Descriptive analysis of responses to items in questionnaires. Why not using a fuzzy rating scale? [1]

María Asunción Lubiano, [a] Sara de la Rosa de Sáa, [a,b]
Manuel Montenegro, [a] Beatriz Sinova, [a] María Ángeles Gil* [a]

[a] *Departamento de Estadística, I.O. y D.M., Facultad de Ciencias Universidad de Oviedo, E-33071 Oviedo, Spain*

[b] *Institut für Stochastik und Wirtschaftsmathematik, Technische Universität Wien, Wiedner Hauptstr. 8-10/E105, 1040 Vienna, Austria*

Abstract

In evaluating aspects like quality perception, satisfaction or attitude which are intrinsically imprecise, the fuzzy rating scale has been introduced as a psychometric tool that allows evaluators to give flexible and quite accurate, albeit non numerical, ratings. The fuzzy rating scale integrates the skills associated with the visual analogue scale, because of the total freedom in assessing ratings, with the ability of fuzzy linguistic variables to capture the natural imprecision in evaluating such aspects.

Thanks to a recent methodology, the descriptive analysis of the responses to a fuzzy rating scale-based questionnaire can be now carried out. This paper aims to illustrate such an analysis through a real-life example, as well as to show that statistical conclusions can often be rather different from the conclusions one could get from either Likert scale-based responses or their fuzzy linguistic encoding. This difference encourages the use of the fuzzy rating scale when statistical conclusions are important, similarly to the use of exact real-valued data instead of grouping them.

Key words: descriptive summary measures, fuzzy data, fuzzy linguistic data, fuzzy rating scale, questionnaire

---

[1] Authors are grateful to Colegio San Ignacio in Oviedo-Asturias (Spain), its staff and its fourth grade students and parents for the support and understanding shown when they allowed us to collect the data in the real-life example. Actually, the nine and ten-year-old students in this school have quite properly caught the spirit and the mechanics behind fuzzy rating scale-based questionnaires. Authors want to acknowledge their colleague Professor Carlos Carleos because of his design and implementation of the computerized format many students have used to answer the questionnaire. Authors also wish to thank the Associate Editor and the reviewers in charge of the original, because of their valuable suggestions to improve the paper content and reading. The research in this paper has been partially supported by/benefited from Principality of Asturias Grants GRUPIN14-101 and Severo Ochoa BP12012 (De la Rosa de Sáa), and the Spanish Ministry of Economy and Competitiveness Grant MTM2013-44212-P. Their financial support is gratefully acknowledged.

* corresponding author: magil@uniovi.es, Fax: (+34) 985103354.

1 Introduction

In rating many characteristics or attributes that cannot be directly measured (like perceived quality, satisfaction, perception, attitude...), different scales have been considered. The best known scales in this setting (see, for instance, Yuen [66]) are the discrete ones, which consist of choosing the most appropriate 'values' within a class according to the rater judgement (like Likert-type scales) and the continuous visual analogue scale, whose aim is to specify an exact level of agreement to a statement or property by choosing a single point along a line between two end-points (i.e., along a compact interval).

When *Likert-type scale* data are analyzed for statistical purposes, they are either treated as categorical ones, so that techniques to analyze them are quite limited, or numerically encoded by means of consecutive integer numbers, which enlarges to some extent the number of procedures to be applied. The main positive features in connection with the use of Likert-type scales are that

▲ surveys/questionnaires based on them are easy-to-conduct, and they require neither a demanding training nor a special framework,

▲ labels for the given 'values' are expressed in terms that properly fit the intrinsic imprecision of the considered characteristics or attributes.

Nevertheless, several weaknesses have been highlighted in the literature on this type of scale (see, for instance, Jamieson [31], Carifio and Perla [9], Calcagnì and Lombardi [8]), namely,

▼ the number of possible 'values' to choose among is small, so the variability, diversity and subjectivity associated with an accurate rating is usually lost, and the choice of the most appropriate 'value' is not necessarily a simple task (raters often prefer to have the opportunity of choosing in between two 'values');

▼ when 'values' are encoded by their relative position in accordance with a certain ranking, differences between codes cannot be interpreted as differences in their magnitude, so only the statistical conclusions addressed to categorical or ordinal data become really reliable and relevant information can be lost.

Several studies (see, for instance, Reips and Funke [46], and Treiblmaier and Filzmoser [60]) have pointed out that the *visual analogue scale* provides researchers with many advantages in contrast to discrete scales. Thus, one can benefit from a metric setting as well as from the fact that a much wider set of statistical methods can be applied to analyze the data coming from this rating. As for the Likert scales, one can find some *pros* and *cons*.

Among the *pros*, one can remark that

▲ the choice is to be made within a continuum, so the variability, diversity and subjectivity is ensured,

▲ statistical conclusions are reliable and no relevant information is generally lost.

Among the *cons*, one can mention that

▼ the choice of the most appropriate point is not a simple task, and it does not seem realistic to demand as much accuracy in connection with such an intrinsically imprecise context;

▼ surveys/questionnaires based on the visual analogue scale require a special framework and either a paper-and-pencil or a computer/web-based form to be filled.

Benoit [3] (see also Benoit and Foulloy [4]) asserts that, from a measurement point of view, the fuzzy scale "establishes a link between strongly defined measurements ... and weakly defined measurements". In this respect, Calcagnì and Lombardi [8] (see also De la Rosa de Sáa *et al.* [15,14]) indicate that fuzzy scales have been applied to overcome the limitations of standard scales by modeling the imprecision of human rating evaluations. Two approaches should be mainly distinguished in applying fuzzy scales, namely,

- the *fuzzy linguistic scales*, which are frequently considered for different goals as an *a posteriori* tool to encode data from a discrete (often a Likert) scale by means of fuzzy numbers (see, for instance, Zadeh [67], Tong and Bonissone [59], Pedrycz [42], Herrera *et al.* [27,26], Lalla *et al.* [32], and also Li [33], Akdag *et al.* [1], Estrella *et al.* [18], Massanet *et al.* [39], Tejeda-Lorente *et al.* [57,58], Villacorta *et al.* [62], Wang *et al.* [63], García-Galán *et al.* [19], Liu *et al.* [34] and Tavana [56], about some very recent developments and applications in connection with perceived quality, satisfaction, etc.);
- the *fuzzy rating scale*, which is considered as an *a priori* tool to directly assess fuzzy values and integrating the continuous nature and free assessment of the visual analogue scales with the ability to cope with imprecision of the fuzzy linguistic ones; this scale has been introduced by Hesketh *et al.* [30] (see also, among others, Hesketh and Hesketh [29], Matsui and Takeya [40], Takemura [53–55], Yamashita [65], Hesketh *et al.* [28] and De la Rosa de Sáa *et al.* [14] for some developments and applications).

The Likert, visual analogue and fuzzy linguistic scales have been commonly involved in research with questionnaires. The fuzzy rating scale has been applied too, but only occasionally in spite of the clear advantages associated with its use that will be detailed in the next section.

As we know, the success of a questionnaire method based on a certain scale depends on the reliability and easy-to-handle use of such a scale. Reasons why the fuzzy rating scale is not so popular yet can probably be found in the following critical requirements:

▽ a certain framework is needed to conduct a fuzzy rating scale-based questionnaire (e.g., it cannot be properly conducted by phone, on the street, etc.);

▽ respondents need a certain training to answer a fuzzy rating scale-based questionnaire;

▽ a special statistical methodology is needed to analyze responses from a fuzzy rating scale-based questionnaire.

Although these requirements could prevent us from applying the fuzzy rating scale, this paper aims to show by means of a real-life example and to a descriptive extent that, actually, they are not substantial drawbacks. As its ultimate goal, it also aims to show that statistical conclusions often differ depending on the considered scale, in the same way that grouping real-valued data by intervals often leads to different conclusions because of many differences being hidden through the grouping.

For this aim, Section 2 recalls the ideas behind the fuzzy rating scale and the way of applying it to questionnaires, along with the basic tools to handle responses for the descriptive data analysis. Section 3 first presents the real-life example, detailing its design and implementation. A key point in the application is the one related to the fact that responses to most of the posed questions have been given in accordance with both a fuzzy rating scale and a 4-point Likert one. Thanks to this double response, we later illustrate some of the relevant descriptive measures to analyze fuzzy data that have been previously introduced, by means of the analysis of the two types of responses collected from the real-life example. The results from this double analysis are summarized, corroborating that statistical conclusions differ depending on the considered scale, and the practical implications from this fact are discussed. The paper ends by commenting some related future research directions in Section 4.

2 Preliminary tools

Responses from fuzzy linguistic/rating scale-based questionnaires are assumed to be fuzzy numbers.

A (bounded) fuzzy number is an ill-defined quantity or value which is characterized by means of a mapping $\widetilde{U} : \mathbb{R} \to [0,1]$ such that it can be formalized by either of the following:

- (VERTICAL VIEW) it is an upper semicontinuous, quasi-concave normal function with a bounded support set, $\operatorname{supp}\widetilde{U} = \{x \in \mathbb{R} : \widetilde{U}(x) > 0\}$, and $\widetilde{U}(x)$ meaning the 'degree of compatibility' of x with the property describing $\widetilde{U}$;
- (HORIZONTAL VIEW) for each $\alpha \in [0,1]$, its α-level set

$$\widetilde{U}_\alpha = \begin{cases} \{x \in \mathbb{R} : \widetilde{U}(x) \geq \alpha\} & \text{if } \alpha \in (0,1] \\ \text{closure}(\operatorname{supp}\widetilde{U}) & \text{if } \alpha = 0 \end{cases}$$

  (i.e., the set of real numbers which are compatible with $\widetilde{U}$ to a degree over α) is a nonempty closed and bounded interval.

As an example of fuzzy number which is used in many different studies one can refer to *trapezoidal fuzzy numbers*, which are given by $\widetilde{U} = \mathrm{Tra}(a,b,c,d)$ $= \mathrm{Tra}(\inf\widetilde{U}_0, \inf\widetilde{U}_1, \sup\widetilde{U}_1, \sup\widetilde{U}_0)$ such that

$$\widetilde{U}_\alpha = [a + \alpha(b-a), c + (1-\alpha)(d-c)] \text{ for each } \alpha \in [0,1]$$

(see Figure 1 for the graphical display of one of these fuzzy numbers).

