.

In order to prove the second condition, let π be the totally uniform OPM on $\mathcal{L}$. Since S_1 and S_2 are scoring functions on $\mathcal{L}$, there exist $d_1, d_2 > 0$ such that $S_1(l_r) = S_1(l_1) + (r-1) \cdot d_1$ and $S_2(l_r) = S_2(l_1) + (r-1) \cdot d_2$. Then,

$$\begin{aligned}
&S^\lambda(l_r) = \lambda \cdot S_1(l_r) + \lambda \cdot (r-1) \cdot d_1 + (1-\lambda) \cdot S_2(l_1) + (1-\lambda) \cdot (r-1) \cdot d_2 = \\
&\lambda \cdot S_1(l_1) + (1-\lambda) \cdot S_2(l_1) + (r-1) \cdot (\lambda \cdot d_1 + (1-\lambda) \cdot d_2) = \\
&S^\lambda(l_1) + (r-1) \cdot (\lambda \cdot d_1 + (1-\lambda) \cdot d_2).
\end{aligned}$$

Hence, the two conditions of scoring functions are satisfied. □

We now introduce four specific scoring functions. The first one, S_b, is based on the comparison between each linguistic term and the **b**est linguistic term, l_g (see Fig. 1). Under a different approach to the present paper, these comparisons are in the basis of the group decision-making procedure introduced and analyzed by García-Lapresta & Pérez-Román [23].

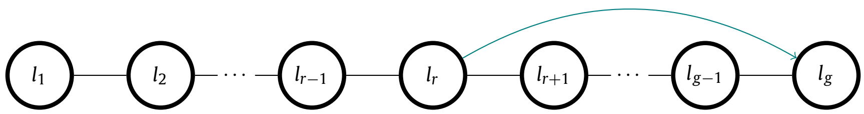

Fig. 1. Scoring function S_b.

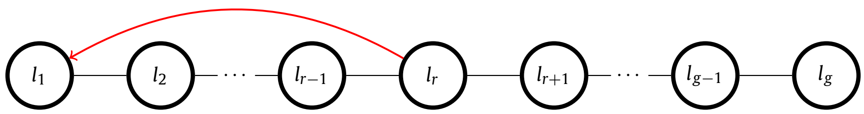

Fig. 2. Scoring function S_w.

Proposition 4. *If $\mathcal{L} = \{l_1, \dots, l_g\}$ is an OQS equipped with a metrizable OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$, then the function $S_b : \mathcal{L} \longrightarrow \mathbb{R}$ defined as*

$$S_b(l_r) = h - \rho(\pi_{rg}),$$

where $\rho(\delta_k) = k$, is a scoring function on $\mathcal{L}$.

Proof. Since π is monotonic, we have

$$r < s \;\Leftrightarrow\; \pi_{sg} \succ \pi_{rg} \;\Leftrightarrow\; \rho(\pi_{sg}) < \rho(\pi_{rg}).$$

Then,

$$S_b(l_r) < S_b(l_s) \;\Leftrightarrow\; h - \rho(\pi_{rg}) < h - \rho(\pi_{sg}) \;\Leftrightarrow\; \rho(\pi_{sg}) < \rho(\pi_{rg}) \;\Leftrightarrow\; r < s.$$

Thus, the first condition is satisfied.

If π is the totally uniform OPM on $\mathcal{L}$, taking into account Remark 1, we have

$$S_b(l_r) = h - \rho(\pi_{rg}) = \rho(\pi_{1g}) - \rho(\pi_{rg}) = g - (g - r + 1) = r - 1.$$

Hence, the two conditions of scoring functions are satisfied. $\square$

Note that $S_b(l_1) = 0$, $S_b(l_g) = h - 1$ and, consequently, $S_b(l_r) \in [0, h-1]$ for every $r \in \{1, \dots, g\}$.

The next scoring function, S_w, is based on the comparison between each linguistic term and the **w**orst linguistic term, l_1 (see Fig. 2).

Proposition 5. *If $\mathcal{L} = \{l_1, \dots, l_g\}$ is an OQS equipped with a metrizable OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$, then the function $S_w : \mathcal{L} \longrightarrow \mathbb{R}$ defined as*

$$S_w(l_r) = \rho(\pi_{r1}) - 1,$$

where $\rho(\delta_k) = k$, is a scoring function on $\mathcal{L}$.

Proof. Since π is monotonic, we have

$$r < s \;\Leftrightarrow\; \pi_{r1} \succ \pi_{s1} \;\Leftrightarrow\; \rho(\pi_{r1}) < \rho(\pi_{s1}).$$

Then, $S_w(l_r) < S_w(l_s) \Leftrightarrow r < s$ and, consequently, the first condition is satisfied.

If π is the totally uniform OPM on $\mathcal{L}$, taking into account Remark 1, we have

$$S_w(l_r) = \rho(\pi_{r1}) - 1 = r - 1.$$

Hence, the two conditions of scoring functions are satisfied. $\square$

Note that $S_w(l_1) = 0$, $S_w(l_g) = h - 1$ and, consequently, $S_w(l_r) \in [0, h-1]$ for every $r \in \{1, \dots, g\}$.

The next scoring function, S_{bw}, is based on the comparison between each linguistic term and the **b**est and **w**orst linguistic terms, l_g and l_1 (see Fig. 3). This approach is related to the TOPSIS method (see [31]) and the Best Worst Method (see [42]).

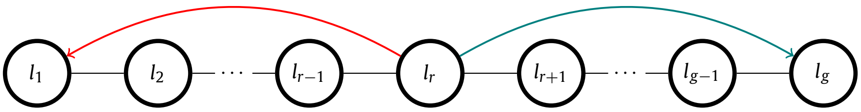

Fig. 3. Scoring function S_{bw}.

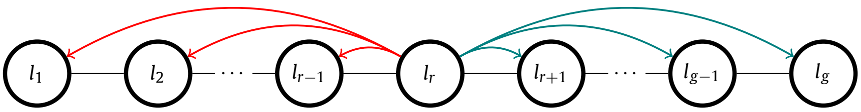

Fig. 4. Scoring function S_a.

Proposition 6. *If $\mathcal{L} = \{l_1, \dots, l_g\}$ is an OQS equipped with a metrizable OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$, then the function $S_{bw} : \mathcal{L} \longrightarrow \mathbb{R}$ defined as*

$$S_{bw}(l_r) = \frac{S_b(l_r) + S_w(l_r)}{2} = \frac{h + \rho(\pi_{r1}) - \rho(\pi_{rg}) - 1}{2},$$

where $\rho(\delta_k) = k$, is a scoring function on $\mathcal{L}$.

