Prioritization of QFD Customer Requirements Based on the Law of Comparative Judgments

Prioritization of QFD Customer Requirements based on the Law of Comparative Judgments

Fiorenzo Franceschini[1] and Domenico Maisano

[1] *fiorenzo.franceschini@polito.it*
Politecnico di Torino, DIGEP (Department of Management and Production Engineering),
Corso Duca degli Abruzzi 24, 10129, Torino (Italy)

Abstract

Quality Function Deployment (QFD) is a useful tool to improve the design/development process of products and services. The initial phases of the QFD process – i.e., those concerning the collection and analysis of the so-called Voice of the Customer – are probably the most critical, because any distortion can propagate to the whole process results, making it ineffective or even misleading. The focus of this paper is on the phase of prioritization of customer requirements (CRs). There are numerous techniques for this task; however (i) the simplest often introduce questionable or unrealistic assumptions, while (ii) the most sophisticated often require too much elaborate and repetitious information from customers, which may lead to inconsistencies.

This paper introduces a new prioritization technique based on the Thurstone's Law of Comparative Judgment. This technique makes it possible to aggregate the evaluations by multiple respondents and transform them into an interval scale, which depicts the relative importance of CRs. The greatest strength of this technique is combining a refined theoretical model with a simple and user-friendly data collection process. The description is supported by a realistic application example concerning the prioritization of QFD's CRs in the design of an aircraft seat.

Keywords: QFD, Customer requirements, Prioritization, Relative importance ratings, Law of Comparative Judgment, Thurstone scaling, Interval scale.

1. Introduction and literature review

Quality Function Deployment (QFD) is a powerful technique to increase the customer satisfaction of product and services. The implementation of QFD may generate significant improvements in the design/development process, such as fewer and earlier design changes, improved cross-functional communications, improved product/service quality, and reduced development time and cost (Hauser and Clausing, 1988; Griffin and Hauser, 1993; Tran and Sherif, 1995; Franceschini, 2002). These improvements are critical success factors to companies in a global marketplace characterized by intense international competition.

The great diffusion of QFD is demonstrated by the literally thousands of scientific publications illustrating a variety of industrial applications, methodological improvements, new variants, and possible integration with other tools.

Typically, QFD utilizes four sets of matrices – the so called Houses of Quality (HoQs). The four HoQs respectively translate (i) customer requirements (CRs) into engineering characteristics and, in turn, into (ii) parts characteristics, (iii) process plans, and (iv) production requirements (Franceschini, 2002). For detailed information, we refer the reader to the vast literature and extensive reviews, e.g., (Chan and Wua, 2002; Sharma et al., 2008).

The customer input, also defined as Voice of the Customer (VoC), is the key starting point for QFD process; if it does not accurately reflect what the customer expects from the product/service of interest, the process may lead to incorrect conclusions (Sireli et al, 2007). Therefore, the first HoQ, also defined as Product Planning HoQ, is of fundamental and strategic importance (Gonzalez et al. 2003). The Product Planning HoQ construction process can be summarized into ten phases, as shown in Fig. 1.

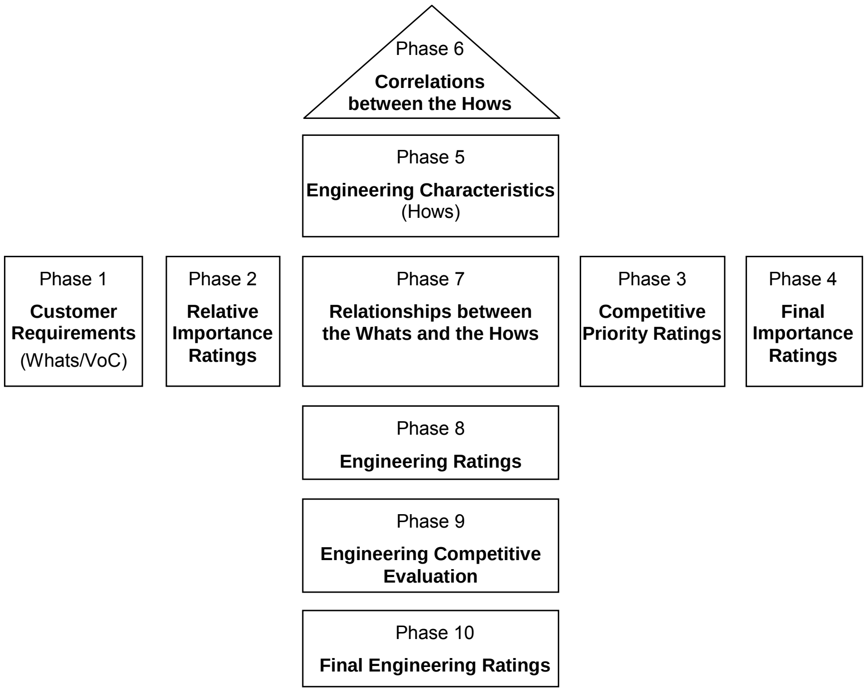

Fig. 1. Main phases of the Product Planning HoQ construction process.

Among these phases – described in detail in the literature, e.g., see (Chan and Wua, 2002; Franceschini, 2002; Franceschini and Rossetto, 2002) – particularly significant are those related to the VoC collection and analysis. The initial phase (i.e., Phase 1, "Customer Requirements", in the scheme in Fig. 1) concerns the VoC collection –through interviews and questionnaires – and analysis, in order to determine an exhaustive list of CRs. For this task, it is necessary to select a representative sample of (potential) customers, with reasonable knowledge of the product/service to be designed. It was found empirically that samples consisting of 20 to 30 respondents are sufficient to cover most CRs; also, for data collected to be reasonable and applicable, respondents have to gain a full understanding of the task required (Urban and Hauser, 1993).

In the second part of this phase, a cross-functional team of experts – composed of members from marketing, design, quality, finance and production – have to review, reorganize and insert the CRs into the Product Planning HoQ.

The next stage, which is the focus of this paper, is that of the prioritization of CRs (i.e., Phase 2, "Relative Importance Ratings", in the scheme in Fig. 1), presuming that the main CRs related to the product/service to be designed have already been identified in Phase 1. The expression "Relative Importance Ratings" indicates that this prioritization is aimed at discriminating a CR based on its importance over the others. On the other hand, Phase 4, "Final Importance Ratings" (in Fig. 1), denotes a prioritization that also takes into account the comparison of quality performance of the products/services of the company and those of its competitors.