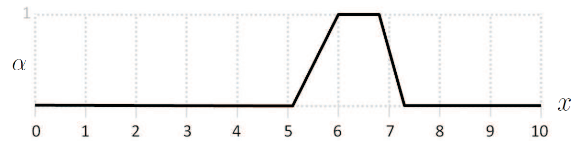


Fig. 1. Graphical display of the trapezoidal fuzzy number Tra(5.1, 6, 6.75, 7.3)

2.1 The fuzzy rating scale

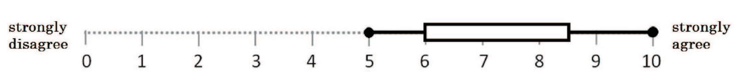
To really exploit the individual differences in responding to questionnaires, and gain reliability by increasing the number of possible responses (see, for instance, Lozano *et al.* [35]), one should consider a rich and expressive scale in which "something can be meaningful although we cannot name it" (Ghneim [20]). As Zadeh [68] argues, "Paradoxically, one of the principal contributions of fuzzy logic, a contribution which is widely unrecognized, is its high power of precisiation of what is imprecise. This capability of fuzzy logic suggests ... that it may find important applications in the realms of ... human-centric fields." The fuzzy rating scale certainly allows respondents to 'precisiate' answers in a continuous way, with infinite possible nuances, as well as to develop mathematical computations with these responses.

The guideline for the mechanism to draw the value that better expresses a response according to the fuzzy rating scale (Hesketh *et al.* [30]) is as follows:

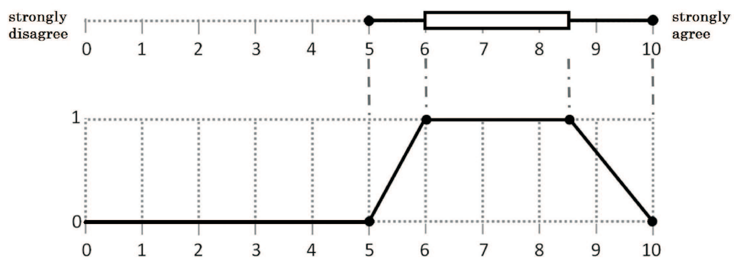
Step 1. A reference bounded interval/segment is first considered. This is often chosen to be $[0, 10]$ or $[0, 100]$, but the choice of the intervals is not at all a constraint. The end-points are often labeled in accordance with their meaning referring to the degree of agreement, satisfaction, quality, and so on.

Step 2. The *core*, or 1-level set, associated with the response is determined. It corresponds to the interval consisting of the real values within the reference one which are considered to be as 'fully compatible' with the response.

Step 3. The *support*, or its closure or 0-level set, associated with the response is determined. It corresponds to the interval consisting of the real values within the referential that are considered to be as 'compatible to some extent' with the response.



Step 4. The two intervals are 'linearly interpolated' to get a trapezoidal fuzzy number.



The main positive features in connection with the use of the fuzzy rating scale are that

▲ fuzzy rating scale values properly capture the inherent imprecision when rating this type of characteristics or attributes,

▲ this scale means a continuum, and the transition from a value to another is gradual and flexible,

▲ this scale is much richer and more expressive than any one based on a (unavoidable finite) natural language or its real/fuzzy-valued encoding,

▲ its flexibility allows us to capture individual differences, whence the natural associated variability, diversity and subjectivity are not lost,

▲ values from this scale can be well handled mathematically and computationally, and one can state arithmetic and distances preserving the meaning of fuzzy numbers and extend many statistical concepts and developments.

On the 'negative side' we have already pointed out that

▼ fuzzy rating scale-based questionnaires should be conducted in an appropriate framework; in this respect, as it happens for the visual analogue scale, one cannot answer such questionnaires by simply speaking, but a paper-and-pencil or computerized form is required to be filled;

▽ respondents should be appropriately trained; although this assertion does not need to be discussed since the real-life example to be detailed in the next section clearly illustrates that the required background to use the fuzzy rating scale is not that demanding (just knowing the notions of triangle/trapezium and the intuitive idea of compatibility);

▽ a specific methodology is required if we wish to analyze responses from a statistical perspective; in connection with this, during the last years several statistical notions and procedures have been developed to analyze, both descriptively and inferentially, fuzzy number-valued data (see Blanco-Fernández *et al.* [6,7] for a recent review), considering each fuzzy datum as a whole, so no relevant information is lost.

Remark 2.1 Although the fuzzy rating scale has been stated on the basis of trapezoidal (or even, in particular, triangular) fuzzy numbers, neither the ideas behind nor the statistical methodology to analyze data make this statement essential. However, such a simple shape, which can be naturally identified with the notion of trapezium, makes the fuzzy rating scale easy to explain and apply, almost irrespective of the background of respondents. Furthermore, their use is also supported by arguments provided by Pedrycz [43], Grzegorzewski [24], Grzegorzewski and Pasternak-Winiarska [25], Ban *et al.* [2], and others, who have soundly discussed how triangular and trapezoidal fuzzy numbers can be considered to describe or approximate fuzzy data. Recently, an additional supporting argument has been given through an empirical sensitivity analysis carried out about the (to some extent descriptive) effect of the shape of fuzzy data on several statistical measures (see Lubiano *et al.* [36]).

2.2 Arithmetic, metrics and modeling of fuzzy-valued random elements

Let $\mathcal{F}_c(\mathbb{R})$ denote the space of (bounded) fuzzy numbers. To develop statistics with fuzzy number-valued data, computations often involve two operations from the usual fuzzy arithmetic based on Zadeh's extension principle [67]:

- If $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$, then the *sum of* $\widetilde{U}$ *and* $\widetilde{V}$ is defined as the fuzzy number $\widetilde{U} + \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$ such that for each $\alpha \in [0,1]$
$$(\widetilde{U}+\widetilde{V})_\alpha = \left[\inf \widetilde{U}_\alpha + \inf \widetilde{V}_\alpha, \sup \widetilde{U}_\alpha + \sup \widetilde{V}_\alpha\right],$$
and it can be easily checked that
$$\mathrm{Tra}(a,b,c,d) + \mathrm{Tra}(a',b',c',d') = \mathrm{Tra}(a+a', b+b', c+c', d+d').$$
- If $\widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ and $\gamma \in \mathbb{R}$, the *product of* $\widetilde{U}$ *by the scalar* γ is defined as $\gamma \cdot \widetilde{U} \in \mathcal{F}_c(\mathbb{R})$ such that for each $\alpha \in [0,1]$
$$(\gamma \cdot \widetilde{U})_\alpha = \gamma \cdot \widetilde{U}_\alpha = \left\{\gamma \cdot y \, : \, y \in \widetilde{U}_\alpha\right\} = \begin{cases} \left[\gamma \inf \widetilde{U}_\alpha, \gamma \sup \widetilde{U}_\alpha\right] & \text{if } \gamma \geq 0, \\ \left[\gamma \sup \widetilde{U}_\alpha, \gamma \inf \widetilde{U}_\alpha\right] & \text{if } \gamma < 0, \end{cases}$$
and it can be easily checked that
$$\gamma \cdot \mathrm{Tra}(a,b,c,d) = \begin{cases} \mathrm{Tra}(\gamma\, a, \gamma\, b, \gamma\, c, \gamma\, d) & \text{if } \gamma \geq 0, \\ \mathrm{Tra}(\gamma\, d, \gamma\, c, \gamma\, b, \gamma\, a) & \text{if } \gamma < 0. \end{cases}$$

As for the visual analogue scale, an interesting feature of the fuzzy rating one is that we can consider nice metric structures. This becomes especially useful for statistical purposes in this setting, since in general $\widetilde{U} + (-1) \cdot \widetilde{U}$ $\neq$ indicator function of $\{0\}$ (the neutral element for the fuzzy sum). As a consequence, it is not possible to establish a well-defined difference between fuzzy numbers preserving all the properties of the difference of real numbers. Among the distances between fuzzy numbers, the following have been shown to be well-adapted for statistical developments:

- If $\widetilde{U}, \widetilde{V} \in \mathcal{F}_c(\mathbb{R})$,
$$\rho_1(\widetilde{U},\widetilde{V}) = \frac{1}{2}\int_{[0,1]} \left(\left|\inf \widetilde{U}_\alpha - \inf \widetilde{V}_\alpha\right| + \left|\sup \widetilde{U}_\alpha - \sup \widetilde{V}_\alpha\right|\right) d\alpha$$
and
$$\rho_2(\widetilde{U},\widetilde{V}) = \sqrt{\frac{1}{2}\int_{[0,1]} \left(\left[\inf \widetilde{U}_\alpha - \inf \widetilde{V}_\alpha\right]^2 + \left[\sup \widetilde{U}_\alpha - \sup \widetilde{V}_\alpha\right]^2\right) d\alpha}$$
are said to be, respectively, the *1-norm distance* and *2-norm distance* between $\widetilde{U}$ and $\widetilde{V}$ (see Diamond and Kloeden [17]) .

In case $\widetilde{U}$ and $\widetilde{V}$ are trapezoidal fuzzy numbers, the squared distance ρ_2 reduces to

$$\left[\rho_2(\widetilde{U},\widetilde{V})\right]^2$$
$$=\frac{1}{6}\left[(\inf \widetilde{U}_0-\inf \widetilde{V}_0)^2+(\inf \widetilde{U}_1-\inf \widetilde{V}_1)^2+(\inf \widetilde{U}_0-\inf \widetilde{V}_0)(\inf \widetilde{U}_1-\inf \widetilde{V}_1)\right]$$
$$+\frac{1}{6}\left[(\sup \widetilde{U}_0-\sup \widetilde{V}_0)^2+(\sup \widetilde{U}_1-\sup \widetilde{V}_1)^2+(\sup \widetilde{U}_0-\sup \widetilde{V}_0)(\sup \widetilde{U}_1-\sup \widetilde{V}_1)\right].$$

- If $\widetilde{U},\widetilde{V}\in\mathcal{F}_c(\mathbb{R})$

$$D(\widetilde{U},\widetilde{V})=\sqrt{\int_{[0,1]}\int_{[0,1]}\left(\left[\widetilde{U}_\alpha^{[\lambda]}-\widetilde{V}_\alpha^{[\lambda]}\right]^2\right)\,d\alpha\,d\lambda},$$

  with $\widetilde{U}_\alpha^{[\lambda]}=\lambda\cdot\sup\widetilde{U}_\alpha+(1-\lambda)\cdot\inf\widetilde{U}_\alpha$ is said to be (see Bertoluzza *et al.* [5], Trutschnig *et al.* [61] for arguments and details about) the (1/3) *mid/spr distance* between $\widetilde{U}$ and $\widetilde{V}$.