Proof. From Proposition 3, taking $\lambda = 0.5$ we obtain that S_{bw} is a scoring function on $\mathcal{L}$. $\square$

Note that $S_{bw}(l_1) = 0$, $S_{bw}(l_g) = h - 1$ and, consequently, $S_{bw}(l_r) \in [0, h-1]$ for every $r \in \{1, \dots, g\}$.

The next scoring function, S_a, is based on the comparison between each linguistic term and **a**ll linguistic terms (see Fig. 4).

Proposition 7. *If $\mathcal{L} = \{l_1, \dots, l_g\}$ is an OQS equipped with a metrizable OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$, then the function $S_a : \mathcal{L} \longrightarrow \mathbb{R}$ defined as*

$$S_a(l_r) = \frac{(g+2) \cdot (g-1)}{2} + \sum_{s<r} \rho(\pi_{sr}) - \sum_{s>r} \rho(\pi_{rs}),$$

where $\rho(\delta_k) = k$, is a scoring function on $\mathcal{L}$.

Proof. The first condition is satisfied because if $r < s$, then

$$\{l_k \in \mathcal{L} \mid k < r\} \subsetneq \{l_k \in \mathcal{L} \mid k < s\} \text{ and } \{l_k \in \mathcal{L} \mid k > s\} \subsetneq \{l_k \in \mathcal{L} \mid k > r\}.$$

If π is the totally uniform OPM on $\mathcal{L}$, taking into account Remark 1, we have

$$S_a(l_r) = \frac{(g+2) \cdot (g-1)}{2} + (2 + \cdots + r) - (2 + \cdots + g - r + 1) =$$
$$\frac{(g+2) \cdot (g-1)}{2} + \frac{(r+2) \cdot (r-1)}{2} - \frac{(g-r+3) \cdot (g-r)}{2} = (r-1) \cdot (g+2).$$

Hence, the two conditions of scoring functions are satisfied. $\square$

Remark 3. The scoring function S_a can assign negative scores to the lowest linguistic terms of the OQS, depending on the OPM used to represent the perception on the OQS.[10] Note that in the context of Social Choice, Copeland [9]; Black [3]; Young [52] assign negative scores to some alternatives in the Borda and Copeland voting systems.

Remark 4. When $g = 2$, i.e., $\mathcal{L} = \{l_1, l_2\}$, there exists only one OPM: the associated with the proximity matrix A_2 (see the Annex). Obviously, that OPM is metrizable and totally uniform. The scoring functions S_b, S_w, S_{bw}, S'_b, S'_w, S'_{bw}, S'_a, S''_b, S''_w, S''_{bw} and S''_a give scores 0 and 1 to l_1 and l_2, respectively; moreover, $S_a(l_1) = 0$ and $S_a(l_2) = 4$. This situation corresponds to the dichotomous case, as in approval voting (see [4]), where the set of alternatives is divided in acceptable and unacceptable.

[10] For $g = 3, 4$, only $S_a(l_1)$ is negative, for all OPMs except those associated with the proximity matrices A_{22} and A_{222}, where $S_a(l_1)$ is positive.

Remark 5. Let $\mathcal{L} = \{l_1, \ldots, l_g\}$ be an OQS equipped with a metrizable OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$. For $g = 3$, the scores obtained from the three metrizable OPMs, those with associated proximity matrices A_{22}, A_{23} and A_{32} (see the Annex), are equal for the scoring functions S_b, S_w, S_{bw}, S'_b, S'_w, S'_{bw}, S''_b, S''_w and S''_{bw}. However, for $g > 3$ the complexity increases and the scores obtained can be different depending on the scoring function applied.

We now introduce the notion of compatibility of an OPM with a scoring function. It implies that if the ordinal proximity between to pairs of linguistic terms is the same, then the absolute difference between their scores should be also the same.

Definition 6. If $\mathcal{L} = \{l_1, \ldots, l_g\}$ is an OQS equipped with a metrizable OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$ and $S : \mathcal{L} \longrightarrow \mathbb{R}$ is a scoring function on $\mathcal{L}$, we say that π is *compatible with* S if

$$\pi_{rs} \succ \pi_{tu} \;\Rightarrow\; |S(l_s) - S(l_r)| < |S(l_u) - S(l_t)|,$$

for all $r, s, t, u \in \{1, \ldots, g\}$.

Remark 6. Let $\mathcal{L} = \{l_1, \ldots, l_g\}$ be an OQS equipped with a metrizable OPM $\pi : \mathcal{L} \times \mathcal{L} \longrightarrow \Delta$.

1. If π is compatible with S, then

   $$\pi_{rs} = \pi_{tu} \;\Leftrightarrow\; |S(l_s) - S(l_r)| = |S(l_u) - S(l_t)|,$$

   for all $r, s, t, u \in \{1, \ldots, g\}$.
2. If π is compatible with S, then it is also compatible with S' and S''.

We now show that for $g = 2, 3$ all the OPMs are compatible with all the scoring functions appearing in this section. However, for $g = 4$ not all the metrizable OPMs are compatible with those scoring functions.

Remark 7. Let $\mathcal{L} = \{l_1, \ldots, l_g\}$ be an OQS.

1. For $g = 2$, the only metrizable OPM is compatible with S_b, S_w, S_{bw} and S_a.
2. For $g = 3$, the three metrizable OPMs are compatible with S_b, S_w, S_{bw} and S_a.
3. For $g = 4$, 16 out of the 25 metrizable OPMs are compatible with S_b and S_w.
   The metrizable OPMs associated with the proximity matrices A'_{234}, A_{235}, A_{322}, A_{323}, A'_{324}, A_{325}, A_{423}, A'_{423} and A'_{432} are not compatible with S_b.
   The metrizable OPMs associated with the proximity matrices A_{223}, A_{234}, A'_{234}, A_{322}, A_{323}, A_{324}, A'_{324}, A'_{432} and A_{523} are not compatible with S_w.
4. For $g = 4$, 19 out of the 25 metrizable OPMs are compatible with S_{bw}.
   The metrizable OPMs associated with the proximity matrices A_{223}, A_{234}, A_{322}, A_{324}, A_{423} and A_{432} are not compatible with S_{bw}.
5. For $g = 4$, 23 out of the 25 metrizable OPMs are compatible with S_a.
   The metrizable OPMs associated with the proximity matrices A_{223} and A_{322} are not compatible with S_a.

We note that, for $g = 4$, S_a behaves better than S_{bw}, and S_{bw} better than S_b and S_w: 92% of the metrizable OPMs are compatible with S_a, 76% of the metrizable OPMs are compatible with S_{bw}, and 64% of the metrizable OPMs are compatible with S_b and S_w.