In Phase 2, a sample of customers – generally the same involved in Phase 1 – have to prioritize the QFD's CRs using several possible approaches. Some of them are point direct scoring method (Hauser and Clausing, 1988; Griffin and Hauser, 1993), analytic hierarchy process (AHP) (Li et al., 2009; Chuang, P.T., 2001), analytic network process (ANP) (Karsak et al., 2002; Lee et al., 2008), outranking methods (Franceschini and Rossetto, 1995; Figueira et al., 2005), fuzzy variants (Chan et al., 1999; Kwong and Bai, 2002; Buyukozkan et al., 2004), and techniques derived from the Kano model (Matzler and Hinterhuber, 1998; Sireli et al., 2007; Chaudha et al., 2011). Without going into these techniques in detail, we remark that they may use different kinds of response data and elaborations from respondents. Even though all these techniques are supposed to reflect the VoC, sometimes they may lead to misleading results, especially when the data collection approach is too complex and elaborate. Here are some examples:

- Techniques based on the AHP and ANP method require CR judgments in the form of paired comparison data, defined on a *ratio* scale; e.g. "CR_1 is twice as important as CR_2" (Chuang, 2001; Franceschini, 2002; Kwong and Bai, 2002; Lee et al., 2008 , Li te al., 2009). These evaluations are inevitably arbitrary and subjective as respondents may find it difficult to express their judgments on this scale. Techniques that integrate the Kano model in the QFD environment require relatively complex questionnaires (Nahm et al., 2013) and the definition of arbitrary weights for the (qualitative) Kano categories (i.e., *basic* or *must-be* (B), *one dimensional* (O), *attractive* (A), *indifferent* (I), *reverse* (R) and *questionable* (Q) (Tan and Shen, 2000)).
- Other sophisticated techniques for the CR prioritization, such as that proposed by Nahm et al. (2013), model the uncertainty in customer requirements, taking into account the uncertainty of customer's judgment. Unfortunately, they generally include complex and structured questionnaires and, sometimes, introduce questionable assumptions in the response data processing.
- In the classical questionnaires for prioritizing QFD's CRs, respondent judgements are defined on a 5-level rating response scale (1=Not at all important, 2=Low importance, 3=Medium

importance, 4=High importance and 5=Very high importance). This response scale has two inherent limitations:

1. Since it is an ordinal scale, it only allows comparisons like “CR_1 is more important than CR_2”. Unfortunately, a typical abuse is “promoting” this scale to an interval or even ratio scale, so as to make incorrect comparisons like “the distance, in terms of importance, between CR_1 and CR_2 is greater than that between CR_3 and CR_4” or “CR_1 is three times more important than CR_2” (Stevens, 1946; Berko, Kloeber, Deckro, 2002; Franceschini, 2007).
2. These scales are used subjectively, as there is no absolute reference shared by all respondents. In general, “indulgent” respondents will tend to assign higher levels of importance, while “severe” respondents will tend to assign lower ones. For example, let us consider the ratings about three CRs (i.e., CR_1, CR_2 and CR_3) by two fictitious respondents (A and B). These ratings on a 5-level scale are respectively A: 3, 2, 1, and B: 5, 4, 2. Despite the relative rankings are identical (i.e., $CR_1 > CR_2 > CR_3$), judgments by A (severe respondent) are concentrated in the lowest levels of the scale, while those of B (indulgent respondent) in the highest. For this reason, it is questionable to aggregate judgments by different respondents through indicators of central tendency, such as the median or the mean value.

The objective of this paper is to introduce a simple technique for the CR prioritization, based on the so-called Thurstone’s “Law of Comparative Judgement”, for aggregating the judgments by multiple respondents and transforming them into a numerical interval scale (Thurstone, 1927).
An important benefit of this technique is combining a simple and user-friendly data collection process – based on the definition of respondent judgements on a 5-level ordinal scale – with a refined theoretical model.
The remainder of this paper is structured into three sections. Sect. 2 provides some background information, which is helpful to grasp the logic of the novel prioritization technique: (i) basic concepts concerning the Thurstone model and (ii) description of a process for deriving response data suitable to this model, keeping data collection as simple and user-friendly as possible. Sect. 3 shows a realistic application example concerning the prioritization of the QFD’s CRs in the design of an aircraft seat for passengers. The concluding section summarizes the original contributions of the paper, focusing on the benefits and limitations of the proposed technique, and possible future research.

2. Background Information

2.1 Basics of Thurstone’s Law of Comparative Judgement

In 1927, Thurstone presented his *Law of Comparative Judgement* (LCJ), i.e., a mathematical model to estimate scale values based on binary choices between *stimuli* (Thurstone, 1927). The explanation of this model will refer to the problem of the relative importance prioritization of QFD's CRs, on the basis of the VoC.

Thurstone postulated that each *stimulus* (CR in this case) will possess some *attribute* (importance level in this case) in varying but unknown degrees. For each of the CRs and among all subjects, it is assumed that a preference will exist. These two conditions imply the assumption of unidimensionality of the scale representing the importance of CRs (McIver and Carmines, 1981). It is also assumed that, for each *i*-th CR, the preference will be distributed normally, i.e., $CR_i \sim N(\mu_i, \sigma_i^2)$, being μ_i and σ_i^2 the unknown mean value and variance of that CR. A person's preference for each CR versus every other CR is thereby obtained. The more persons who select one CR of a pair over the other CR, the greater the importance for that CR, and thus the greater its scale weight (Edwards, 1957).

Thurstone's LCJ is an indirect form of measurement based on a transformation of individual preferences (input data) into scale values on a *psychological continuum*. Such indirect approaches are referred to as *scaling* processes. There are many scaling models; the most well known are the Rasch model (Rasch, 1966; Jansen, 1984) and conjoint analysis (Luce and Tukey, 1964). In addition, the LCJ model is based on deriving group scale values from dispersed individual choice data. Therefore, it can be also considered as a statistic choice model.

In Thurstone's terminology, choices are mediated by a *discriminal process*. He defines this as the process by which an individual identifies, distinguishes, or reacts to *stimuli*. Let us consider the theoretical distributions of the discriminal process for any two CRs, CR_i and CR_j (see Fig. 2(a)). In the LCJ model, the distribution associated with a given CR is characterized by a dispersion (or variance) of that CR, which reflects the subject-to-subject variability. Dispersions may be different for different CRs. Let μ_i and μ_j correspond to the (unknown) scale values of the two CRs and σ_i^2 and σ_j^2 the (unknown) variances.

The difference ($CR_{ij} = CR_i - CR_j$) will follow a normal distribution with parameters:

$$\mu_{ij} = \mu_i - \mu_j \text{ and } \sigma_{ij} = \sqrt{\sigma_i^2 + \sigma_j^2 - 2 \cdot \rho_{ij} \cdot \sigma_i \cdot \sigma_j} \ . \quad (1)$$

where:

μ_i and μ_j denote the (unknown) mean values of CR_i and CR_j;

σ_i^2 and σ_j^2 denote the (unknown) variances of CR_i and CR_j;

ρ_{ij} denotes the (unknown) correlation between the pairs of discriminal processes CR_i and CR_j.

Considering the area subtended by the distribution of CR_{ij}, let us draw a vertical line passing through the point with $CR_{ij} = CR_i - CR_j = 0$ (see Fig. 2(b)). The area to the right of the line depicts the observed proportion of times (p_{ij}) that $CR_{ij} \geq 0$. Of course, the area to the left depicts the complementary proportion $(1 - p_{ij})$.