  In case $\widetilde{U}$ and $\widetilde{V}$ are trapezoidal fuzzy numbers, the squared distance D reduces to

$$\left[D(\widetilde{U},\widetilde{V})\right]^2$$
$$=\frac{1}{3}\Big[(\mathrm{mid}_0\,\widetilde{U}-\mathrm{mid}_0\,\widetilde{V})^2+(\mathrm{mid}_1\,\widetilde{U}-\mathrm{mid}_1\,\widetilde{V})^2$$
$$+(\mathrm{mid}_0\,\widetilde{U}-\mathrm{mid}_0\,\widetilde{V})(\mathrm{mid}_1\,\widetilde{U}-\mathrm{mid}_1\,\widetilde{V})\Big]$$
$$+\frac{1}{9}\Big[(\mathrm{spr}_0\,\widetilde{U}_0-\mathrm{spr}_0\,\widetilde{V})^2+(\mathrm{spr}_1\,\widetilde{U}-\mathrm{spr}_1\,\widetilde{V})^2$$
$$+(\mathrm{spr}_0\,\widetilde{U}-\mathrm{spr}_0\,\widetilde{V})(\mathrm{spr}_1\,\widetilde{U}-\mathrm{spr}_1\,\widetilde{V})\Big],$$

  where $\mathrm{mid}_0/\mathrm{mid}_1$ and $\mathrm{spr}_0/\mathrm{spr}_1$ denote, respectively, the centre (mid-point) and the spread (radius) of the core/support of the fuzzy number.

To develop a well-stated methodology to analyze fuzzy data, we need a formal model for the random mechanism generating fuzzy number-valued data. This model should integrate randomness (to generate data) and fuzziness (because of the intrinsic nature of these data). Random fuzzy sets (originally coined as fuzzy random variables by Puri and Ralescu [45]) result in a well-defined and sound model within the probabilistic setting, what allows us to extend or preserve almost all the basis (although extending statistical methods for the analysis of fuzzy data is not a straightforward task).

Given a random experiment mathematically modeled by means of a probability space $(\Omega,\mathcal{A},P)$, a *random fuzzy number* (or one-dimensional random fuzzy set, for short *RFN*) associated with it is a mapping $\mathcal{X}:\Omega\to\mathcal{F}_c(\mathbb{R})$ such that, for all $\alpha\in[0,1]$, the α-level mapping $\mathcal{X}_\alpha$ is a compact random interval (that is, for all $\alpha\in[0,1]$ the real-valued mappings $\inf\mathcal{X}_\alpha$ and $\sup\mathcal{X}_\alpha$ are random variables). In case we deal with trapezoidal-valued random fuzzy numbers, the sufficient and necessary condition is that the real-valued mappings $\inf\mathcal{X}_0$, $\inf\mathcal{X}_1$, $\sup\mathcal{X}_1$ and $\sup\mathcal{X}_0$ are random variables.

One can prove (see, for instance, Colubi *et al.* [11]) that $\mathcal{X}:\Omega\to\mathcal{F}_c(\mathbb{R})$ is an RFN if and only if it is a Borel-measurable mapping with respect to a certain σ-field and the above mentioned metrics, which enable us to properly refer to the induced distribution of an RFN.

2.3 Relevant descriptive summary measures of a fuzzy dataset

In summarizing the distribution of an RFN over a sample of individuals, some of the best know measures extend the mean and median (central tendency/location measures) and the variance and median absolute deviation (dispersion/scale measures) of real-valued random variables.

Assume that over the sample of individuals $(\omega_1, \ldots, \omega_n)$ the RFN $\mathcal{X}$ provides us with the sample of fuzzy number-valued data $\widetilde{\boldsymbol{x}_n} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$ (i.e., $\widetilde{x}_i = \mathcal{X}(\omega_i)$, $i = 1, \ldots, n$).

- SAMPLE AUMANN-TYPE MEAN:

  Following the definition by Puri and Ralescu [45], and particularizing it to trapezoidal fuzzy data, the *sample Aumann-type mean* is defined as the fuzzy number $\overline{\widetilde{\boldsymbol{x}_n}} \in \mathcal{F}_c(\mathbb{R})$ given by

$$\overline{\widetilde{\boldsymbol{x}_n}} = \mathrm{Tra}\left(\frac{1}{n}\sum_{i=1}^{n}\inf(\widetilde{x}_i)_0, \frac{1}{n}\sum_{i=1}^{n}\inf(\widetilde{x}_i)_1, \frac{1}{n}\sum_{i=1}^{n}\sup(\widetilde{x}_i)_1, \frac{1}{n}\sum_{i=1}^{n}\sup(\widetilde{x}_i)_0\right).$$

- SAMPLE 1-NORM MEDIAN:

  Following Sinova *et al.* [51], and particularizing it to trapezoidal fuzzy data, the *sample 1-norm median* is defined as the fuzzy number $\widetilde{\mathrm{Me}}(\widetilde{\boldsymbol{x}_n}) \in \mathcal{F}_c(\mathbb{R})$, such that for all $\alpha \in [0,1]$

$$\left(\widetilde{\mathrm{Me}}(\widetilde{\boldsymbol{x}_n})\right)_\alpha$$
$$= \Big[\mathrm{Me}\Big\{\alpha \inf(\widetilde{x}_1)_1 + (1-\alpha)\inf(\widetilde{x}_1)_0, \ldots, \alpha \inf(\widetilde{x}_n)_1 + (1-\alpha)\inf(\widetilde{x}_n)_0\Big\},$$
$$\mathrm{Me}\Big\{\alpha \sup(\widetilde{x}_1)_1 + (1-\alpha)\sup(\widetilde{x}_1)_0, \ldots, \alpha \sup(\widetilde{x}_n)_1 + (1-\alpha)\sup(\widetilde{x}_n)_0\Big\}\Big],$$

  $\mathrm{Me}\{\cdot\}$ denoting the median of the corresponding real-valued dataset, with the usual convention of using the mid-point of possible medians in case it is not unique.

  The practical computation of $\widetilde{\mathrm{Me}}(\widetilde{\boldsymbol{x}_n})$ cannot be generally simplified, even for trapezoidal fuzzy data. This is due to the fact that, in contrast to what happens for the L^2 metrics ρ_2 and D, the expression of ρ_1 cannot be especially simplified for trapezoidal data. Actually, except from very especial cases, the sample 1-norm median is usually approximated by computing for each of a large number of equidistant levels (in the examples below, this number is taken to be 101) the infima and suprema of the α-levels and finally rebuilding the fuzzy number from the intervals they determine. Computations involved in the real-life example have been carried out by using R functions. Other location measures have been defined (see Sinova *et al.* [49,52]), although they usually involve more computational tasks.

- SAMPLE FRÉCHET-TYPE VARIANCE:

  Following Lubiano *et al.* [37], the *sample D-Fréchet-type variance* particularized to trapezoidal fuzzy data is defined as the real number

$$s^2(\widetilde{\boldsymbol{x_n}}) = \frac{1}{n-1}\sum_{i=1}^{n}\left[D(\widetilde{x}_i, \overline{\widetilde{\boldsymbol{x_n}}})\right]^2$$

$$= \frac{1}{3}\Big[s^2(\mathrm{mid}_0\,\widetilde{x}_1,\ldots,\mathrm{mid}_0\,\widetilde{x}_n) + s^2(\mathrm{mid}_1\,\widetilde{x}_1,\ldots,\mathrm{mid}_1\,\widetilde{x}_n) + s\Big((\mathrm{mid}_0\,\widetilde{x}_1,\ldots,\mathrm{mid}_0\,\widetilde{x}_n),(\mathrm{mid}_1\,\widetilde{x}_1,\ldots,\mathrm{mid}_1\,\widetilde{x}_n)\Big)\Big]$$

$$+\frac{1}{9}\Big[s^2(\mathrm{spr}_0\,\widetilde{x}_1,\ldots,\mathrm{spr}_0\,\widetilde{x}_n) + s^2(\mathrm{spr}_1\,\widetilde{x}_1,\ldots,\mathrm{spr}_1\,\widetilde{x}_n) + s\Big((\mathrm{spr}_0\,\widetilde{x}_1,\ldots,\mathrm{spr}_0\,\widetilde{x}_n),(\mathrm{spr}_1\,\widetilde{x}_1,\ldots,\mathrm{spr}_1\,\widetilde{x}_n)\Big)\Big]$$

with

$$s^2(a_1,\ldots,a_n) = \frac{1}{n-1}\sum_{i=1}^{n}(a_i-\overline{a})^2, \qquad \overline{a} = \frac{1}{n}\sum_{i=1}^{n}a_i,$$

$$s((a_1,\ldots,a_n),(b_1,\ldots,b_n)) = \frac{1}{n-1}\sum_{i=1}^{n}(a_i-\overline{a})\cdot(b_i-\overline{b}).$$

- SAMPLE MEDIAN 1-NORM DEVIATION:

  Following De la Rosa de Sáa *et al.* [13,16], the *sample median 1-norm deviation* is defined as the real number ρ_1-MDD($\widetilde{\boldsymbol{x_n}}$) such that

  $$\rho_1\text{-MDD}(\widetilde{\boldsymbol{x_n}}) = \mathrm{Me}\left\{\rho_1\Big(\widetilde{x}_1, \widetilde{\mathrm{Me}}(\widetilde{\boldsymbol{x_n}})\Big),\ldots,\rho_1\Big(\widetilde{x}_n, \widetilde{\mathrm{Me}}(\widetilde{\boldsymbol{x_n}})\Big)\right\},$$

  with the usual convention of using the mid-point of possible medians Me$\{\cdot\}$ in case it is not unique.

  As for computing the 1-norm median, the practical computations for the real-life example to be detailed later have been carried out by using R functions.

In addition to these summary measures, we are going to consider two additional ones. The first one refers to the diversity of values, whereas the second one establishes an indicator of the similarity of the distributions of two fuzzy datasets.

Assume that over the sample of individuals $(\omega_1,\ldots,\omega_n)$ the RFN $\mathcal{X}$ takes on k different values with sample relative frequencies f_j, $j = 1,\ldots,k$.

- SAMPLE GINI-SIMPSON DIVERSITY INDEX:

  Following Gini [23] and Simpson [47] (who adapted Gini's index in some biological developments to quantify the diversity of answers), the *sample Gini-Simpson quadratic index* is the real number given by

  $$\mathrm{Div}_{\mathrm{Gini}}(\widetilde{\boldsymbol{x_n}}) = 1 - \sum_{j=1}^{k} f_j^2.$$

  It should be noticed that the diversity index only takes into account the different values the random element takes on in the sample and their associated sample frequencies, irrespectively of the nature of these values. Consequently, this summary measure does not need a special adaptation.

Another popular diversity index is the one based on Shannon's entropy. However, as it has been proved for different random samplings (see Pérez *et al.* [44], Caso and Gil [10] and Gil and Gil [21]), the Gini-Simpson index can be unbiasedly estimated, whereas Shannon's one cannot.