4. Discussion

In some decision-making procedures, it can be necessary to convert qualitative information into quantitative one by using numerical values, intervals, fuzzy sets, etc. For this purpose, some methods and procedures have been proposed. One of the most popular is the 2-tuple linguistic model of Herrera & Martínez [29]. This model uses a 2-tuple (s, α), where s is a linguistic term and α is a number that represents the symbolic translation within the range $[-0.5, 0.5)$. Agents show their opinions on the alternatives through a uniform OQS and it is after an aggregation process over the subindices of the linguistic terms (consecutive integer numbers) when the 2-tuples are applied.

In our proposal, OQSs are non-necessarily uniform. We note that in this framework several decision-making procedures have been devised in a pure ordinal fashion, without using numbers, by means of OPMs: García-Lapresta & Pérez-Román [23], García-Lapresta & González del Pozo [16], García-Lapresta et al. [20] and García-Lapresta & Marques Pereira [19], among others.

The aim of the proposed scoring functions is to simplify the management of qualitative information, avoiding the use of fuzzy membership functions or numerical values that do not consider the individuals' perceptions of the scales. The considered scoring functions are based on the notion of OPM, in a way that the scores assigned to the terms of OQSs try to collect, as faithfully as possible, the individuals' perceptions on the ordinal proximities between the linguistic terms of the OQSs.

The proposed scoring functions are based on pairwise comparisons between the corresponding term and: a single pole, the best term (S_b) or the worst one (S_w); two poles, both the best and worst linguistic terms (S_{bw}); and $g-1$ poles, the remaining terms of the OQS (S_a).

The choice of a scoring function depends on the decision-maker and the nature of the problem. For example, the scoring function S_{bw} requires fewer comparisons than the scoring function S_a, which simplifies its calculation. In addition, since S_{bw} takes into account the comparison between each linguistic term and the best and worst linguistic terms, the information provided is more complete than that obtained with the scoring functions S_b and S_w.

Another factor to consider when selecting a scoring function is its compatibility. As discussed in Remark 7, not all metrizable OPMs are compatible with the proposed scoring functions. For example, when $g=4$, S_a is compatible with more metrizable OPMs than S_{bw}, and the last one is compatible with more metrizable OPMs than S_b and S_w. Taking into account these drawbacks, if the metrizable OPM that best represents the perceptions on the proximities between the terms of the OQS has been obtained, a good strategy could be to select a scoring function so that the OPM will be compatible with it.

An important issue is how to determine the metrizable OPM that will be used. Clearly, different agents may have distinct perceptions about the OQS. García-Lapresta et al. [17] introduce an aggregation procedure that generates a collective metrizable OPM from the individuals' metrizable OPMs. They also provide a procedure that generates metrizable OPMs through suitable sequences of 2-4 questions for the case of OQSs with four linguistic terms (the case of three linguistic terms only requires a single question). When the OQS has more than four linguistic terms, that procedure is difficult to put in practice due to the number of questions and their complexity. In order to solve this problem, García-Lapresta et al. [18] introduce a visual procedure, managed through sliders, for obtaining the proximities between the terms of the OQS. An appropriate software generates the metrizable OPM which represents the ordinal arrangement of the proximities among terms.

5. The procedure

In multi-attribute group decision making, different approaches have been proposed to manage linguistic information (see [1], [56], [16], [48], [20], among others).

Many papers in the literature tackle with decision-making problems where experts' importance may be not necessarily the same for all of them. In some of these procedures, a numerical weight is assigned to each expert (see [2], [25], [35], among others), while in other cases the proposals are just based on hierarchies in which the relative weight of each expert is unknown (see [51], [13], among others). Likewise, Franceschini & García-Lapresta [12, Subsect. 3.3] introduce an approach in which the importance of each expert is based on the assessments provided by a decision-maker to the experts through an OQS equipped with an OPM.

In this section, we present a new decision-making procedure to rank alternatives that are evaluated by a group of experts with different knowledge. The procedure manages two different OQSs: one to determine the weights of experts in the procedure according to their expertise, and another one to assess the alternatives by the experts. These scales can be distinct and even they can be formed by a different number of linguistic terms.

The proposed procedure assigns scores to the linguistic terms of the scales taking into account the ordinal proximities between the terms of each scale through appropriate metrizable OPMs. The assignment is carried out through scoring functions. The decision-making procedure is illustrated by a flowchart in Fig. 5.

To rank the alternatives the procedure is divided in six steps.

1. A decision-maker (DM) evaluates a set of experts $E=\{e_1,\dots,e_m\}$ by means of an OQS $\mathcal{L}^e=\left\{l^e_1,\dots,l^e_{g_e}\right\}$ equipped with a metrizable OPM[11]

$$\pi^e : \mathcal{L}^e \times \mathcal{L}^e \longrightarrow \Delta^e = \left\{\delta^e_1,\dots,\delta^e_{h_e}\right\}.$$

.
With $v_k \in \mathcal{L}^e$ we denote the assessment obtained by the expert $e_k \in E$.
2. A weight $w_k \in [0,1]$ is assigned to each expert $e_k \in E$ through a scoring function $S^e : \mathcal{L}^e \longrightarrow \mathbb{R}$ as follows:

$$w_k = \frac{S^e(v_k)}{S^e(v_1)+\dots+S^e(v_m)}. \tag{1}$$

Note that $w_1+\dots+w_m=1$.
3. The experts of E evaluate a set of alternatives $X=\{x_1,\dots,x_n\}$ through an OQS $\mathcal{L}^a=\left\{l^a_1,\dots,l^a_{g_a}\right\}$ equipped with a metrizable OPM

[11] This OPM can be obtained from the perceptions about the scale provided by the DM. In turn, the assessments given by the DM can be based on objective information on the experts.