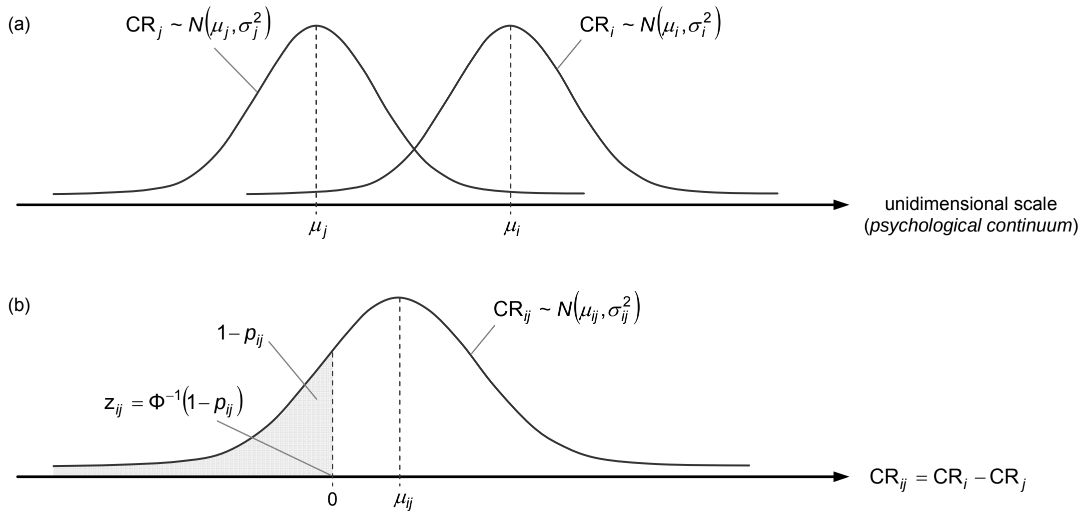

Fig. 2. (a) Theoretical distributions of the discriminal process for two CRs (i.e., CR_i and CR_j). (b) Link between $CR_{ij} = CR_i - CR_j$, and z_{ij} , i.e., the unit normal deviate corresponding to the probability $1 - p_{ij}$, being $p_{ij} = \Pr(CR_{ij} \geq 0)$.

In the standard method of Thurstone scaling, the paired comparison approach is used to collect response data. Under the protocol, respondents are forced to express a preference for one CR over another (i.e., by asking them to rank order CRs two at a time rather than all at once). All possible $\binom{n}{2} = \frac{n \cdot (n-2)}{2}$ pairs are assessed, n being the number of CRs of interest.

Paired comparison data of each respondent are reported into a "binary" matrix (B). For the purpose of example, Fig. 3(a) shows three matrices (B_1, B_2 and B_3) related to three fictitious respondents. The element of the single respondent's matrix is 1 when the CR in the i-th row is preferred to that in the j-th column. If two CRs have identical level of importance (e.g., CR_3 and CR_4 in matrix B_1), their mutual paired comparisons are conventionally 0.5.

After the total pairs of CRs have been determined for a large number of respondents (N), respondents' matrices can be summed into a single frequency matrix (F), whose general element f_{ij} represents the number of times that CR_i was preferred to CR_j. Fig. 3(b) reports the matrix F aggregating the judgment matrices B_1, B_2 and B_3. The general element f_{ij}, which appears in the i-th row and j-th column, denotes the observed number of times that CR_i was judged better or worse than CR_j.

Matrix P (Fig. 3(c)) is constructed from matrix F ($p_{ij} = \frac{f_{ij}}{N}$). The element p_{ij} is the observed proportion of times that CR_i was chosen over CR_j. Symmetric cells now sum to unity.

Interpreting p_{ij} in probabilistic terms, it can be stated that $p_{ij} = \Pr(CR_{ij} \geq 0)$. Since CR_{ij} follows a normal distribution, a standardized variable can be defined:

$$z_{ij} = \frac{CR_{ij} - \mu_{ij}}{\sigma_{ij}} \quad , \tag{2}$$

where the element z_{ij} is the unit normal deviate. For $CR_{ij} = 0$ the unit normal deviate is determined by the theoretical proportion $(1 - p_{ij})$, i.e., $z_{ij} = \Phi^{-1}(1 - p_{ij})$, Φ being the cumulative distribution function of the standard normal distribution (see Fig. 2(b)). The element z_{ij} will be positive for all values of $(1 - p_{ij})$ over 0.50 and negative for all values of $(1 - p_{ij})$ under 0.50.

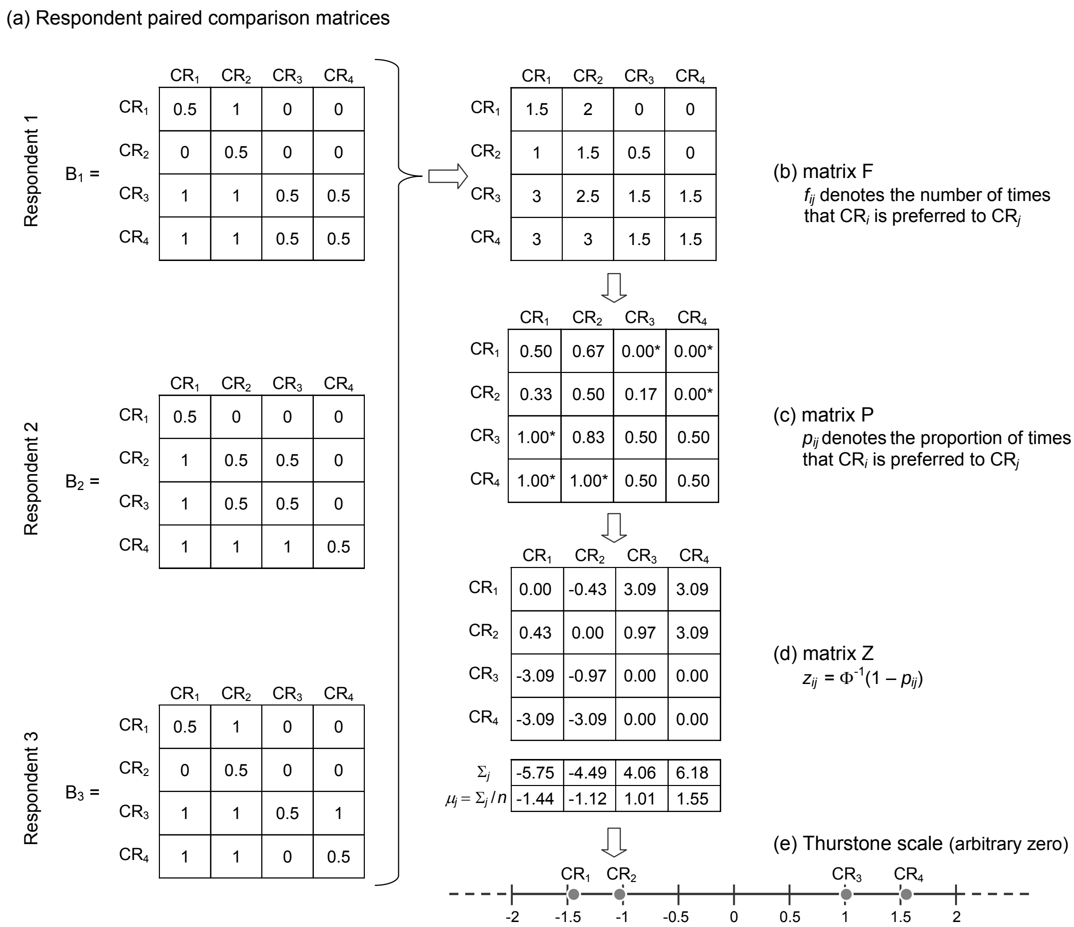

Fig. 3. Main steps of Thurstone scaling.