Recently, Sinova *et al.* [48] have introduced the fuzzy characterizing function of the distribution of a random fuzzy number as an extension of the moment-generating function of a real-valued random variable. On the basis of its properties, one can state a descriptive index of the dissimilarity between the distributions of two fuzzy datasets.

- SAMPLE DISSIMILARITY BETWEEN DISTRIBUTIONS:

  Consider two sample fuzzy datasets. Assume that the first dataset is associated with a (trapezoidal-valued) RFN $\mathcal{X}$ and it is denoted by $\widetilde{\boldsymbol{x_n}} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)$, the second one being associated with a (trapezoidal-valued) RFN $\mathcal{Y}$ and denoted by $\widetilde{\boldsymbol{y_m}} = (\widetilde{y}_1, \ldots, \widetilde{y}_m)$.

  Following the definition of the fuzzy characterizing function of the distribution of an RFN by Sinova *et al.* [48], for a fixed small value $\varepsilon > 0$, the *ε-sample dissimilarity between the distributions of $\widetilde{\boldsymbol{x_n}}$ and $\widetilde{\boldsymbol{y_m}}$* is given by the index based on the ρ_2 distance between the corresponding sample fuzzy characterizing functions, which can be proven to be given (up to the factor $1/\varepsilon$) by

$$
\begin{aligned}
&\varrho_\varepsilon(\widetilde{\boldsymbol{x_n}}, \widetilde{\boldsymbol{y_m}}) \\
&= \frac{1}{\varepsilon} \max_{t\in[-\varepsilon,\varepsilon]} \left(\frac{1}{2n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \frac{e^{t[\inf(\widetilde{x}_i)_1 + \inf(\widetilde{x}_{i'})_1]} - e^{t[\inf(\widetilde{x}_i)_0 + \inf(\widetilde{x}_{i'})_0]}}{t[\inf(\widetilde{x}_i)_1 + \inf(\widetilde{x}_{i'})_1] - t[\inf(\widetilde{x}_i)_0 + \inf(\widetilde{x}_{i'})_0]} \right. \\
&+ \frac{1}{2m^2} \sum_{j=1}^{m} \sum_{j'=1}^{m} \frac{e^{t[\inf(\widetilde{y}_j)_1 + \inf(\widetilde{y}_{j'})_1]} - e^{t[\inf(\widetilde{y}_j)_0 + \inf(\widetilde{y}_{j'})_0]}}{t[\inf(\widetilde{y}_j)_1 + \inf(\widetilde{y}_{j'})_1] - t[\inf(\widetilde{y}_j)_0 + \inf(\widetilde{y}_{j'})_0]} \\
&- \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{e^{t[\inf(\widetilde{x}_i)_1 + \inf(\widetilde{y}_j)_1]} - e^{t[\inf(\widetilde{x}_i)_0 + \inf(\widetilde{y}_j)_0]}}{t[\inf(\widetilde{x}_i)_1 + \inf(\widetilde{y}_j)_1] - t[\inf(\widetilde{x}_i)_0 + \inf(\widetilde{y}_j)_0]} \\
&+ \frac{1}{2n^2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \frac{e^{t[\sup(\widetilde{x}_i)_1 + \sup(\widetilde{x}_{i'})_1]} - e^{t[\sup(\widetilde{x}_i)_0 + \sup(\widetilde{x}_{i'})_0]}}{t[\sup(\widetilde{x}_i)_1 + \sup(\widetilde{x}_{i'})_1] - t[\sup(\widetilde{x}_i)_0 + \sup(\widetilde{x}_{i'})_0]} \\
&+ \frac{1}{2m^2} \sum_{j=1}^{m} \sum_{j'=1}^{m} \frac{e^{t[\sup(\widetilde{y}_j)_1 + \sup(\widetilde{y}_{j'})_1]} - e^{t[\sup(\widetilde{y}_j)_0 + \sup(\widetilde{y}_{j'})_0]}}{t[\sup(\widetilde{y}_j)_1 + \sup(\widetilde{y}_{j'})_1] - t[\sup(\widetilde{y}_j)_0 + \sup(\widetilde{y}_{j'})_0]} \\
&\left. - \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{e^{t[\sup(\widetilde{x}_i)_1 + \sup(\widetilde{y}_j)_1]} - e^{t[\sup(\widetilde{x}_i)_0 + \sup(\widetilde{y}_j)_0]}}{t[\sup(\widetilde{x}_i)_1 + \sup(\widetilde{y}_j)_1] - t[\sup(\widetilde{x}_i)_0 + \sup(\widetilde{y}_j)_0]} \right)^{1/2}.
\end{aligned}
$$

  On the basis of the results in Sinova *et al.* [48], it can be concluded that this index vanishes if and only if the two sample distributions coincide.

  The correction factor $1/\varepsilon$ is not really relevant at all, but it can be considered for descriptive purposes to make differences more evident. Thus, the choice of very small values of ε usually leads to very small distances, irrespectively of the sample distributions being more or less close.

3 Descriptive data analysis of a real-life example

When one attempts to develop descriptive statistics with fuzzy data, some distinctive features with respect to the real-valued case should be taken into account, namely: one cannot make use of a difference operator which is well-defined and preserves all the properties from the real-valued case, and there is no 'universally accepted' total order between fuzzy numbers. The measures in the preceding section overcome these eventual concerns with the use of appropriate metrics.

To illustrate the suggested descriptive approach, we first present a real-life example involving fuzzy rating scale-based data. Aiming to compare the values of the descriptive measures in the preceding section with their counterpart(s) for Likert-type data, and to descriptively conclude about the differences between the use of the scales, the example involves two types of responses.

3.1 Real-life application

The study to be described is related to the well-known questionnaire TIMSS-PIRLS 2011, which was the result of the union in 2011 of the TIMSS (Trends in International Mathematics and Science Study) and PIRLS (Progress in International Reading Literacy Study) to jointly assess the same fourth grade students in mathematics, science and reading, and to have the most appropriate basis for studying the relationships among these abilities. This was possible because TIMSS and PIRLS questionnaires were conducted on the same students, with some additional questionnaires having been filled by their parents, teachers and school management team. Managing data on the same students makes possible to perform valuable investigations and researchers can apply a variety of modeling techniques to explore these important issues.

In 2011, the Spanish Institute of Educational Evaluation (INEE) commissioned some members of our Department at the University of Oviedo (Spain) to develop a data analysis with the data collected through some of the TIMSS/PIRLS questionnaires conducted in Spanish schools (see Corral *et al.* [12] for a summary of conclusions). These questionnaires are standard and most of the involved questions have to be answered according to the 4-point Likert scale given by A1 = DISAGREE A LOT, A2 = DISAGREE A LITTLE, A3 = AGREE A LITTLE and A4 = AGREE A LOT.

These studies have been interesting, but our colleagues have been contemplating whether the use of a fuzzy rating scale approach for the responses would yield somewhat different statistical conclusions or enlighten about some apparently curious conclusions. This work will focus on the first fact. To carry out an introductory analysis about the likely differences, we have chosen the following nine questions (three per subject) from the Student questionnaire, as gathered in Table 1.

Table 1
Questions selected from the TIMSS-PIRLS 2011 Student Questionnaires

	READING IN SCHOOL
$R.1$	I like to read things that make me think
$R.2$	I learn a lot from reading
$R.3$	Reading is harder for me than any other subject
	MATHEMATICS IN SCHOOL
$M.1$	I like mathematics
$M.2$	My teacher is easy to understand
$M.3$	Mathematics is harder for me than any other subject
	SCIENCE IN SCHOOL
$S.1$	My teacher taught me to discover science in daily life
$S.2$	I read about science in my spare time
$S.3$	Science is harder for me than any other subject

The questionnaire form involving these nine questions, along with a few more questions related to the facilities to study the students enjoy at home, has been adapted to allow a double-type response, namely, the original Likert and a fuzzy rating scale-based one (see Figure 2 for Question $M.2$, and the webpage http://bellman.ciencias.uniovi.es/SMIRE/FuzzyRatingScaleQuestionnaire-SanIgnacio.html for the full paper-and-pencil and computerized datasets).

In this way, each of the nine questions in Table 1 is assumed to be filled in accordance with both the 4-point Likert scale and the fuzzy rating one with reference interval $[0, 10]$.

To ease the relationship between the two scales, each numerically encoded Likert response, re-scaled to $[0, 10]$, has been superimposed upon the reference interval of the fuzzy rating scale part.

The questionnaire involving these double-response questions has been conducted on 69 fourth grade students from Colegio San Ignacio (Oviedo-Asturias, Spain).

A preliminary analysis has been considered in Gil *et al.* [22], and this paper aims to enter in more detail and complete it with the ultimate goal of corroborating that statistical results often vary with the scale employed. Indeed, this is something one can immediately expect from the fact that the fuzzy rating scale is much richer and more expressive, and it captures a higher subjectivity

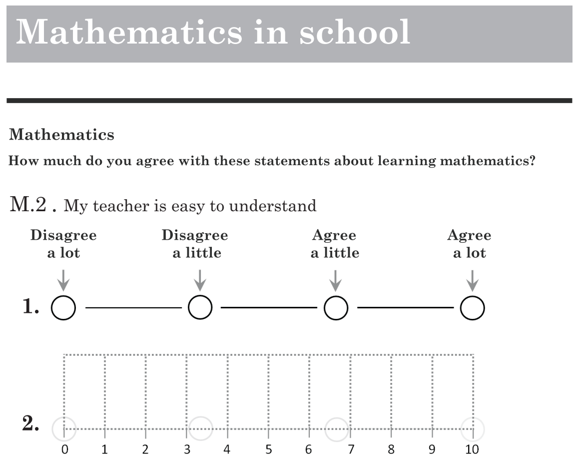


Fig. 2. Example of the double-response form to a question

and variability in responding, than Likert ones.

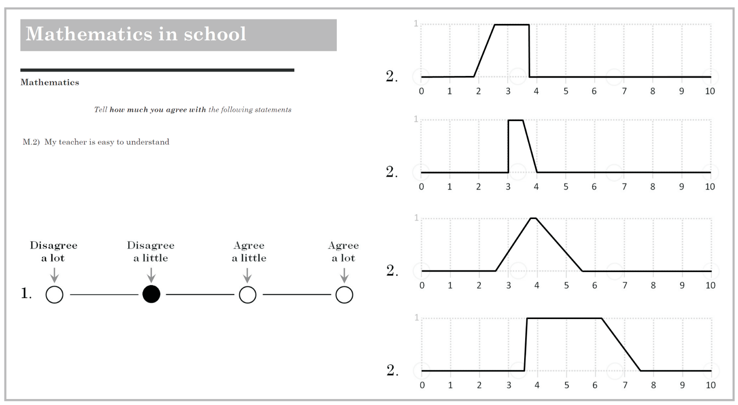


Fig. 3. Example of 4 double responses to Question $M.2$ for which the Likert-type ones coincide while the fuzzy rating scale-type clearly differ

To support and illustrate this last assertion, one can consider a combined graphical display of the double response to Question $M.2$ for which the Likert scale-based response chosen by four students has corresponded to A2 = DISAGREE A LITTLE, and the fuzzy rating scale-based responses for the same students have been definitely different (see Figure 3).