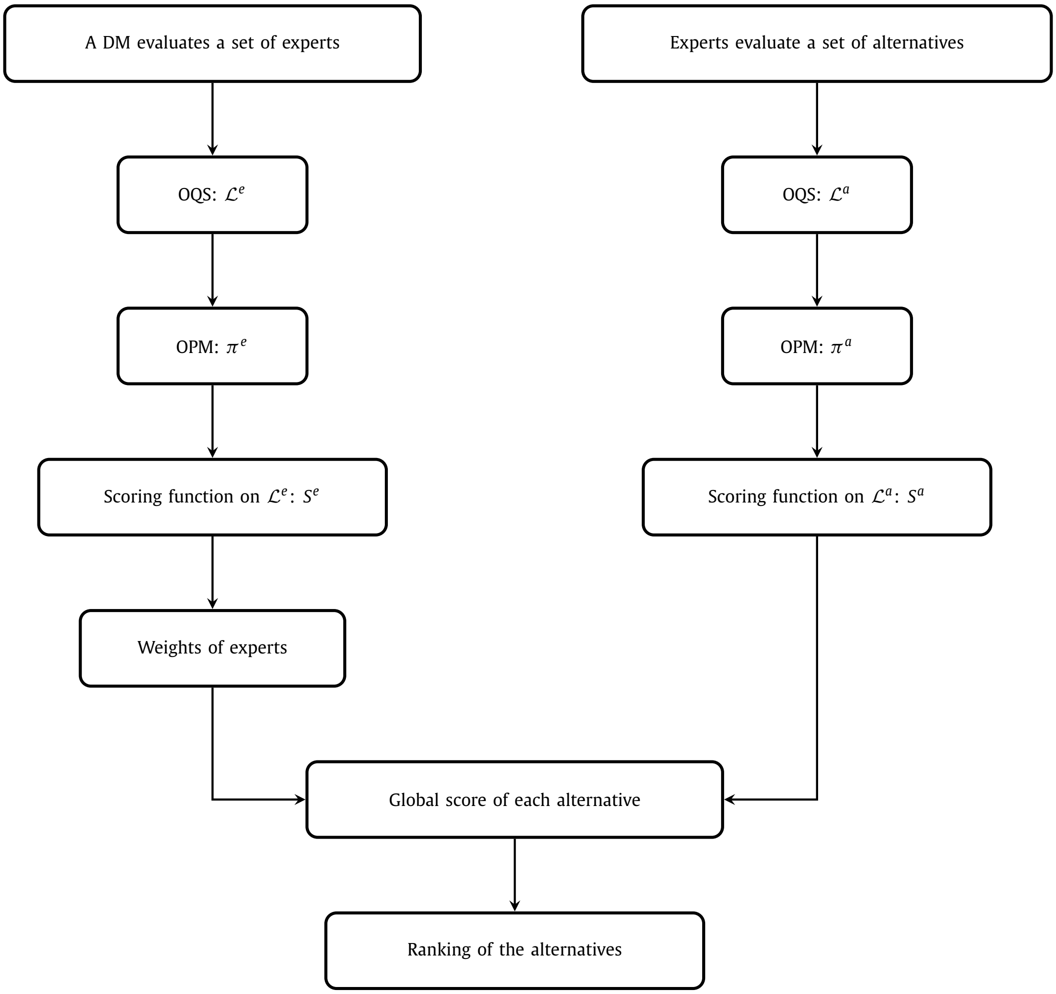

Fig. 5. Flowchart of the procedure.

$$\pi^a : \mathcal{L}^a \times \mathcal{L}^a \longrightarrow \Delta^a = \left\{\delta_1^a, \dots, \delta_{h_a}^a\right\}.$$

The experts' opinions on the alternatives are collected in a matrix of m rows and n columns with coefficients in $\mathcal{L}^a$:

$$\begin{pmatrix} v_1^1 & \cdots & v_i^1 & \cdots & v_n^1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ v_1^k & \cdots & v_i^k & \cdots & v_n^k \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ v_1^m & \cdots & v_i^m & \cdots & v_n^m \end{pmatrix},$$

where $v_i^k \in \mathcal{L}^a$ is the linguistic assessment given by the expert $e_k \in E$ to the alternative $x_i \in X$.

4. A score is assigned to each linguistic term of the OQS $\mathcal{L}^a$ through a scoring function $S^a : \mathcal{L}^a \longrightarrow [0, 1]$.
5. A global score is assigned to each alternative $x_i \in X$ through the function $U : X \longrightarrow [0, 1]$ defined as

$$U(x_i) = \sum_{k=1}^{m} w_k \cdot S^a\left(v_i^k\right). \tag{2}$$

6. Finally, the alternatives are ranked through the weak order $\succeq$ defined as

$$x_i \succeq x_j \Leftrightarrow U(x_i) \geq U(x_j).$$

Remark 8. In group decision-making it is desirable to reach as much consensus as possible before applying the procedure that generates the outcome. In this way, *consensus reaching processes* are a useful tool that have been widely studied in the literature: Herrera et al. [28], Herrera-Viedma et al. [30], Wu & Chiclana [50] and Cabrerizo et al. [6], among many others.

Table 1
OQS used to evaluate the experts.

Linguistic terms in $\mathcal{L}^e$	
l_1^e	Low
l_2^e	Moderate
l_3^e	High
l_4^e	Very high

Table 2
Assessments given by the decision-maker to the experts.

e_1	e_2	e_3	e_4
l_2^e	l_3^e	l_4^e	l_3^e

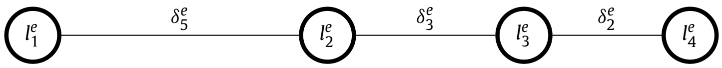

Fig. 6. Metrizable OPM associated with the proximity A_{532}.

In the setting of OQSs, García-Lapresta & Pérez-Román [22] propose a consensus reaching process that takes into account the perceptions of the closeness between the terms of the scale by means of OPMs. Once agents evaluate the alternatives, the consensus is measured through medians of ordinal proximities between the assessments provided by the agents. Given an ordinal threshold, the procedure selects the set of alternatives where the degrees of consensus are lower than the overall degree of consensus. For each of these alternatives, a moderator invites the agents whose assessments are greater or lower than the median assessment to modify their opinions in order to increase the consensus in the group.

In the framework of the present paper, it is possible to implement the consensus reaching process of García-Lapresta & Pérez-Román [22] in each criterion. But, it is also possible to apply a conventional process if the numerical scores generated by a scoring function are considered, instead of the original qualitative assessments.

6. Case study

In order to illustrate how the proposed procedure works, we present an illustrative case study. We have considered a set of four experts, $E = \{e_1, \dots, e_4\}$, that assess a set of five alternatives $X = \{x_1, \dots, x_5\}$.

In the first step of the procedure, a DM evaluates the experts' expertise by means of the 4-term OQS $\mathcal{L}^e = \{l_1^e, l_2^e, l_3^e, l_4^e\}$ contained in Table 1, whose linguistic terms are commonly used to measure agents' abilities and knowledge about certain aspects in different areas (see [26], [43], among others).

Table 2 shows the linguistic assessments given by the DM to the four experts according their expertise.

To apply the proposed procedure, we need to establish an appropriate OPM that collects how the proximities between the four linguistic terms are perceived. In this case, we have considered that the 4-term OQS $\mathcal{L}^e$ is equipped with the metrizable OPM π^e associated with the proximity matrix A_{532}:

$$A_{532} = \begin{pmatrix} \delta_1^e & \delta_5^e & \delta_6^e & \delta_7^e \\ & \delta_1^e & \delta_3^e & \delta_4^e \\ & & \delta_1^e & \delta_2^e \\ & & & \delta_1^e \end{pmatrix}$$

which can be visualized in Fig. 6.