In detail, for $CR_{ij} = 0$, Eq. 2 becomes:

$$z_{ij} = \frac{0 - \mu_{ij}}{\sigma_{ij}} = -\frac{\mu_{ij}}{\sigma_{ij}} \quad \rightarrow \quad \mu_{ij} = -z_{ij} \cdot \sigma_{ij}, \text{ with } z_{ij} = \Phi^{-1}(1 - p_{ij}). \tag{3}$$

Combining the second formula in Eq. 3 with the expression of σ_{ij} in Eq.1, we obtain:

$$\mu_{ij} = \mu_i - \mu_j = -z_{ij} \cdot \sqrt{\sigma_i^2 + \sigma_j^2 - 2 \cdot \rho_{ij} \cdot \sigma_i \cdot \sigma_j} \ , \tag{4}$$

Matrix P is used to construct matrix Z (see Fig. 3(d)), the basic transformation matrix. Zeros are entered in the diagonal cells in matrix Z because we can ordinarily assume that here $\mu_i - \mu_i = 0$.

Apart from the aforementioned assumptions, the Thurstone model considered here is based on the following further hypotheses:

- CRs are judged differently by subjects; if all subjects would express the same preference for each outcome, the model would not be viable (proportions of 1.00 and 0.00 in the matrix P cannot be used because the *z* values corresponding to these proportions are $\pm\infty$). This is the case for the pair-wise comparisons CR_1 and CR_3, CR_1 and CR_4, and CR_2 and CR_4 in the matrix P in Fig. 3(c): in every comparison the second CR is unanimously preferred to the first. A simplified approach for tackling this problem is associating values of $p_{ij} \leq 0.001$ with $z_{ij} = \Phi^{-1}(1 - 0.001) = 3.09$ and values of $p_{ij} \geq 0.999$ with $z_{ij} = \Phi^{-1}(1 - 0.999) = -3.09$ (see the items marked with "*" in the matrix P in Fig. 3(c)). More sophisticated solutions to deal with this issue have been proposed (Edwards, 1957; Krus and Kennedy, 1977).
- As a further practical assumption, it is assumed that the CRs standard deviations are all equal ($\sigma_i = \sigma_j = \ldots = \sigma$). Therefore Eq. 4 turns into:

$$\mu_{ij} = -z_{ij} \cdot \sqrt{2 \cdot \sigma^2 \cdot (1 - \rho_{ij})}, \tag{5}$$

- It is further assumed that the intercorrelations are all equal to one another ($\rho_{ij} = \rho, \forall i, j$), so that Eq. 5 turns into:

$$\mu_{ij} = -z_{ij} \cdot \sqrt{2 \cdot \sigma^2 \cdot (1 - \rho)}, \tag{6}$$

  More precisely, Thurstone (1927) states that in a paired judgment in which the evaluation of one of the *stimuli* has no influence on the evaluation of the other *stimulus*, the correlation ρ_{ij} is likely to be very low and possibly even zero. Also, the assumption that the intercorrelations are all equal to zero is relatively safe when (i) the set of *stimuli* is rather variegated[1] and (ii) the group of respondents is not too small. Since, in the case of the QFD's CR prioritization, both these conditions are generally satisfied, it does not seem unreasonable to assume that $\rho_{ij} = \rho = 0, \forall i, j$.

Then, under the assumptions we have made, $\sqrt{2 \cdot \sigma^2 \cdot (1 - \rho)}$ (or $\sqrt{2 \cdot \sigma^2}$ in the case ρ is assumed to be zero) will be a constant and is the common scale factor of the various arithmetic mean pairs of CRs. Without any loss of generality, this common scale factor is set to 1, so that:

$$\mu_{ij} = \mu_i - \mu_j = -z_{ij}. \tag{7}$$

Eq. 7, with the assumptions involved in its derivation, is commonly referred to as Case V of the LCJ (Thurstone, 1927).

Now we can show that Thurstone scale values for each CR can be obtained from the elements of the matrix Z. Actually, if we sum the entries in the *j*-th column of the matrix Z, we obtain:

[1] The adjective "variegated" indicates that the *stimuli* of interest represent different basic concepts, not the same one, just stated in different ways.

$$\sum_{i=1}^{n} z_{ij} = \sum_{i=1}^{n} \left(\mu_j - \mu_i\right) = n \cdot \mu_j - \sum_{i=1}^{n} \mu_i \,, \qquad (8)$$

where $\sum_{i=1}^{n} z_{ij}$ means that the j-th column is held constant and the summation is over the n rows of the table. The first term on the right is the sum of the scale value of the j-th CR and the second term is the sum of the scale values of all n CRs on the psychological continuum. Dividing both sides of Eq. 8 by n, we have:

$$\bar{z}_{\cdot j} = \frac{\sum_{i=1}^{n} z_{ij}}{n} = \mu_j - \frac{\sum_{i=1}^{n} \mu_i}{n} = \mu_j - \mu \,, \qquad (9)$$

being:

$\bar{z}_{\cdot j}$ the arithmetic mean of the entries in the j-th column of the matrix Z;

μ the arithmetic mean of the (n) μ_i values;

μ_j the mean value of the j-th CR.

Thus we see that the mean of the z values in the j-th column of the matrix Z expresses the mean value of j-th CR in terms of its deviation from the mean of all the μ_i values (i.e., μ). This procedure can be applied to every column of matrix Z, in order to obtain the scale values of every CR. These values are shown in the second row at the bottom of matrix Z (Fig. 3(d)) and graphically represented in Fig. 3(e).

As a check upon calculations, we observe that the sum of the scale values in deviation form is equal to zero ($\sum_{j=1}^{n} \bar{z}_{\cdot j} = \sum_{j=1}^{n} \mu_j - \sum_{j=1}^{n} \mu = n\mu - n\mu = 0$). CRs with negative scale values are thus judged to be less favourable than the average of the scale values of all CRs and those with positive scale values are judged to be more favourable than the average. Since the scale origin – taken as the mean of the scale values of the CRs on the psychological continuum – is arbitrary, we can apply a permissible scale transformation (i.e., monotonically increasing linear function (Stevens, 1946)), so as to obtain numerical values easier to handle. This will not change the relative position of the scale values on the psychological continuum.