Because of the 4th grade students at Colegio San Ignacio being grouped in three different classrooms, three of us have conducted the questionnaire with the support of a teacher. Two groups have used the computer-administered format, whereas the other one has filled the paper-and-pencil one.

The training of the nine-year-old children has taken up to fifteen minutes. Since they ignored what real-valued functions mean, we have taken advantage of their knowledge of the notion of trapezium already at this stage, what has been sufficient to make them understand the meaning of the upper and lower bases in the fuzzy rating.

Along the training, some students asked about the possibility of using triangles or singletons as responses to the fuzzy rating scale-based questionnaire. Actually and surprisingly, there have been more missing responses with the Likert than with the fuzzy rating scale.

As it will be explained later, the obtained responses have been quite reasonable, which allows us to conclude that **the requirements for training do not mean an important handicap**.

3.2 Descriptive data analysis of the real-life example: a comparative view

The complete datasets for the study can be found in http://bellman.ciencias.uniovi.es/SMIRE/FuzzyRatingScaleQuestionnaire-SanIgnacio.html. In addition to Likert and fuzzy rating scales (for short FRS) datasets, in the statistical analysis in this section we will consider the datasets obtained when the Likert-type responses are

- encoded by means of the re-scaled to $[0, 10]$ usual numerical encoding (i.e., $A1 \equiv 0$, $A2 \equiv 10/3$, $A3 \equiv 20/3$ and $A4 \equiv 10$), for short NELikert, and
- encoded by means of the re-scaled to $[0, 10]$ set of four terms with its usual semantics (see Figure 4), for short FLELikert.

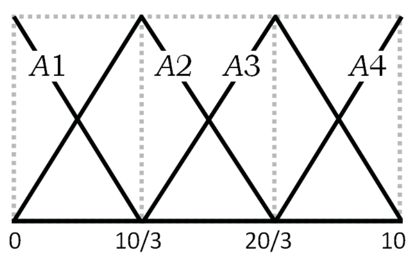


Fig. 4. A usual fuzzy linguistic encoding of a 4-point Likert scale

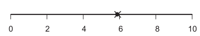
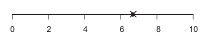
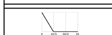
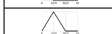
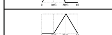
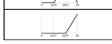
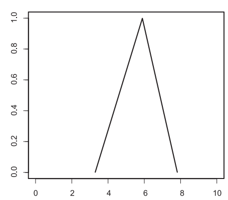
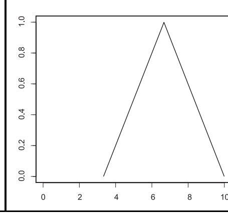
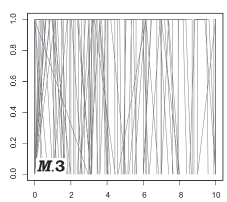
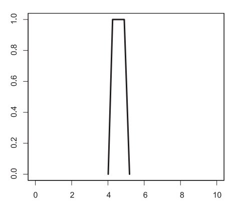
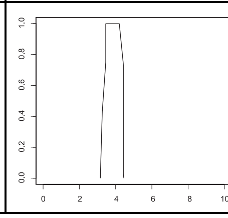
Other balanced and unbalanced fuzzy linguistic encodings have been considered, like those in Calcagnì and Lombardi [8], Villacorta *et al.* [62], and Wang *et al.* [63]. However, we have validated different encodings through the percentage of discrete responses matching with the closest one in ρ_2's sense, and the one in Figure 4 has been the one yielding over 80% of matches.

3.2.1 Analyzing location and dispersion of the datasets

In summarizing the distribution of the sample of responses from the questionnaire, we are first going to estimate their sample Aumann-type mean, 1-norm median, Fréchet-type variance and median 1-norm deviation, all of them leading to strongly consistent estimators of the corresponding population parameters. For each question only data from the students who have provided with both the Likert-type and the FRS responses have been considered in the comparative analysis.

The only measure making sense for Likert responses is the median (and it could be trivially deduced from the one for the NELikert data), so we will only consider the NELikert, the FLELikert and the FRS data.

Table 2
Sample mean, median, variance and median distance deviation for NELikert responses to Question $M.3$ (1st row); sample Aumann-type mean, 1-norm median, D Fréchet-type variance and ρ_1-MDD for FLELikert responses to Question $M.3$ (2nd row); and for FRS responses to Question $M.3$ (3rd row)

NELikert sample data	NELikert mean	NELikert median	NELikert variance	NELikert MDD
$M.3$ / # ind. 0 / 15 10/3 / 15 20/3 / 10 10 / 29			16.3918	3.33
FLELikert sample data	**FLELikert Aumann-type mean**	**FLELikert 1-norm median**	**FLELikert variance**	**FLELikert ρ_1-MDD**
$M.3$ / # ind.  / 15  / 15  / 10  / 29			11.9399	2.5
FRS sample data	**FRS Aumann-type mean**	**FRS 1-norm median**	**FRS variance**	**FRS ρ_1-MDD**
			12.2326	3.0357

To easily visualize some of the differences one can find depending on the scale, we first describe in detail how to proceed with all datasets for Question $M.3$. In this way, the available information (data), and the location and dispersion measures have been gathered in Table 2 (1st row for NELikert responses, 2nd row for FLELikert ones and third row for the FRS-based ones), the measures for fuzzy-valued data being obtained by considering the Aumann-type mean, 1-norm median, D-variance and ρ_1-MDD as formalized in the preceding section.

By looking at the results in Table 2 one can easily show that

▷ the values of the measures for the FRS responses are rather or quite different from those for the encoded Likert ones; for instance, the mean value for the NELikert responses equals 5.8935 and the one for the FLELikert is $\mathrm{Tri}(3.2859, 5.8935, 7.8261) = \mathrm{Tra}(3.2859, 5.8935, 5.8935, 7.8261)$, whereas the one for the FRS equals $\mathrm{Tra}(4.0149,\ 4.2562, 4.8982, 5.1895)$. If the fuzzy means are defuzzified through their weighted averaging based on levels (see Yager [64], or more recently Nasibov and Shikhlinskaya [41]),

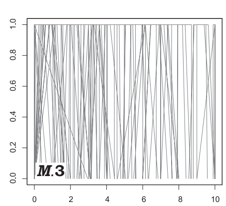
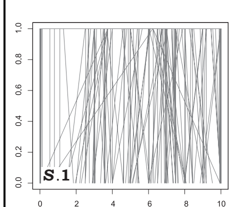
$$\mathrm{wabl}(\widetilde{U}) = \int_{[0,1]} [(\inf \widetilde{U}_\alpha + \sup \widetilde{U}_\alpha)/2]\, d\alpha,$$

which in the trapezoidal case coincides with the average of the end-points of the 0- and 1-level, we get the value 6.5377 for the FLELikert and 4.5897 for the FRS;

- ▷ the median for the NELikert and FLELikert dataset are the encoded A3, which is visually very different to that for the FRS;
- ▷ as formally proven in Sinova *et al.* [51], the 1-norm median is more robust than the Aumann-type mean since it is somewhat less influenced by the possible 'outliers' (in this case, some high values);
- ▷ regarding the variance and MDD, the FRS responses lead to a value in between the one for the NELikert and the one for the FLELikert; although it has been empirically shown (see De la Rosa de Sáa *et al.* [14]) that the variance is mostly lower for the FRS than for the other two ones, this is not true for this case. This can be possibly due to the diversity being much higher for the FRS responses and also (as we will comment later) to the fact that $M.3$ is one of the questions for which the distance between the FRS-based median and mean value is larger; close descriptive conclusions can be drawn for the MDD.

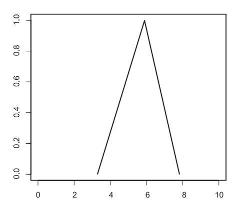
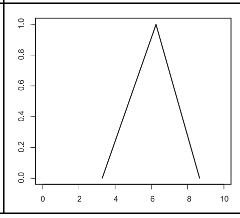
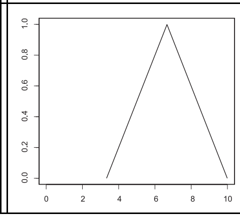
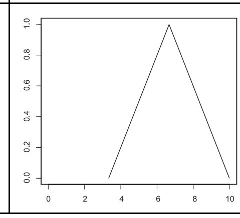
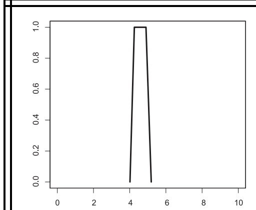
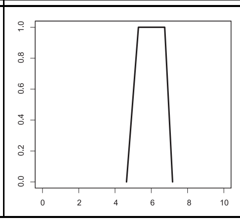
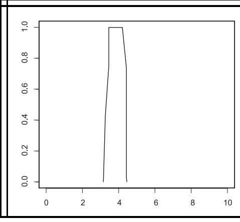
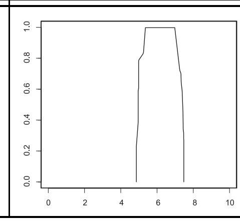
On the other hand, quite different summary measures can be obtained depending on the considered scale, as one can see by looking at the sample means and medians of different types of responses to Questions $M.3$ and $S.1$ (see Table 3, the NELikert and FLELikert sample distributions being immediate to obtain on the basis of the Likert ones).

Table 3
Sample data according to Likert and FRS for Questions $M.3$ and $S.1$

<table>
<tr><th colspan="2">$M.3$</th><th colspan="2">$S.1$</th></tr>
<tr><th>Likert sample data</th><th>FRS sample data</th><th>Likert sample data</th><th>FRS sample data</th></tr>
<tr>
<td><table><tr><th>$M.3$</th><th># ind.</th></tr><tr><td>A1</td><td>15</td></tr><tr><td>A2</td><td>15</td></tr><tr><td>A3</td><td>10</td></tr><tr><td>A4</td><td>29</td></tr></table></td>
<td>
</td>
<td><table><tr><th>$S.1$</th><th># ind.</th></tr><tr><td>A1</td><td>7</td></tr><tr><td>A2</td><td>12</td></tr><tr><td>A3</td><td>28</td></tr><tr><td>A4</td><td>19</td></tr></table></td>
<td>
</td>
</tr>
</table>

Once more, differences between the central tendency measures of these distributions are much more clearly visualized through the FRS data than through the encoded Likert ones. In this way, Table 4 clearly shows that although for NELikert and FLELikert the sample means are very close, and the sample medians coincide, this does not happen when sample FRS data are analyzed since their mean and median are quite different.