To determine the scores of the OQS used to evaluate the experts' expertise, it can be applied any of the scoring functions proposed in Section 3. In this case study, we have considered the scoring function S_{bw}, whose scores are collected in Table 3. This scoring function is the arithmetic mean of the scoring functions S_b and S_w, so it provides more complete information than the formers. In addition, the scoring function S_a can be also considered. In the case of the metrizable OPM associated with the proximity matrix A_{532}, we obtain a negative score for the linguistic term l_1.[12]

Then, taking into account the assessments given by the DM to the experts (see Table 2), we determine the experts' weights by means of Eq. (1). These weights are presented in Table 4.

[12] Since it seems not be reasonable to assign a negative score to the experts, the scoring function S_a has not been considered at this step of the procedure. However, it is important to note that the normalized scoring function S_a'' never assigns negative scores. Therefore, S_a'' could be applied to the mentioned metrizable OPM.

Table 3
S_{bw} scores for each linguistic term in $\mathcal{L}^e$ obtained from the metrizable OPM associated with the proximity matrix A_{532}.

Term	S_{bw}
l_1^e	0
l_2^e	3.5
l_3^e	5
l_4^e	6

Table 4
Experts' weights.

Expert	Weight
e_1	0.17949
e_2	0.25641
e_3	0.30769
e_4	0.25641

Table 5
OQS used to evaluate the alternatives.

Linguistic terms in $\mathcal{L}^a$	
l_1^a	Poor
l_2^a	Fair
l_3^a	Good
l_4^a	Very good
l_5^a	Excellent

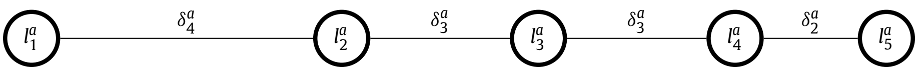

Fig. 7. Metrizable OPM associated with the proximity matrix A_{4332}.

Table 6
Assessments given the experts to the alternatives.

	x_1	x_2	x_3	x_4	x_5
e_1	l_1^a	l_4^a	l_2^a	l_2^a	l_2^a
e_2	l_2^a	l_2^a	l_2^a	l_3^a	l_5^a
e_3	l_2^a	l_3^a	l_2^a	l_4^a	l_3^a
e_4	l_3^a	l_5^a	l_3^a	l_5^a	l_4^a

On the other hand, the experts evaluate the alternatives through the 5-term OQS $\mathcal{L}^a = \{l_1^a, l_2^a, l_3^a, l_4^a, l_5^a\}$ contained in Table 5. This is a typical scale for Quality of Experience (QoE) measurement (see [41], [49], among others).

Taking into account the meaning of the linguistic terms shown in Table 5, we have considered that the 5-term OQS is equipped with the metrizable OPM π^a associated with the proximity matrix A_{4332}:

$$A_{4332} = \begin{pmatrix} \delta_1^a & \delta_4^a & \delta_7^a & \delta_9^a & \delta_{10}^a \\ & \delta_1^a & \delta_3^a & \delta_6^a & \delta_8^a \\ & & \delta_1^a & \delta_3^a & \delta_5^a \\ & & & \delta_1^a & \delta_2^a \\ & & & & \delta_1^a \end{pmatrix}$$

that can be visualized in Fig. 7.

The assessments given by the experts to the alternatives are collected in Table 6.

Table 7
Scores for each linguistic term in $\mathcal{L}^a$ obtained from the metrizable OPM associated with the proximity matrix A_{4332}.

Term	S_a	S''_a
l_1^a	-16	0
l_2^a	1	0.30909
l_3^a	16	0.58182
l_4^a	30	0.83636
l_5^a	39	1

Table 8
Global scores of the alternatives.

Alternative	$U(x_i)$
x_1	0.32354
x_2	0.66480
x_3	0.37902
x_4	0.71841
x_5	0.70536

To obtain the scores corresponding to the metrizable OPM associated with the proximity matrix A_{4332}, we use the scoring function S_a whose scores have been normalized between 0 and 1 through the scoring function S'' (see Proposition 2), i.e., S''_a.

Then, from the experts' weights and the scores assigned to each linguistic term of the 5-term OQS, we calculate the global score for each alternative by means of Eq. (2). These final scores are contained in Table 8.

Finally, we rank the alternatives according to their global scores:

$$x_4 \succ x_5 \succ x_2 \succ x_3 \succ x_1.$$

Taking into account the global scores of the alternatives shown in Table 8, the scores of the linguistic terms in Table 7 and the meaning of the linguistic terms in Table 5, we can also provide the range of their linguistic global assessments: x_1, between fair and good; the rest of alternatives, between good and very good.

7. Extension to multiple criteria

The proposed procedure can be extended to more general scenarios. We now consider a set of experts $E = \{e_1, \dots, e_m\}$ that evaluate a set of alternatives $X = \{x_1, \dots, x_n\}$ regarding a set of criteria $C = \{c_1, \dots, c_q\}$ by means of OQSs $\mathcal{L}^p = \left\{l_1^p, \dots, l_{g_p}^p\right\}$ equipped with metrizable OPMs

$$\pi^p : \mathcal{L}^p \times \mathcal{L}^p \longrightarrow \Delta^p = \left\{\delta_1^p, \dots, \delta_{h_p}^p\right\},$$

for each criterion $c_p \in C$.

The opinions given the experts to all the alternatives regarding the criterion $c_p \in C$ are collected in a matrix of m rows and n columns with coefficients in $\mathcal{L}^p$:

$$\begin{pmatrix} v_1^{1,p} & \cdots & v_i^{1,p} & \cdots & v_n^{1,p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ v_1^{k,p} & \cdots & v_i^{k,p} & \cdots & v_n^{k,p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ v_1^{m,p} & \cdots & v_i^{m,p} & \cdots & v_n^{m,p} \end{pmatrix},$$

where $v_i^{k,p} \in \mathcal{L}^p$ is the assessment given by the expert $e_k \in E$ to the alternative $x_i \in X$ with respect to the criterion $c_p \in C$.