2.2 Practical response mode

As shown in Sect. 2.1, in the standard method of Thurstone scaling, the paired comparison approach is used to collect response data. A drawback of this approach is that it can be tedious and complex to manage for n greater than 4 or 5, since it requires so much repetitious information from respondents. An alternative response mode, which also yields data suitable for Thurstone scaling, is based on two steps:

1. Turning each respondent's judgments, typically expressed on a 5-level rating scale (see Fig. 4(a)), into rank order data (see Fig. 4(b)).
2. For each respondent, rank order data can be transformed into paired comparison data and reported into a matrix (see Fig. 4(c)). The element of the single respondent's matrix is 1 when the CR in the *i*-th row is preferred to that in the *j*-th column. If two CRs have identical level of importance, their mutual paired comparisons are conventionally 0.5.

The response mode based on a 5-level rating scale – as well as being less tedious and time consuming than the paired comparison approach – forces the respondent to be *transitive* (e.g., if $CR_1 > CR_2$ and $CR_2 > CR_3$, then $CR_1 > CR_3$). Also, it is generally familiar to respondents and therefore less subject to misinterpretation.

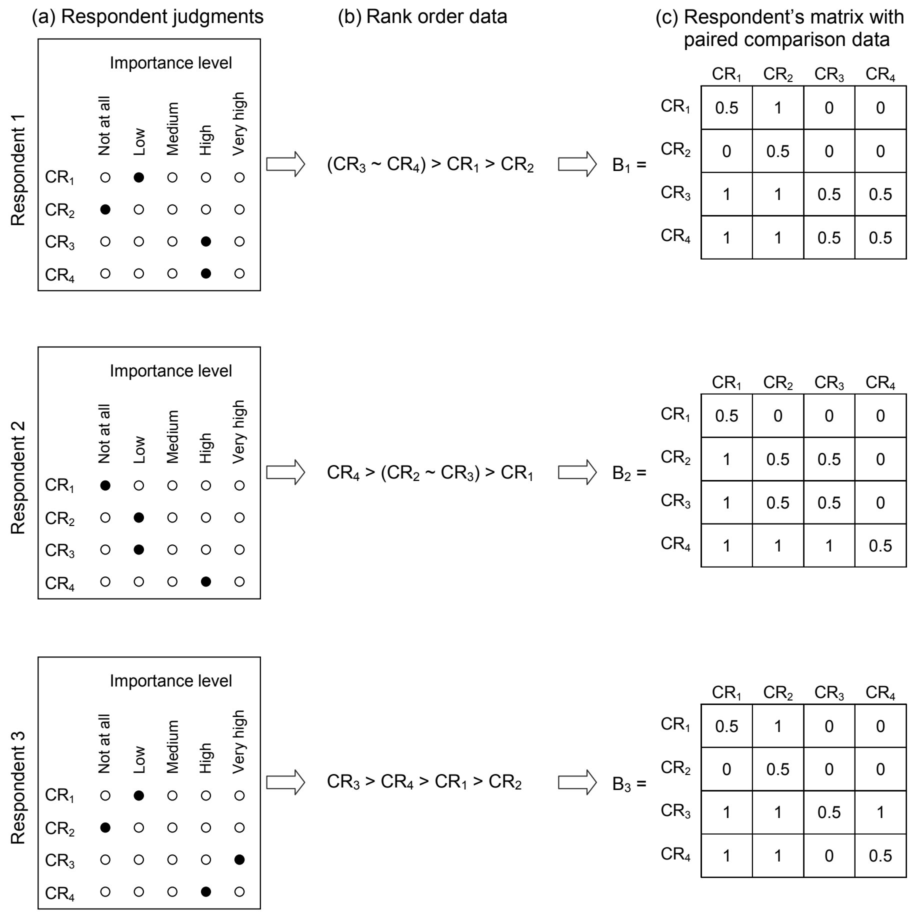

Fig. 4. Process for deriving paired comparison data based on respondents' judgments, defined on a 5-level rating scale. The binary matrices (B_1, B_2 and B_3) in (c) are derived from the results of the questionnaires in (a). The same three fictitious respondents introduced in Fig. 3 are considered. In (b), symbols "~" and ">" respectively mean "indifferent to" and "preferred to".

3. Application example

To exemplify the performance of the proposed approach, this section illustrates an example about the CR prioritization for a civilian aircraft seat, from the perspective of passengers.

Through market survey, a sample of 30 respondents – i.e., regular air passengers – are selected to identify the CRs by individual interview, focus groups and existing information. Finally, 12 major CRs (reported in Tab. 1) are identified to represent the major concerns of customers.

Then, a questionnaire for assessing the level of importance of each of the 12 CRs is submitted to each of the respondents. Results, defined on a 5-level rating scale, are reported in Tab. A.1 (in the appendix).

Abbr.	Description
CR_1	Comfortable (does not give you back ache)
CR_2	Enough leg room
CR_3	Comfortable when you recline
CR_4	Does not hit person behind when you recline
CR_5	Comfortable seat belt
CR_6	Seat belt feels safe
CR_7	Arm rests not too narrow
CR_8	Arm rest folds right away
CR_9	Does not make you sweat
CR_{10}	Does not soak up a spilt drink
CR_{11}	Hole in tray for coffee cup
CR_{12}	Magazines can be easily removed from rack

Tab. 1. List of the major CRs related to an aircraft seat, from the perspective of passengers.

For each respondent, judgements are then transformed into ranked order data and, in turn, into paired comparison data, according to the procedure described in Sect. 2.2. For example, as regards the respondent 1, rank order data are ($CR_5 \sim CR_7$) > ($CR_1 \sim CR_3 \sim CR_6$) > ($CR_2 \sim CR_8 \sim CR_9$) > CR_{10} > ($CR_4 \sim CR_{11} \sim CR_{12}$), which are transformed into the matrix in Fig. A.1 (in the appendix).

We remark that, consistently with the convention introduced in Sect. 2.1, the mutual paired comparisons of two CRs with identical importance are both 0.5 (e.g., see CR_1 and CR_8 or CR_1 and CR_{12} in the matrix in Fig. A.1 (in the appendix).

Next, the paired comparison data matrices relating to the 30 respondents are summed into a single frequency matrix (F), in Fig. A.2 (in the appendix).

The matrix F is transformed into the matrix P (in Fig. A.3, in the appendix) and subsequently into the matrix Z (in Fig. A.4, in the appendix).

According to the convention illustrated in Sect. 2.1, for $p_{ij} \leq 0.001$ and $p_{ij} \geq 0.999$, z_{ij} values have been set to 3.090 and -3.090 respectively (see the p_{ij} values marked with “*”, in the matrix P (Fig. A.3, in the appendix).