Table 4
Comparative display of the sample mean and median for NELikert, FLELikert and FRS responses to Questions $M.3$ and $S.1$

$M.3$ NELikert mean	$S.1$ NELikert mean	$M.3$ NELikert median	$S.1$ NELikert median
5.8935	6.2572	20/3	20/3
$M.3$ FLELikert Aumann-type mean	$S.1$ FLELikert Aumann-type mean	$M.3$ FLELikert 1-norm median	$S.1$ FLELikert 1-norm median
			
$M.3$ FRS Aumann-type mean	$S.1$ FRS Aumann-type mean	$M.3$ FRS 1-norm median	$S.1$ FRS 1-norm median
			

The global analysis of the sample means and medians for the nine questions can be found in Table 5.

On the other hand, the analysis of the sample variances and MDDs for the same questions can be found in Table 6.

A pairwise linear correlation analysis between the values of the sample variances for the three scales in Table 6 yields to very strong increasing linear relationships between them. More specifically,

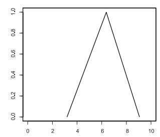
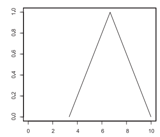
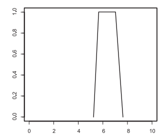
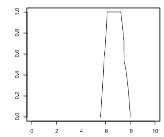
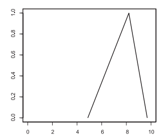
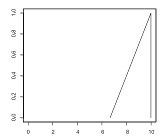
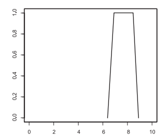
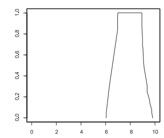
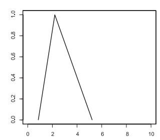
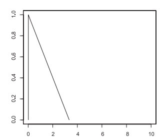
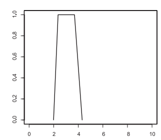
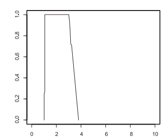
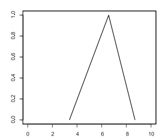
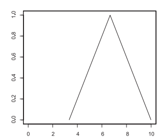
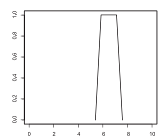
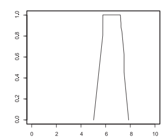
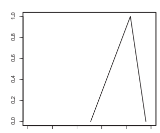
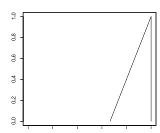
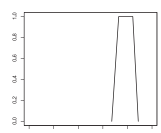
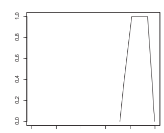
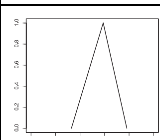
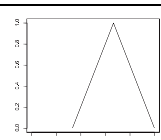
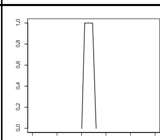
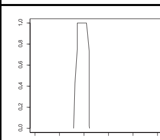
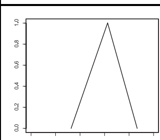
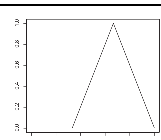
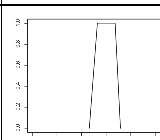
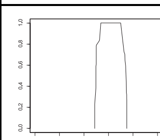
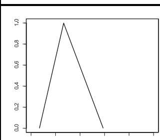
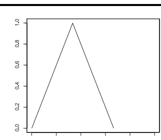
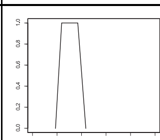
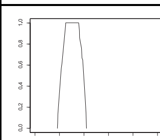
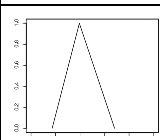
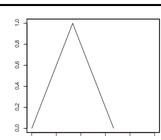
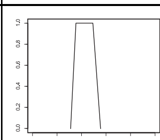
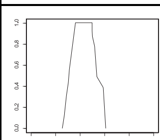
$$r_{\text{NELikert variance,FLELikert } D\text{-variance}} = .9978,$$

$$r_{\text{NELikert variance,FRS } D\text{-variance}} = .9791,$$

$$r_{\text{FLELikert } D\text{-variance,FRS } D\text{-variance}} = .9708.$$

By looking at Table 6, one can immediately conclude that the variance is lower for the FLELikert- and FRS-based responses than for the NELikert ones for the nine questions, but the comparison is not uniform between the variances for FLELikert and FRS data. Moreover, one can conclude that the MDD is lower for the FLELikert responses than for the NELikert ones for the nine questions, but the comparison is not uniform with respect to the MDD for FRS data.

Table 5
Sample means and medians for the responses to Questions $R.1$ to $S.3$

	NELikert mean	NELikert median	FLELikert Aumann mean	FLELikert 1-norm median	FRS Aumann mean	FRS 1-norm median
$R.1$	6.3738	20/3				
$R.2$	8.2099	10				
$R.3$	2.1885	0				
$M.1$	6.5672	20/3				
$M.2$	8.3341	10				
$M.3$	5.8935	20/3				
$S.1$	6.2572	20/3				
$S.2$	2.6553	10/3				
$S.3$	3.9392	10/3				

If we now compute distances between the sample Aumann-type mean and 1-norm median of the FRS responses, we can easily realize that the variance is smaller for the FRS than for the FLELikert data except when the distance between the mean and the median of the FRS responses is large (over .7, see Table 7), that is, the mean is not very representative as a central tendency summarization.

Table 6
Sample D-variances and ρ_1-MDDs for the responses to Questions $R.1$ to $S.3$

QUESTION	$R.1$	$R.2$	$R.3$	$M.1$	$M.2$	$M.3$	$S.1$	$S.2$	$S.3$
NELikert variance	6.2191	4.8236	10.2874	9.4243	6.2381	16.3918	9.9034	7.1199	12.2823
FLELikert D-variance	4.8253	3.4621	**7.3660**	7.2054	**4.4168**	**11.9399**	7.4159	5.2613	9.1435
FRS D-variance	**4.7650**	**3.1600**	8.2858	**7.0894**	5.2719	12.2326	**6.6063**	**5.1205**	**8.2269**

QUESTION	$R.1$	$R.2$	$R.3$	$M.1$	$M.2$	$M.3$	$S.1$	$S.2$	$S.3$
NELikert MDD	**0**	**0**	**0**	3.33	**0**	3.33	3.33	3.33	3.33
FLELikert ρ_1-MDD	**0**	**0**	**0**	2.5	**0**	**2.5**	2.5	2.5	2.5
FRS ρ_1-MDD	1.2934	1.3625	1.7149	**2.2140**	1.2119	3.0357	**1.8289**	**1.4314**	**2.3375**

Moreover, the MDD is generally smaller for the FRS than for the FLELikert/NELikert data when the distance between the mean and the median is rather small (see Table 7).

Table 7
Distances between the Aumann-type mean and the 1-norm median for the FRS responses to Questions $R.1$ to $S.3$

QUESTION	$R.1$	$R.2$	$R.3$	$M.1$	$M.2$	$M.3$	$S.1$	$S.2$	$S.3$
ρ_1	.3522	.3519	**.8708**	.1607	**.8216**	**.7352**	.2534	.0726	.2705
ρ_2	.3608	.4193	**.9129**	.1829	**.8527**	**.7421**	.3047	.0803	.3360
D	.3554	.3023	**.8871**	.1111	**.8328**	**.7386**	.2615	.0642	.1999

3.2.2 Diversity in responding

A first evidence supporting the use of the fuzzy rating scale is that the richness and diversity/variability/subjectivity of the available information clearly increase w.r.t. the discrete scales.

This is amply confirmed by looking at Table 3, which displays together, for instance on the left side, the responses to Question $M.3$ in accordance with the 4-point Likert scale (on the left-left), and those collected by using the fuzzy rating scale (on the left-right).

If we quantify the diversity of values by considering the Gini-Simpson index and, for each question, we only consider data from the students who have provided with both the Likert-type and the FRS responses, we obtain the results in Table 8. Notice that the frequencies of the different values coincide for Likert datasets and their NELikert and FLELikert encodings, so there is no need for an extra computation.

Table 8
Sample Gini-Simpson diversity index for Likert responses (and their usual encodings) and FRS responses to Questions $R.1$ to $S.3$

QUESTION	$R.1$	$R.2$	$R.3$	$M.1$	$M.2$	$M.3$	$S.1$	$S.2$	$S.3$
$\mathrm{Div}_{\mathrm{Gini}}(\mathrm{Likert})$	.5960	.5587	.5730	.6915	.5266	.7078	.6921	.6362	.7287
$\mathrm{Div}_{\mathrm{Gini}}(\mathrm{FRS})$	.9853	.9833	.9842	.9837	.9793	.9725	.9846	.9839	.9844

The results in Table 8 are coherent with what has been proved in De la Rosa de Sáa *et al.* [14]: under quite general conditions **the use of the fuzzy rating scale allows individual diversities to be more evident than with the use of Likert one or its numerical/fuzzy linguistic encodings**.

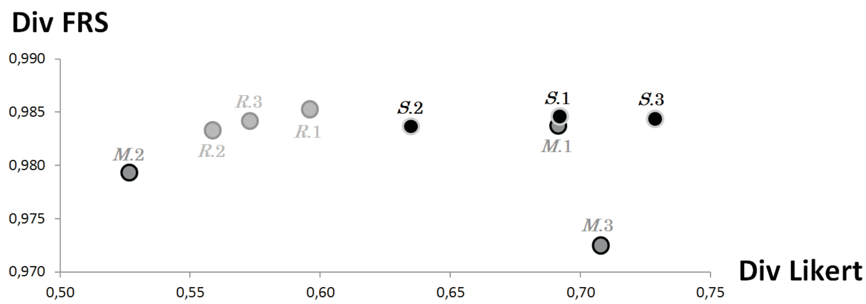


Fig. 5. Scatter diagram for Gini-Simpson diversity index for Likert (abscise) *vs* Gini-Simpson diversity index for FRS (ordinate) responses to Questions $R.1$ to $S.3$

Furthermore, one can see that, in contrast to what happens for the variance, the scatter diagram for the pairs of values of the Gini-Simpson diversity index for the Likert *vs* FRS scale (see Figure 5) does not show any specific connection between the values of the index for both scales (in fact, the linear correlation coefficient equals $r_{\mathrm{Div}_{\mathrm{Gini}}(\mathrm{Likert}),\mathrm{Div}_{\mathrm{Gini}}(\mathrm{FRS})} = -.121$, and quadratic or cubic relationships are also unappropriate). This is mainly due to the lack of a relationship between the diversities for questions related to math.