Since criteria may have different importance in decision-making procedures, they can have different weights. Although usually a DM provides the weights of the criteria, we propose to allow experts to evaluate the importance of each criterion through an OQS $\mathcal{L}^c = \{l_1^c, \dots, l_{g_c}^c\}$ equipped with a metrizable OPM

$$\pi^c : \mathcal{L}^c \times \mathcal{L}^c \longrightarrow \Delta^c = \left\{\delta_1^c, \dots, \delta_{h_c}^c\right\}.$$

If $u_p^k \in \mathcal{L}^c$ is the assessment given by the expert $e_k \in E$ to the criterion $c_p \in C$, a weight $w_p^k \in [0,1]$ is assigned to c_p by e_k as follows:

$$w_p^k = \frac{S^c\left(u_p^k\right)}{S^c\left(u_1^k\right) + \cdots + S^c\left(u_q^k\right)},$$

where S^c is one of the scoring functions introduced in Section 3.

Note that $w_1^k + \cdots + w_q^k = 1$ for every expert $e_k \in E$.

In accordance with the Step 5 of the procedure proposed in Section 5, the criteria weights are considered to obtain the global score of each alternative through the function $U_c : X \longrightarrow [0,1]$ defined as

$$U_c(x_i) = \sum_{k=1}^{m} \sum_{p=1}^{q} w_k \cdot w_p^k \cdot S^a\left(v_i^{k,p}\right) \in [0,1],$$

where w_k is obtained by means of Eq. (1) from the assessments given by a DM to the experts, as in the Steps 1 and 2 of the procedure proposed in Section 5.

Finally, the alternatives are ranked through the following weak order

$$x_i \succeq x_j \Leftrightarrow U_c(x_i) \geq U_c(x_j).$$

8. Concluding remarks

In this paper, we have introduced a new group decision-making procedure to rank a set of alternatives based on the opinions of a group of experts with different knowledge. In our proposal, experts and alternatives are evaluated by means of OQSs equipped with metrizable OPMs, which are considered to assign a numerical score to each linguistic term through appropriate scoring functions.

To deal with qualitative information coming from OQSs, it is quite common the use of equidistant numerical scores[13] or even, as we have seen in Remark 2, the practice carried out by several organizations and data providers of grouping the highest linguistic terms of the scales. Therefore, it is important to note that the results obtained by means of these methods should be interpreted with a great care and attention.

The main contribution of this paper is the introduction of a new methodology that assigns a score to each linguistic term of an OQS, taking into account the proximities between the linguistic terms of the OQS by means of the concept of metrizable OPM. To do that, this paper proposes and analyzes several scoring functions that are based on the ordinal proximities between the considered linguistic term and the best, the worst, the best and the worst, and the rest of linguistic terms of the OQS: S_b, S_w, S_{bw} and S_a, respectively (Propositions 4, 5, 6 and 7). Additionally, we have proposed two normalization procedures of scoring functions, S' and S'' (Propositions 1 and 2). All these scoring functions can be used in different decision-making problems where agents evaluate a set of alternatives by means of OQSs.

On the other hand, the procedure introduced in this paper does not consider the possible uncertainty when the DM evaluates the experts or when the experts evaluate the alternatives. For that reason, regarding further research, it could be interesting to extend the proposed procedure to situations in which the DM or experts are allowed to select two consecutive linguistic terms if they hesitate (see [16]).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

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

Acknowledgements

The authors are grateful to Miguel Martínez-Panero, David Pérez-Román and the participants to the *18th Applied Stochastic Models and Data Analysis International Conference (ASMDA 2019)*, the *1st International 'Alan Turing' Conference on Decision Support and Recommender Systems* and the *31th European Conference on Operational Research - EURO 2021* for their comments and suggestions. The authors are also grateful to two anonymous reviewers for their useful comments and suggestions. The financial support of the Spanish *Ministerio de Ciencia e Innovación* (project PID2021-122506NB-I00) is acknowledged.

[13] According to Definition 5, this practice is suitable when considering the totally uniform OPM of an OQS, but not in general (see Remark 6).

Appendix A. Metrizable OPMs for $g = 2$, $g = 3$ and $g = 4$

Metrizable OPM for $g = 2$:

$$A_2 = \begin{pmatrix} \delta_1 & \delta_2 \\ & \delta_1 \end{pmatrix}.$$

Metrizable OPMs for $g = 3$:

$$A_{22} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_3 \\ & \delta_1 & \delta_2 \\ & & \delta_1 \end{pmatrix}, \quad A_{23} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_4 \\ & \delta_1 & \delta_3 \\ & & \delta_1 \end{pmatrix}, \quad A_{32} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_4 \\ & \delta_1 & \delta_2 \\ & & \delta_1 \end{pmatrix}.$$

Metrizable OPMs for $g = 4$:

$$A_{222} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_3 & \delta_4 \\ & \delta_1 & \delta_2 & \delta_3 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix}, \quad A_{223} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_4 & \delta_6 \\ & \delta_1 & \delta_2 & \delta_5 \\ & & \delta_1 & \delta_3 \\ & & & \delta_1 \end{pmatrix},$$

$$A'_{223} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_3 & \delta_5 \\ & \delta_1 & \delta_2 & \delta_4 \\ & & \delta_1 & \delta_3 \\ & & & \delta_1 \end{pmatrix}, \quad A_{224} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_3 & \delta_6 \\ & \delta_1 & \delta_2 & \delta_5 \\ & & \delta_1 & \delta_4 \\ & & & \delta_1 \end{pmatrix},$$

$$A_{232} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_4 & \delta_5 \\ & \delta_1 & \delta_3 & \delta_4 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix}, \quad A_{233} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_4 & \delta_6 \\ & \delta_1 & \delta_3 & \delta_5 \\ & & \delta_1 & \delta_3 \\ & & & \delta_1 \end{pmatrix},$$

$$A_{234} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_5 & \delta_7 \\ & \delta_1 & \delta_3 & \delta_6 \\ & & \delta_1 & \delta_4 \\ & & & \delta_1 \end{pmatrix}, \quad A'_{234} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_4 & \delta_6 \\ & \delta_1 & \delta_3 & \delta_5 \\ & & \delta_1 & \delta_4 \\ & & & \delta_1 \end{pmatrix},$$

$$A_{235} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_4 & \delta_7 \\ & \delta_1 & \delta_3 & \delta_6 \\ & & \delta_1 & \delta_5 \\ & & & \delta_1 \end{pmatrix}, \quad A_{243} = \begin{pmatrix} \delta_1 & \delta_2 & \delta_5 & \delta_7 \\ & \delta_1 & \delta_4 & \delta_6 \\ & & \delta_1 & \delta_3 \\ & & & \delta_1 \end{pmatrix},$$

$$A_{322} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_5 & \delta_6 \\ & \delta_1 & \delta_2 & \delta_4 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix}, \quad A'_{322} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_4 & \delta_5 \\ & \delta_1 & \delta_2 & \delta_3 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix},$$