Thurstone scale values for each CR are finally calculated through the mean value of the column elements of the matrix Z (see the second row at the bottom of matrix Z, in Fig. A.4 in the appendix). Since the unit and the origin of the resulting interval scale are both arbitrary, we can transform the scale values so that they are included in the interval [1; 5], according to the transformation:

$$\frac{\mu_j' - 1}{5 - 1} = \frac{\mu_j - \min(\mu_j)}{\max(\mu_j) - \min(\mu_j)}, \tag{10}$$

being:

μ_j the scale value related to the j-th CR, resulting from Thurstone scaling;

μ_j' the transformed scale value related to the j-th CR in the interval scale [1, 5].

Fig. 5 provides a graphical representation of the Thurstone scale values, before and after the transformation in Eq. 10. This transformation is nothing else than a monotonically increasing linear function of the type:

$$\mu_j' = a + b \cdot \mu_j, \tag{11}$$

being:

$$a = \frac{\max(\mu_j) - 5 \cdot \min(\mu_j)}{\max(\mu_j) - \min(\mu_j)} \text{ and } b = \frac{4}{\max(\mu_j) - \min(\mu_j)} > 0 .$$

This transformation eases Phases 4 and 8 of the Product Planning HoQ construction process (in Fig. 1), since they are traditionally based on CR importance levels defined on a 1 to 5 scale.

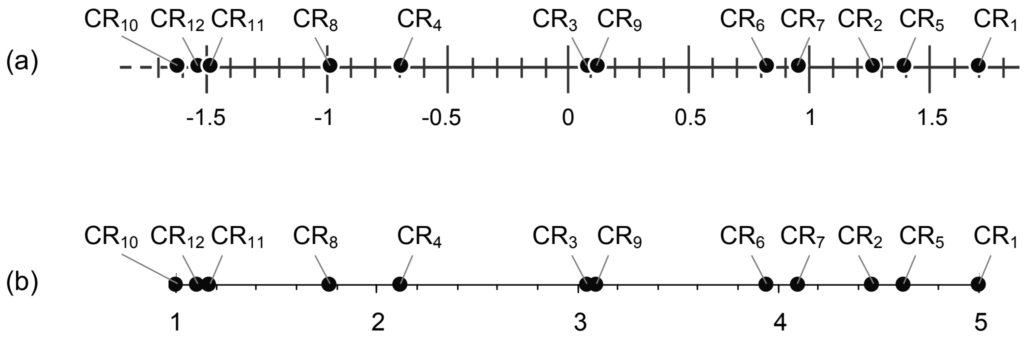

Fig. 5. Resulting Thurstone scale values (a) before and (b) after the scale transformation in Eq. 10.

4. Concluding remarks

The following three subsections respectively discuss (i) the benefits of the proposed procedure, (ii) its limitations and (iii) some ideas for future research.

4.1 Benefits

- The proposed procedure allows aggregation of the typical CR judgments – generally expressed on ordinal response scales – into a continuous interval scale, avoiding the typical abuses (e.g., arbitrary promotion of the scale properties) of the classical approaches (Franceschini et al., 2007).
- Unlike other methods, such as the AHP, ANP or Kano model, the proposed procedure does not require complicated elaborations by respondents. Particularly in populations in which educational attainment and numeracy are limited, a simple measurement strategy may have considerable practical advantages over more complex techniques, such as ease of comprehension and greater reliability due to reduced measurement error. However, the fact remains that the Thurstone model can be extended to more complex response modes, such as questionnaires in

which CRs are ordered or compared in pairs by each respondent.

- The Thurstone model is relatively robust to incomplete data, like the omission of a portion of judgments by respondents. When the incidence of incomplete data is high, the model presented can be replaced by more refined ones, and it is also possible to check the internal consistency of the results obtained (Thurstone, 1927; Edwards, 1957). However, if CRs were identified correctly, the amount of omitted judgments should not be too large. The opposite could mean that the CRs in use do not reflect the real needs of the customer.

4.2 Limitations

- Like any model, that of Thurstone is based on several assumptions, such as, (i) the phenomena to be scaled must lie on a latent unidimensional scale, (ii) the model is based on the normal distribution of *stimuli*, and (iii) dispersion and correlation of the *stimuli* are assumed to be equal. Some of these assumptions can be relaxed when using more sophisticated but also complex variants of the proposed model (Maydeu-Olivares and Böckenholt, 2008).
- The 5-level rating scale for CR judgments is simple and intuitive for the respondent but has a relatively limited resolution. In some cases, this can make the analysis uncertain, since it may generate a significant number of "ties" (i.e., CRs with identical levels of importance), when judgments are transformed into paired comparison data. The problem can be solved by using alternative response scales with a larger number of levels, or questionnaires in which the CRs are ranked or compared in pairs by each respondent.
- As for most of the statistical models, the larger the sample of respondents, the more reasonable and robust results will be. Thurstone (1927) recommends that there be at least a few tens of respondents. This seems to be in line with the typical amount of customers involved in the initial phases of QFD process.

4.3 Ideas for future research

Further research should investigate the relationship between results acquired from different techniques for CR prioritization. Moreover, the use of Thurstone LCJ could be extended to other prioritization processes within QFD, such as Phases 3 and 9 in Fig. 1.

References

Buyukozkan, G., Ertay, T., Kahraman, C., Ruan, D. (2004) Determining the importance weights for the design requirements in the house of quality using the fuzzy analytic network approach. International Journal of Intelligent Systems, 19(5): 443-461.

Burke E., Kloeber Jr J.M., Deckro R.F. (2002) Using and Abusing QFD Scores. Quality Engineering, 15(1):9-21.

Chan, L.K., Wua, M.L. (2002) Quality Function Deployment: A Comprehensive Review of Its Concepts and Methods. Quality Engineering, 15(1): 23-35.

Chan, L.K., Kao, H.P., Wu, M.L. (1999) Rating the importance of customer needs in quality function