3.2.3 Analyzing the similarity between sample distributions

Finally, by means of the sample index of dissimilarity between two distributions, some pairwise differences of FRS-based responses are examined. More concretely, possible differences by sex (girls/boys), format of the questionnaire students have filled (paper-and-pencil/computerized) and type of bedroom students have at home (shared/individual) have been described in terms of the index $\varrho_{.001}$. Outputs for the computations have been collected in Table 9.

Since there is not a maximum value the index can take on, we can only state descriptive comparative conclusions instead of separate ones. The last ones would require an inferential reasoning which should be yet developed. As an example of some of the conclusions we can draw,

Table 9
Analyzing the influence of sex, format and type of bedroom on the FRS responses

POSED QUESTION	differences by sex $\varrho_{.001}$	differences by format $\varrho_{.001}$	differences by room $\varrho_{.001}$
$R.1$	.3874	1.0148	.5859
$R.2$	.2397	1.1045	1.2238
$R.3$	.6087	.7244	.4755
$M.1$	1.2692	.8622	.9392
$M.2$	.3713	1.3347	.3153
$M.3$	.6207	1.5316	.6548
$S.1$	.6784	1.5403	.2659
$S.2$	.2754	.6827	.5868
$S.3$	.4223	1.5221	.8058

▷ sex is especially influential in students liking or not mathematics (by looking at responses in more detail, boys seem to have a more positive view about the subject than girls),
▷ the bedroom type is more influential in the attitude with respect to learning from reading (those having an individual room being more positive about this activity), and
▷ rather surprisingly, the version of the questionnaire seems to be influential for the responses to almost all the questions.

3.3 Some remarks about the descriptive analysis

It should be highlighted that the choice of the reference bounded interval does not essentially affect the main conclusions in the study in this paper. More concretely,

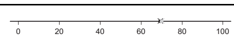
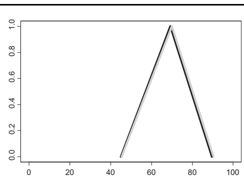
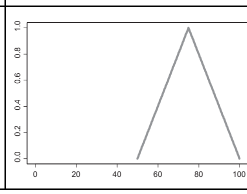
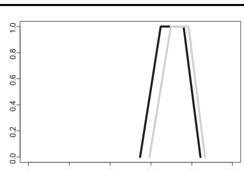
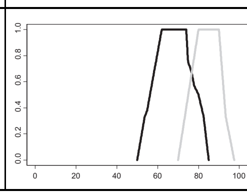
- the central tendency measures (i.e., the Aumann-type mean and the 1-norm median) are equivariant by translation and scale;
- the D-variance/ρ_1-MDD is invariant by translation and squared equivariant/equivariant by scale;
- Gini's diversity index is invariant by translation and nonnull scale, small/large dissimilarities in a reference interval correspond to small/large dissimilarities in any other one.

The comparative descriptive study in Section 3 has dealt with an even-point Likert scale, the 4 points having been predetermined by the TIMSS/PIRLS team in connection with the considered questionnaire. Consequently, respondents have had no chance to choose a neutral position within the Likert scale. Nevertheless, it should be pointed out that the fact that a neutral label is included in the questionnaire does not affect in general the main implication from this paper: statistical conclusions can differ depending on the scale.

To briefly illustrate these assertions by means of a real-life example, we can consider the datasets obtained after conducting a double-response questionnaire similar to the one in Section 3, but for which respondents can choose among five possible Likert-type responses and, simultaneously, draw a fuzzy rating scale-type one having $[0, 100]$ (see, for more details, Sinova *et al.* [50], and the web http://bellman.ciencias.uniovi.es/smire/FuzzyRatingScaleQuestionnaire-Restaurants.html). A sample of 70 people with different age, background and work type and position has been considered for responding to a restaurant customer satisfaction questionnaire. Some of the questions have consisted of making the customers rate the degree of agreement with the statements "$QF2$. The menu has a good variety of items", "$QF3$. The quality of food is excellent", "$QR2$. Employees are patient when taking my order", and "$QR7$. The service is excellent".

Table 10 shows that, as for the 4-point and reference interval [0,10], differences can be found between statistical conclusions depending on the considered scale.

Table 10
Comparative display of the sample mean-variance/sample median-MDD for NELikert, FLELikert and FRS responses to Questions $QF3$ *vs* $QR7$/$QF2$ *vs* $QR2$

$QF3$ *vs* $QR7$ NELikert mean	$QF3$ *vs* $QR7$ NELikert D-variance	$QF2$ *vs* $QR2$ NELikert median	$QF2$ *vs* $QR2$ NELikert ρ_1-MDD
	546.58 *vs* 427.54	75 (Somewhat agree)	25
$QF3$ *vs* $QR7$ FLELikert Aumann-type mean	$QF3$ *vs* $QR7$ FLELikert D-variance	$QF2$ *vs* $QR2$ FLELikert 1-norm median	$QF2$ *vs* $QR2$ FLELikert ρ_1-MDD
	462.13 *vs* 356.04		18.75
$QF3$ *vs* $QR7$ FRS Aumann-type mean	$QF3$ *vs* $QR7$ FRS D-variance	$QF2$ *vs* $QR2$ FRS 1-norm median	$QF2$ *vs* $QR2$ FRS ρ_1-MDD
	371.86 *vs* 272.93		16.79 *vs* 8.96

As a summary implication, this paper corroborates what has been stated from other statistical perspectives (see De la Rosa de Sáa *et al.* [15], for the representativeness of the mean, and Lubiano *et al.* [38], for some studies from an inferential viewpoint): conclusions based on responses from the fuzzy rating scale do not coincide in general with those based on responses from either Likert or fuzzy linguistic Likert scales. This implication has an important analogy with what happens when grouping real-valued data by intervals: both, 'identifying' fuzzy rating scale-valued responses with one of a few possible Likert labels (or their fuzzy linguistic counterpart) and 'identifying' real-valued responses with one of a few possible non-overlapping interval-valued ones, entail a loss of information so that some existing differences can be ignored, whence statistical conclusions are not usually well preserved under such an identification.

4 Concluding remarks

This paper has explained in detail an approach to descriptively analyze data obtained from the use of a fuzzy rating scale-based questionnaire. It should be remarked that there are many other studies to be developed, although they are beyond the extent and length of this paper and will also depend in practice on the real interests users can have. Among them, there are still many statistical methods to be developed for both descriptive and inferential fuzzy data analysis, and this is a clear future direction to consider.

References

[1] H. Akdag, T. Kalayci, S. Karagöz, H. Zülfikar, D. Giz, The evaluation of hospital service quality by fuzzy MCDM, Appl. Soft Comput. 23 (2014) 239–248.

[2] A. Ban, L. Coroianu, P. Grzegorzewski, Trapezoidal approximation and aggregation, Fuzzy Sets Syst. 177 (2011) 45–59.

[3] E. Benoit, Expression of uncertainty in fuzzy scales based measurements, Measurement 46 (2013) 3778–3782.

[4] E. Benoit, L. Foulloy, The role of fuzzy scales in measurement theory, Measurement 46 (2013) 2921–2926.

[5] C. Bertoluzza, N. Corral, A. Salas, On a new class of distances between fuzzy numbers, Mathware & Soft Comput. 2 (1995) 71–84.

[6] A. Blanco-Fernández, M.R. Casals, A. Colubi, N. Corral, M. García-Bárzana, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, M. Montenegro, A.B. Ramos-Guajardo, S. De la Rosa de Sáa, B. Sinova, A distance-based statistical analysis of fuzzy number-valued data, Int. J. Approx. Reason. 55 (2014) 1487–1501.

[7] A. Blanco-Fernández, M.R. Casals, A. Colubi, N. Corral, M. García-Bárzana, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, M. Montenegro, A.B. Ramos-Guajardo, S. De la Rosa de Sáa, B. Sinova, Rejoinder on "A distance-based statistical analysis of fuzzy number-valued data", Int. J. Approx. Reason. 55 (2014) 1601–1605.

[8] A. Calcagnì, L. Lombardi, Dynamic Fuzzy Rating Tracker (DYFRAT): a novel methodology for modeling real-time dynamic cognitive processes in rating scales, Appl. Soft Comput. 24 (2014) 948–961.

[9] J. Carifio, R.J. Perla, Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes, Journal of Social Sciences 3 (2007) 106–116.

[10] C. Caso, M.A. Gil, The Gini-Simpson index of diversity - Estimation in the stratified sampling, Commun. Stat. A-Theor. 17 (1988) 2981–2995.

[11] A. Colubi, J.S. Domínguez-Menchero, M. López-Díaz, D.A. Ralescu, On the formalization of fuzzy random variables, Inform. Sci. 133 (2001) 3–6.

[12] N. Corral-Blanco, E. Zurbano-Fernández, A. Blanco-Fernández, I. García-Honrado, A.B. Ramos-Guajardo, Structure of the family educational environment: its influence on performance and differential performance,

in: PIRLS - TIMSS 2011 International Study on Progress in Reading Comprehension, Mathematics and Sciences IEA, Vol. II, Spanish Report, Secondary Analysis, Ministerio de Educación, Cultura y Deporte, Instituto Nacional de Evaluación Educativa, 2013, pp. 9–31.

[13] S. De la Rosa de Sáa, P. Filzmoser, M.A. Gil, M.A. Lubiano, On the robustness of absolute deviations with fuzzy data, in: P. Grzegorzewski, M. Gagolewski, O. Hryniewicz, M.A. Gil (Eds.), Strengthening Links Between Data Analysis and Soft Computing, Springer, Heidelberg, 2015, pp. 133–141.

[14] S. De la Rosa de Sáa, M.A. Gil, G. González-Rodríguez, M.T. López, M.A. Lubiano, Fuzzy rating scale-based questionnaires and their statistical analysis, IEEE T. Fuzzy Syst. 23 (2015) 111–126.

[15] S. De la Rosa de Sáa, M.A. Gil, M.T. López, M.A. Lubiano, Fuzzy rating *vs.* fuzzy conversion scales: an empirical comparison through the MSE, in: R. Kruse, M.R. Berthold, C. Moewes, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.), Synergies of Soft Computing and Statistics for Intelligent Data Analysis, Springer, Heidelberg, 2013, pp. 135–143.