$$A_{323} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_4 & \delta_5 \\ & \delta_1 & \delta_2 & \delta_4 \\ & & \delta_1 & \delta_3 \\ & & & \delta_1 \end{pmatrix}, \quad A_{324} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_5 & \delta_7 \\ & \delta_1 & \delta_2 & \delta_6 \\ & & \delta_1 & \delta_4 \\ & & & \delta_1 \end{pmatrix},$$

$$A'_{324} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_4 & \delta_6 \\ & \delta_1 & \delta_2 & \delta_5 \\ & & \delta_1 & \delta_4 \\ & & & \delta_1 \end{pmatrix}, \quad A_{325} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_4 & \delta_7 \\ & \delta_1 & \delta_2 & \delta_6 \\ & & \delta_1 & \delta_5 \\ & & & \delta_1 \end{pmatrix},$$

$$A_{332} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_5 & \delta_6 \\ & \delta_1 & \delta_3 & \delta_4 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix}, \quad A_{342} = \begin{pmatrix} \delta_1 & \delta_3 & \delta_6 & \delta_7 \\ & \delta_1 & \delta_4 & \delta_5 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix},$$

$$A_{422} = \begin{pmatrix} \delta_1 & \delta_4 & \delta_5 & \delta_6 \\ & \delta_1 & \delta_2 & \delta_3 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix}, \quad A_{423} = \begin{pmatrix} \delta_1 & \delta_4 & \delta_6 & \delta_7 \\ & \delta_1 & \delta_2 & \delta_5 \\ & & \delta_1 & \delta_3 \\ & & & \delta_1 \end{pmatrix},$$

$$A'_{423} = \begin{pmatrix} \delta_1 & \delta_4 & \delta_5 & \delta_6 \\ & \delta_1 & \delta_2 & \delta_4 \\ & & \delta_1 & \delta_3 \\ & & & \delta_1 \end{pmatrix}, \quad A_{432} = \begin{pmatrix} \delta_1 & \delta_4 & \delta_6 & \delta_7 \\ & \delta_1 & \delta_3 & \delta_5 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix},$$

$$A'_{432} = \begin{pmatrix} \delta_1 & \delta_4 & \delta_5 & \delta_6 \\ & \delta_1 & \delta_3 & \delta_4 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix}, \quad A_{523} = \begin{pmatrix} \delta_1 & \delta_5 & \delta_6 & \delta_7 \\ & \delta_1 & \delta_2 & \delta_4 \\ & & \delta_1 & \delta_3 \\ & & & \delta_1 \end{pmatrix},$$

$$A_{532} = \begin{pmatrix} \delta_1 & \delta_5 & \delta_6 & \delta_7 \\ & \delta_1 & \delta_3 & \delta_4 \\ & & \delta_1 & \delta_2 \\ & & & \delta_1 \end{pmatrix}.$$

References

[1] R. de Andrés, J.L. García-Lapresta, L. Martínez, A multi-granular linguistic model for management decision-making in performance appraisal, Soft Comput. 14 (2010) 21–34.
[2] K.J. Arrow, H. Raynaud, Social Choice and Multicriterion Decision-Making, Cambridge, 1986.
[3] D. Black, The Theory of Committees and Elections, Publishers, Boston, 1958.
[4] S.J. Brams, P.C. Fishburn, Approval voting, Am. Polit. Sci. Rev. 72 (1978) 831–847.
[5] J.J. Buckley, Fuzzy hierarchical analysis, Fuzzy Sets Syst. 17 (1985) 233–247.
[6] F.J. Cabrerizo, F. Chiclana, R. Al-Hmouz, A. Morfeq, A.S. Balamash, E. Herrera-Viedma, Fuzzy decision making and consensus: challenges, J. Intell. Fuzzy Syst. 29 (2015) 1109–1118.
[7] F.J. Cabrerizo, I.J. Pérez, E. Herrera-Viedma, Managing the consensus in group decision making in an unbalanced fuzzy linguistic context with incomplete information, Knowl.-Based Syst. 23 (2010) 169–181.
[8] S. Casacci, A. Pareto, Methods for quantifying ordinal variables: a comparative study, Qual. Quant. 49 (2015) 1859–1872.
[9] A.H. Copeland, A reasonable social welfare function, Technical Report Mimeo, University of Michigan, USA, 1951.
[10] K.E. Flynn, M.A. Smith, D. Vanness, A typology of preferences for participation in healthcare decision making, Soc. Sci. Med. 63 (2006) 1158–1169.
[11] F. Franceschini, M. Galetto, M. Varetto, Qualitative ordinal scales: the concept of ordinal range, Qual. Eng. 16 (2004) 515–524.
[12] F. Franceschini, J.L. García-Lapresta, Decision-making in semi-democratic contexts, Inf. Fusion 52 (2019) 281–289.
[13] F. Franceschini, D. Maisano, L. Mastrogiacomo, A new proposal for fusing individual preference orderings by rank-ordered agents: a generalization of the Yager's algorithm, Eur. J. Oper. Res. 249 (2016) 209–223.
[14] T. Gadrich, E. Bashkansky, ORDANOVA: analysis of ordinal variation, J. Stat. Plan. Inference 142 (2012) 3174–3188.
[15] T. Gadrich, E. Bashkansky, R. Zitikis, Assessing variation: a unifying approach for all scales of measurement, Qual. Quant. 49 (2015) 1145–1167.
[16] J.L. García-Lapresta, R. González del Pozo, An ordinal multi-criteria decision-making procedure under imprecise linguistic assessments, Eur. J. Oper. Res. 279 (2019) 159–167.
[17] J.L. García-Lapresta, R. González del Pozo, D. Pérez-Román, Metrizable ordinal proximity measures and their aggregation, Inf. Sci. 448 (2018) 149–163.
[18] J.L. García-Lapresta, R. Iurev, D. Pérez-Román, Managing perceptions on the linguistic terms of qualitative scales, (2023), submitted.
[19] J.L. García-Lapresta, R.A. Marques Pereira, An extension of Majority Judgment to non-uniform qualitative scales, Eur. J. Oper. Res. 301 (2022) 667–674.
[20] J.L. García-Lapresta, P. Moreno-Albadalejo, D. Pérez-Román, V. Temprano-García, A multi-criteria procedure in new product development using different qualitative scales, Appl. Soft Comput. 106 (2021) 107279.
[21] J.L. García-Lapresta, D. Pérez-Román, Ordinal proximity measures in the context of unbalanced qualitative scales and some applications to consensus and clustering, Appl. Soft Comput. 35 (2015) 864–872.
[22] J.L. García-Lapresta, D. Pérez-Román, A consensus reaching process in the context of non-uniform ordered qualitative scales, Fuzzy Optim. Decis. Mak. 16 (2017) 449–461.
[23] J.L. García-Lapresta, D. Pérez-Román, Aggregating opinions in non-uniform ordered qualitative scales, Appl. Soft Comput. 67 (2018) 652–657.
[24] R. González del Pozo, L.C. Dias, J.L. García-Lapresta, Using different qualitative scales in a multi-criteria decision-making procedure, Mathematics 8 (2020) 458.
[25] S. Greco, M. Ehrgott, J. Figueira, Multiple Criteria Decision Analysis: State of the Art Surveys, International Series in Operations Research & Management Science, vol. 78, Springer, 2016.
[26] M.E. Groetch, L. Christie, P.A. Vargas, S.M. Jones, S.H. Sicherer, Food allergy educational needs of pediatric dietitians: a survey by the consortium of food allergy research, J. Nutr. Educ. Behav. 42 (2010) 259–264.
[27] F. Herrera, E. Herrera-Viedma, L. Martínez, A fuzzy linguistic methodology to deal with unbalanced linguistic term sets, IEEE Trans. Fuzzy Syst. 16 (2008) 354–370.
[28] F. Herrera, E. Herrera-Viedma, J.L. Verdegay, Linguistic measures based on fuzzy coincidence for reaching consensus in group decision making, Int. J. Approx. Reason. 16 (1997) 309–334.
[29] F. Herrera, L. Martínez, A 2-tuple fuzzy linguistic representation model for computing with words, IEEE Trans. Fuzzy Syst. 8 (2000) 746–752.
[30] E. Herrera-Viedma, F. Herrera, F. Chiclana, A consensus model for multiperson decision making with different preference structures, IEEE Trans. Syst. Man Cybern., Part A, Syst. Hum. 32 (2002) 394–402.
[31] C.L. Hwang, K. Yoon, Multiple Attributes Decision Making Methods and Applications, Springer, 1981.
[32] A. Jahan, K.L. Edwards, A state-of-the-art survey on the influence of normalization techniques in ranking: improving the materials selection process in engineering design, Mater. Des. 65 (2015) 335–342.
[33] S. Labovitz, The assignment of numbers to rank order categories, Am. Sociol. Rev. 35 (1970) 515–524.
[34] R.J. Li, Fuzzy method in group decision making, Comput. Math. Appl. 38 (1999) 91–101.
[35] Y. Li, G. Kou, G. Li, H. Wang, Multi-attribute group decision making with opinion dynamics based on social trust network, Inf. Fusion 75 (2021) 102–115.
[36] Y. Liu, C.M. Eckert, C. Earl, A review of fuzzy AHP methods for decision-making with subjective judgements, Expert Syst. Appl. 161 (2020) 113738.
[37] X. Ma, J. Qin, L. Martínez, W. Pedrycz, A linguistic information granulation model based on best-worst method in decision making problems, Inf. Fusion 89 (2023) 210–227.