deployment by fuzzy and entropy methods. International Journal of Production Research, 37(11): 2499-2518.
Chaudha, A., Jain, R., Singh, A.R., Mishra, P.K. (2011) Integration of Kano's model into quality function deployment (QFD). International Journal of Advanced Manufacturing Technology, 53(5-8): 689-698.
Chuang, P.T. (2001) Combining the analytic hierarchy process and quality function deployment for a location decision from a requirement perspective. International Journal of Advanced Manufacturing Technology, 18(11): 842-849.
Edwards, A.L. (1957) Techniques of Attitude Scale Construction. Irvington Publishers, New York.
Figueira, J., Greco, S., Ehrgott, M. (2005) Multiple criteria decision analysis: state of the art surveys. Springer, New York.
Franceschini, F. (2002) Advanced Quality Function Deployment, St. Lucie Press/CRC Press LLC, Boca Raton, FL.
Franceschini, F., Rossetto, S. (1995) QFD: the problem of comparing technical/engineering design requirements. Research in Engineering Design, 7(4): 270-278.
Franceschini, F., Rossetto, S. (2002) QFD: an interactive algorithm for the prioritization of product's technical design characteristics. Integrated Manufacturing Systems, 13(1): 69-75.
Franceschini, F., Galetto, M., Maisano, D. (2007) Management by Measurement: Designing Key Indicators and Performance Measurement Systems. Springer, Berlin
Gonzalez, M.E., Quesada, G., Bahill, A.T. (2003) Improving product design using quality function deployment: the school furniture case in developing countries. Quality Engineering, 16(1): 47-58.
Griffin, A., Hauser, J.R. (1993) The voice of customer. Marketing Science, 12(1): 1-27.
Hauser, J.R., Clausing, D. (1988) The house of quality. Harvard Business Review, 66(3): 63-73.
Karsak, E.E., Sozer, S., Alptekin, S.E. (2002) Product planning in quality function deployment using a combined analytic network process and goal programming approach. Computers & Industrial Engineering, 44: 171-190.
Krus, D.J., Kennedy, P.H. (1977) Normal scaling of dominance matrices: The domain-referenced model. Educational and Psychological Measurement, 37:189 -193.
Kwong, C.K., Bai, H. (2002) A fuzzy AHP approach to the determination of importance weights of customer requirements in quality function deployment. Journal of Intelligent Manufacturing, 13(5): 367-377.
Jansen, P.G.W. (1984) Relationships between the Thurstone, Coombs, and Rasch approaches to item scaling. Applied Psychological Measurement, 8:373-383.
Lee, Y.T., Wu, W.W., Tzeng, G.H. (2008) An effective decision-making method using a combined QFD and ANP approach. WSEAS Transactions on Business and Economics, 12(5): 541-551.
Li, Y., Tang, J., Luo, X., Xu, J. (2009) An integrated method of rough set, Kano's model and AHP for rating customer requirements' final importance. Expert Systems with Applications, 36(3): 7045-7053.
Luce, R.D., Tukey, J.W., (1964) Simultaneous conjoint measurement: a new type of fundamental measurement. Journal of Mathematical Psychology, 1(1):1-27.
Maydeu-Olivares, A., Böckenholt, U. (2008) Modeling subject health outcomes. Top 10 reasons to use Thurstone's method. Medical Care, 46(4): 346-348.
Matzler, K., Hinterhuber, H.H. (1998) How to make product development projects more successful by integrating Kano's model of customer satisfaction into quality function deployment. Technovation 18(1): 25-38.
McIver, J.P., Carmines, E.G. (1981) Unidimensional Scaling. Beverly Hills, CA: Sage Publications; 1981.
Nahm, Y.E., Ishikawa, H., Inoue, M. (2013) New rating methods to prioritize customer requirements in QFD with incomplete customer preferences. International Journal of Advanced Manufacturing Technology, 65(9-12): 1587-1604.
Rasch, G. (1966) An item analysis which takes individual differences into account. British Journal of Mathematical and Statistical Psychology, 19(1): 49 -57.
Sireli, Y., Kauffmann, P., Ozan, E. (2007) Integration of Kano's model into QFD for multiple product design. IEEE Transactions on Engineering Management, 54(2): 380-390.
Sharma, J.R., Rawani, A.M., Barahate, M. (2008) Quality function deployment: a comprehensive literature review. International Journal of Data Analysis Techniques and Strategies, 1(1): 78-103.
Stevens, S.S. (1946) On the Theory of Scales of Measurement. Science, 103(2684): 677-680.
Tan, K.C., Shen, X.X. (2000) Integrating Kano's model in the planning matrix of quality function deployment. Total Quality Management, 11(8): 1141-1151.
Thurstone, L.L. (1927) A law of comparative judgments. Psychological Review, 34: 273-286.
Tran, T.L., Sherif, J.S. (1995) Quality function deployment (QFD): an effective technique for requirements

acquisition and reuse. Proceedings of the 2nd IEEE Software Engineering Standards Symposium, Montreal, Canada, 191-200.
Urban, G.L., Hauser, J.R. (1993) Design and marketing of new products (Vol. 2). Englewood Cliffs, NJ: Prentice Hall.

Appendix

See the following Figures and Tables.

Respondent No.	CR_1	CR_2	CR_3	CR_4	CR_5	CR_6	CR_7	CR_8	CR_9	CR_{10}	CR_{11}	CR_{12}
1	4	3	4	1	5	4	5	3	3	2	1	1
2	4	3	3	2	4	3	5	1	4	1	1	2
3	5	5	4	3	5	5	4	3	4	1	1	1
4	5	5	4	2	5	5	4	1	3	1	1	2
5	5	4	2	1	4	4	5	2	3	2	2	2
6	5	5	4	3	5	3	3	2	3	1	2	1
7	5	3	3	1	5	4	5	3	4	2	1	2
8	4	3	1	3	5	3	2	1	3	1	1	1
9	5	5	4	4	4	5	2	3	3	1	1	1
10	3	5	2	1	5	2	3	1	2	1	1	2
11	5	3	5	3	3	4	3	1	3	1	2	2
12	5	4	3	1	5	3	3	2	3	1	1	2
13	5	4	3	2	5	3	5	1	1	1	2	1
14	4	5	1	1	3	3	3	2	3	1	1	1
15	4	5	4	3	4	4	5	3	4	1	1	2
16	5	4	3	3	5	5	5	3	3	2	3	1
17	4	4	4	2	5	5	5	3	2	2	2	1
18	5	5	4	2	3	3	2	2	4	1	1	1
19	5	4	4	2	3	4	3	2	3	1	1	2
20	4	5	3	3	5	3	5	2	4	2	1	1
21	5	4	4	2	3	3	3	1	1	1	1	1
22	5	5	4	4	5	4	3	2	3	1	2	1
23	5	4	3	2	4	4	5	2	4	2	3	1
24	5	5	3	3	3	5	3	2	2	2	2	1
25	5	4	2	1	4	3	2	2	4	1	1	1
26	4	4	1	2	5	3	4	2	2	1	1	2
27	4	5	2	2	4	3	3	3	3	1	1	1
28	4	4	2	3	5	3	4	2	4	2	3	1
29	5	5	2	1	4	5	3	1	3	2	1	1
30	5	4	4	1	5	3	5	1	3	2	1	1

Tab. A.1. Levels of importance assigned by 30 respondents to the CRs, through a 5-level rating scale (1=Not at all important, 5=Very high importance).

B_1 =	CR_1	CR_2	CR_3	CR_4	CR_5	CR_6	CR_7	CR_8	CR_9	CR_{10}	CR_{11}	CR_{12}
CR_1	0.5	1	0.5	1	0	0.5	0	1	1	1	1	1
CR_2	0	0.5	0	1	0	0	0	0.5	0.5	1	1	1
CR_3	0.5	1	0.5	1	0	0.5	0	1	1	1	1	1
CR_4	0	0	0	0.5	0	0	0	0	0	0	0.5	0.5
CR_5	1	1	1	1	0.5	1	0.5	1	1	1	1	1
CR_6	0.5	1	0.5	1	0	0.5	0	1	1	1	1	1
CR_7	1	1	1	1	0.5	1	0.5	1	1	1	1	1
CR_8	0	0.5	0	1	0	0	0	0.5	0.5	1	1	1
CR_9	0	0.5	0	1	0	0	0	0.5	0.5	1	1	1
CR_{10}	0	0	0	1	0	0	0	0	0	0.5	1	1
CR_{11}	0	0	0	0.5	0	0	0	0	0	0	0.5	0.5
CR_{12}	0	0	0	0.5	0	0	0	0	0	0	0.5	0.5

Fig. A.1. Paired comparison data relating to the judgments by respondent 1, in Tab. A.1.