[16] S. De la Rosa de Sáa, M.A. Lubiano, B. Sinova, P. Filzmoser, Robust scale estimators for fuzzy data, Adv. Data Anal. Classif. (2016) (in press, doi:10.1007/s11634-015-0210-1).

[17] P. Diamond, P. Kloeden, Metric spaces of fuzzy sets, Fuzzy Sets Syst. 100 (1999) 63–71.

[18] F.J. Estrella, M. Espinilla, F., Herrera, L. Martínez, FLINTSTONES: A fuzzy linguistic decision tools enhancement suite based on the 2-tuple linguistic model and extensions, Inform. Sci., 280 (2014) 152–170.

[19] S. García-Galán, R.P. Prado, J.E. Muñoz Expósito, Rules discovery in fuzzy classifier systems with PSO for scheduling in grid computational infrastructures, Appl. Soft Comput., 29 (2015) 424–435.

[20] J. Ghneim (Ed.) (2013). Fuzzy concept, in: Encyclopedia of Psychometrics, Wikipedia, 2014, pp. 61–68.

[21] M.A. Gil, P. Gil, On some information measures of degree $\beta = 2$: estimation in simple-stage cluster sampling, Statist. Probab. Lett. 8 (1989) 157–162.

[22] M.A. Gil, M.A. Lubiano, S. De la Rosa de Sáa and B. Sinova, Analyzing data from a fuzzy rating scale-based questionnaire. A case study, Psicothema, 27 (2015) 182–191.

[23] C.W. Gini, Variabilità e mutabilità, Studi Economico-Giuridici della R. Università di Cagliari, 3 (1912) 3–159.

[24] P. Grzegorzewski, Trapezoidal approximations of fuzzy numbers preserving the expected interval - algorithms and properties, Fuzzy Sets Syst. 159 (2008) 1354–1364.

[25] P. Grzegorzewski, K. Pasternak-Winiarska, Trapezoidal approximations of fuzzy numbers with restrictions on the support and core, in: Proceedings 7th Conference EUSFLAT-2011 and LFA-2011, Atlantis Press, Paris, 2011, pp. 749–756.

[26] F. Herrera, E. Herrera-Viedma, L. Martínez, A fuzzy linguistic methodology to deal with unbalanced linguistic term sets, IEEE T. Fuzzy Syst. 16 (2008) 354–370.

[27] F. Herrera, E. Herrera-Viedma, J.L. Verdegay, Choice processes for non-homogeneous group decision making in linguistic setting, Fuzzy Sets Syst. 94 (1998) 287–308.

[28] B. Hesketh, B. Griffin, V. Loh, A future-oriented retirement transition adjustment framework, J. Vocat. Behav. 79 (2011) 303–314.

[29] T. Hesketh, B. Hesketh, Computerized fuzzy ratings: the concept of a fuzzy class, Behav. Res. Meth. Ins. C. 26 (1994) 272–277.

[30] T. Hesketh, R. Pryor, B. Hesketh, An application of a computerized fuzzy graphic rating scale to the psychological measurement of individual differences, Int. J. Man-Mach. Stud. 29 (1988) 21–35.

[31] S. Jamieson, Likert scales: how to (ab)use them, Med. Educ. 38 (2004) 1217–1218.

[32] M. Lalla, G. Facchinetti, G. Mastroleo, Vagueness evaluation of the crisp output in a fuzzy inference system, Fuzzy Sets Syst. 159 (2008) 3297–3312.

[33] Q. Li, A novel Likert scale based on fuzzy sets theory, Expert Syst. Appl. 40 (2013) 1609–1618.

[34] R. Liu, L. Cui, G. Zeng, H. Wu, C. Wang, S. Yan, B. Yan, Applying the fuzzy SERVQUAL method to measure the service quality in certification & inspection industry, Appl. Soft Comput. 26 (2015) 508–512.

[35] L.M. Lozano, E. García-Cueto, J. Muñiz, Effect of the number of response categories on the reliability and validity of rating scales, Methodology 4 (2008) 73–79.

[36] M.A. Lubiano, S. De la Rosa de Sáa, B. Sinova, M.A. Gil, Empirical sensitivity analysis on the influence of the shape of fuzzy data on the estimation of some statistical measures, in: P. Grzegorzewski, M. Gagolewski, O. Hryniewicz, M.A. Gil (Eds.), Strengthening Links Between Data Analysis and Soft Computing, Springer, Heidelberg, 2015, pp. 123–131.

[37] M.A. Lubiano, M.A. Gil, M. López-Díaz, M.T. López, The $\overrightarrow{\lambda}$-mean squared dispersion associated with a fuzzy random variable, Fuzzy Sets Syst. 111 (2000) 307–317.

[38] M.A., Lubiano, M. Montenegro, B. Sinova, S. De la Rosa de Sáa, M.A. Gil, Hypothesis testing for means in connection with fuzzy rating scale-based data: algorithms and applications, Eur. J. Oper. Res. (2016) (in press, doi:10.1016/j.ejor.2015.11.016).

[39] S. Massanet, J.V. Riera, J. Torrens, E. Herrera-Viedma, A new linguistic computational model based on discrete fuzzy numbers for computing with words, Inform. Sci. 258 (2014) 277–290.

[40] T. Matsui, M. Takeya, Structural analysis method for fuzzy rating scale data using fuzzy integration, in: Proceedings IEEE International Conference on Systems, Man and Cybernetics, 1, 1994, pp. 493–500.

[41] E.N. Nasibov, R.Y. Shikhlinskaya, Adjustment of the parameters of WABL-aggregation for locating the center of gravity of a polynomial type fuzzy number, Autom. Control Comp. S. 37 (2003) 34–42.

[42] W. Pedrycz, A fuzzy cognitive structure for pattern recognition, Pattern Recogn. Lett. 9 (1989) 305–313.

[43] W. Pedrycz, Why triangular membership functions?, Fuzzy Sets Syst. 64 (1994) 21–30.

[44] R. Pérez, M.A. Gil, P. Gil, Estimating the uncertainty associated with a variable in a finite population, Kybernetes 15 (1986) 251–256.

[45] M.L. Puri, D.A. Ralescu, Fuzzy random variables, J. Math. Anal. Appl. 114 (1986) 409–422.

[46] U.D. Reips, F. Funke, Interval-level measurement with visual analogue scales in Internet-based research: VAS Generator, Behav. Res. Meth. 40 (2008) 699–704.

[47] E.H. Simpson, Measurement of diversity, Nature 163 (1949) 688.

[48] B. Sinova, M.R. Casals, M.A. Gil, M.A. Lubiano, The fuzzy characterizing function of the distribution of a random fuzzy number, Appl. Math. Model. 39 (2015) 4044-4056.

[49] B. Sinova, S. De la Rosa de Sáa, M.A. Gil, A generalized L^1-type metric between fuzzy numbers for an approach to central tendency of fuzzy data, Inform. Sci. 242 (2013) 22–34.
[50] B. Sinova, S. De la Rosa de Sáa, M.A. Gil, M.A. Lubiano, An overview on the statistical central tendency for fuzzy data sets, Int. J. Uncertain. Fuzz. 23 Suppl. 1 (2015) 105–132.
[51] B. Sinova, M.A. Gil, A. Colubi, S. Van Aelst, The median of a random fuzzy number. The 1-norm distance approach, Fuzzy Sets Syst. 200 (2012) 99-115.
[52] B. Sinova, M.A. Gil, S. Van Aelst, M-estimates of location for the robust central tendency of fuzzy data, IEEE T. Fuzzy Syst. (2016) (in press, doi:10.1109/TFUZZ.2015.2489245).
[53] K. Takemura, A fuzzy linear regression analysis for fuzzy input-output data using the least squares method under linear constraints and its application to fuzzy rating data, J. Advanced Comput. Intel. and Intel. Inform. 3, (1999) 36–41.
[54] K. Takemura, Ambiguous comparative judgment: fuzzy set model and data analysis, Jpn. Psychol. Res. 49 (2007) 148–156.
[55] K. Takemura, Ambiguity and social judgment: fuzzy set model and data analysis, in: E.P. Dadios (Ed.), Fuzzy Logic - Algorithms, Techniques and Implementations, InTech-Open Access Publisher, 2012, pp. 3–24.
[56] M. Tavana, D. Di Caprio, F.J. Santos-Arteaga, A bilateral exchange model: The paradox of quantifying the linguistic values of qualitative characteristics, Inform. Sci. 296 (2015) 201–218.
[57] A. Tejeda-Lorente, J. Bernabé-Moreno, C. Porcel, P. Galindo-Moreno, E. Herrera-Viedma, A dynamic recommender system as reinforcement for personalized education by a fuzzly linguistic web system, Procedia Comput. Sci. 55 (2015) 1143–1150.
[58] A. Tejeda-Lorente, C. Porcel, E. Peis, R. Sanz, E. Herrera-Viedma, A quality based recommender system to disseminate information in a university digital library, Inform. Sci. 261 (2014) 52–69.
[59] R.M. Tong, P.P. Bonissone, A linguistic approach to decision making with fuzzy sets, IEEE T. Syst. Man Cyb. 10 (1980) 716–723.
[60] H. Treiblmaier, P. Filzmoser, Benefits from using continuous rating scales in online survey research, Proc. ICIS'2011, Shanghai (2011) Paper 1.
[61] W. Trutschnig, G. González-Rodríguez, A. Colubi, M.A. Gil, A new family of metrics for compact, convex (fuzzy) sets based on a generalized concept of mid and spread, Inform. Sci. 179 (2009) 3964–3972.
[62] P.J. Villacorta, A.D. Masegosa, D. Castellanos, M.T. Lamata, A new fuzzy linguistic approach to qualitative Cross Impact Analysis, Appl. Soft Comput. 24 (2014) 19–30.
[63] X. Wang, X. Liu, L. Zhang, A rapid fuzzy rule clustering method based on granular computing, Appl. Soft Comput. 24 (2014) 534–542.
[64] R.R. Yager, A procedure for ordering fuzzy subsets of the unit interval, Inform. Sci. 24 (1981) 143–161.
[65] T. Yamashita, Fuzzy ratings and crisp feedback in fuzzy AHP for supporting human decision making, J. Advanced Comput. Intel. and Intel. Inform. 10 (2006) 219–224.
[66] K.K.F. Yuen, Compound linguistic scale, Appl. Soft Comput. 21 (2014) 38–56.
[67] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning, I, Inform. Sci. 8 (1975) 199–249; II, Inform. Sci. 8 (1975) 301–353; III, Inform. Sci. 9 (1975) 43–80.
[68] L.A. Zadeh, Is there a need for fuzzy logic?, Inform. Sci. 178 (2008) 2751–2779.