[38] L. Martínez, F. Herrera, An overview on the 2-tuple linguistic model for computing with words in decision making: extensions, applications and challenges, Inf. Sci. 207 (2012) 1–18.
[39] P. McCullagh, Regression models for ordinal data, J. R. Stat. Soc., Ser. B, Methodol. 42 (1980) 109–127.
[40] C. Merbitz, J. Morris, J.C. Grip, Ordinal scales and foundations of misinference, Arch. Phys. Med. Rehabil. 70 (1989) 308–312.
[41] K. Mitra, A. Zaslavsky, C. Åhlund, Context-aware QoE modelling, measurement, and prediction in mobile computing systems, IEEE Trans. Mob. Comput. 14 (2013) 920–936.
[42] J. Rezaei, Best-worst multi-criteria decision-making method, Omega 53 (2015) 49–57.
[43] M. Roduta Roberts, C.B. Alves, M.-W. Chu, M. Thompson, L.M. Bahry, A. Gotzmann, Testing expert-based versus student-based cognitive models for a grade 3 diagnostic mathematics assessment, Appl. Meas. Educ. 27 (2014) 173–195.
[44] T.L. Saaty, The Analytic Hierarchy Proces, McGraw-Hill, 1980.
[45] L.F.d.O.M. Santos, L. Osiro, R.H.P. Lima, A model based on 2-tuple fuzzy linguistic representation and Analytic Hierarchy Process for supplier segmentation using qualitative and quantitative criteria, Expert Syst. Appl. 79 (2017) 53–64.
[46] S. Siegel, Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill, 1956.
[47] S.S. Stevens, Mathematics, Measurement, and Psychophysics, Wiley, 1951.
[48] R. Ureña, G. Kou, Y. Dong, F. Chiclana, E. Herrera-Viedma, A review on trust propagation and opinion dynamics in social networks and group decision making frameworks, Inf. Sci. 478 (2019) 461–475.
[49] S. Wang, K.S. Ong, P. Surman, J. Yuan, Y. Zheng, X.W. Sun, Quality of experience measurement for light field 3D displays on multilayer LCDs, J. Soc. Inf. Disp. 24 (2016) 726–740.
[50] J. Wu, F. Chiclana, Multiplicative consistency of intuitionistic reciprocal preference relations and its application to missing values estimation and consensus building, Knowl.-Based Syst. 71 (2014) 187–200.
[51] R.R. Yager, Fusion of multi-agent preference orderings, Fuzzy Sets Syst. 117 (2001) 1–12.
[52] H.P. Young, An axiomatization of Borda's rule, J. Econ. Theory 9 (1974) 43–52.
[53] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning I, Inf. Sci. 8 (1975) 199–249.
[54] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning II, Inf. Sci. 8 (1975) 301–357.
[55] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning III, Inf. Sci. 9 (1975) 43–80.
[56] Z. Zhang, C. Guo, L. Martínez, Managing multigranular linguistic distribution assessments in large-scale multiattribute group decision making, IEEE Trans. Syst. Man Cybern. Syst. 47 (2016) 3063–3076.
[57] Z. Zhang, Z. Li, Y. Gao, Consensus reaching for group decision making with multi-granular unbalanced linguistic information: a bounded confidence and minimum adjustment-based approach, Inf. Fusion 74 (2021) 96–110.
[58] H. Zhu, J. Zhao, Y. Xu, 2-dimension linguistic computational model with 2-tuples for multi-attribute group decision making, Knowl.-Based Syst. 103 (2016) 132–142.
[59] C. Zuheros, C.C. Li, F.J. Cabrerizo, Y. Dong, E. Herrera-Viedma, F. Herrera, Computing with words: revisiting the qualitative scale, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 26 (2018) 127–143.