F =	CR_1	CR_2	CR_3	CR_4	CR_5	CR_6	CR_7	CR_8	CR_9	CR_{10}	CR_{11}	CR_{12}
CR_1	15.0	19.5	28.0	30.0	17.0	25.0	20.5	30.0	28.0	30.0	30.0	30.0
CR_2	10.5	15.0	25.0	29.0	14.0	20.0	18.5	29.0	25.0	30.0	30.0	30.0
CR_3	2.0	5.0	15.0	23.5	5.5	9.5	9.5	24.0	15.0	27.0	26.0	27.0
CR_4	0.0	1.0	6.5	15.0	1.5	2.5	5.0	17.0	8.0	22.0	22.0	21.5
CR_5	13.0	16.0	24.5	28.5	15.0	20.0	20.0	30.0	25.5	30.0	30.0	30.0
CR_6	5.0	10.0	20.5	27.5	10.0	15.0	14.5	29.5	20.0	30.0	29.5	29.5
CR_7	9.5	11.5	20.5	25.0	10.0	15.5	15.0	27.5	21.0	30.0	30.0	30.0
CR_8	0.0	1.0	6.0	13.0	0.0	0.5	2.5	15.0	5.0	22.0	20.0	21.5
CR_9	2.0	5.0	15.0	22.0	4.5	10.0	9.0	25.0	15.0	28.0	27.0	28.0
CR_{10}	0.0	0.0	3.0	8.0	0.0	0.0	0.0	8.0	2.0	15.0	14.0	15.5
CR_{11}	0.0	0.0	4.0	8.0	0.0	0.5	0.0	10.0	3.0	16.0	15.0	15.0
CR_{12}	0.0	0.0	3.0	8.5	0.0	0.5	0.0	8.5	2.0	14.5	15.0	15.0

Fig. A.2. Matrix F, obtained from the paired comparison data originated from the respondents' judgments (in Tab. A.1).

		CR_1	CR_2	CR_3	CR_4	CR_5	CR_6	CR_7	CR_8	CR_9	CR_{10}	CR_{11}	CR_{12}
	CR_1	0.500	0.650	0.933	1.000*	0.567	0.833	0.683	1.000*	0.933	1.000*	1.000*	1.000*
	CR_2	0.350	0.500	0.833	0.967	0.467	0.667	0.617	0.967	0.833	1.000*	1.000*	1.000*
	CR_3	0.067	0.167	0.500	0.783	0.183	0.317	0.317	0.800	0.500	0.900	0.867	0.900
	CR_4	0.000*	0.033	0.217	0.500	0.050	0.083	0.167	0.567	0.267	0.733	0.733	0.717
	CR_5	0.433	0.533	0.817	0.950	0.500	0.667	0.667	1.000*	0.850	1.000*	1.000*	1.000*
P =	CR_6	0.167	0.333	0.683	0.917	0.333	0.500	0.483	0.983	0.667	1.000*	0.983	0.983
	CR_7	0.317	0.383	0.683	0.833	0.333	0.517	0.500	0.917	0.700	1.000*	1.000*	1.000*
	CR_8	0.000*	0.033	0.200	0.433	0.000*	0.017	0.083	0.500	0.167	0.733	0.667	0.717
	CR_9	0.067	0.167	0.500	0.733	0.150	0.333	0.300	0.833	0.500	0.933	0.900	0.933
	CR_{10}	0.000*	0.000*	0.100	0.267	0.000*	0.000*	0.000*	0.267	0.067	0.500	0.467	0.517
	CR_{11}	0.000*	0.000*	0.133	0.267	0.000*	0.017	0.000*	0.333	0.100	0.533	0.500	0.500
	CR_{12}	0.000*	0.000*	0.100	0.283	0.000*	0.017	0.000*	0.283	0.067	0.483	0.500	0.500

Fig. A.3. Matrix P, obtained from the matrix F (in Fig. A.2).

		CR_1	CR_2	CR_3	CR_4	CR_5	CR_6	CR_7	CR_8	CR_9	CR_{10}	CR_{11}	CR_{12}
	CR_1	0.000	-0.385	-1.501	-3.090	-0.168	-0.967	-0.477	-3.090	-1.501	-3.090	-3.090	-3.090
	CR_2	0.385	0.000	-0.967	-1.834	0.084	-0.431	-0.297	-1.834	-0.967	-3.090	-3.090	-3.090
	CR_3	1.501	0.967	0.000	-0.784	0.903	0.477	0.477	-0.842	0.000	-1.282	-1.111	-1.282
	CR_4	3.090	1.834	0.784	0.000	1.645	1.383	0.967	-0.168	0.623	-0.623	-0.623	-0.573
	CR_5	0.168	-0.084	-0.903	-1.645	0.000	-0.431	-0.431	-3.090	-1.036	-3.090	-3.090	-3.090
Z =	CR_6	0.967	0.431	-0.477	-1.383	0.431	0.000	0.042	-2.128	-0.431	-3.090	-2.128	-2.128
	CR_7	0.477	0.297	-0.477	-0.967	0.431	-0.042	0.000	-1.383	-0.524	-3.090	-3.090	-3.090
	CR_8	3.090	1.834	0.842	0.168	3.090	2.128	1.383	0.000	0.967	-0.623	-0.431	-0.573
	CR_9	1.501	0.967	0.000	-0.623	1.036	0.431	0.524	-0.967	0.000	-1.501	-1.282	-1.501
	CR_{10}	3.090	3.090	1.282	0.623	3.090	3.090	3.090	0.623	1.501	0.000	0.084	-0.042
	CR_{11}	3.090	3.090	1.111	0.623	3.090	2.128	3.090	0.431	1.282	-0.084	0.000	0.000
	CR_{12}	3.090	3.090	1.282	0.573	3.090	2.128	3.090	0.573	1.501	0.042	0.000	0.000
	Σ_j	20.451	15.132	0.974	-8.339	16.722	9.894	11.460	-11.876	1.414	-19.522	-17.851	-18.459
	$\mu_j = \Sigma_j / n$	1.704	1.261	0.081	-0.695	1.394	0.825	0.955	-0.990	0.118	-1.627	-1.488	-1.538
	μ'	5.000	4.468	3.051	2.119	4.627	3.944	4.100	1.765	3.095	1.000	1.167	1.106

Fig. A.4. Matrix Z containing the unit normal deviates (z_{ij}) corresponding to the complementary probabilities ($1 - p_{ij}$) to those in matrix P (p_{ij}). Values of $p_{ij} \leq 0.001$ and ≥ 0.999 (marked with "*" in Fig. A.3) have been conventionally associated with z_{ij} = 3.090 and -3.090 respectively.