Decision concordance with incomplete expert rankings in manufacturing applications

Original
Decision concordance with incomplete expert rankings in manufacturing applications / Franceschini, Fiorenzo; Maisano, Domenico Augusto Francesco. - In: RESEARCH IN ENGINEERING DESIGN. - ISSN 0934-9839. - STAMPA. - 31:4(2020), pp. 471-490. [10.1007/s00163-020-00340-x]

Availability:
This version is available at: 11583/2846813 since: 2020-09-28T09:44:37Z

Publisher:
Springer

Published
DOI:10.1007/s00163-020-00340-x

Decision concordance with incomplete expert rankings in manufacturing applications

Fiorenzo Franceschini[1] and Domenico Maisano[2]

[1] *fiorenzo.franceschini@polito.it* [2] *domenico.maisano@polito.it*

Politecnico di Torino, DIGEP (Department of Management and Production Engineering),
Corso Duca degli Abruzzi 24, 10129, Torino (Italy)

ABSTRACT

The manufacturing field encompasses a number of problems in which some experts formulate their rankings of a set of objects, which should be aggregated into a collective judgment. E.g., consider the aggregation of (i) the opinions of designers on alternative design concepts, (ii) the opinions of reliability/safety engineers on the criticality of a set of failures, (iii) the perceptions of a panel of customers on alternative aesthetic features of a product, etc.. For these problems, the Kendall's *concordance coefficient* (*W*) can be used to express the degree of agreement between experts in a simple and practical way. Unfortunately, this indicator is applicable to *complete* rankings only, while experts often find it more practical to formulate *incomplete* rankings, e.g., identifying only the most/less relevant objects and/or deliberately excluding some of them, if they are not sufficiently relevant or well known.

This research aims at extending the use of the traditional *W* to incomplete rankings, preserving its practical meaning and simplicity. In a nutshell, the proposed methodological approach associates a so-called "midrank" to all objects, even the ones that are not easily comparable with the other ones; subsequently, *W* can be applied to these midranks. The description is supported by several pedagogical examples.

Keywords: Incomplete ranking, Incomparability, Coefficient of concordance, Midrank, Degree of completeness.

1. INTRODUCTION

A widely debated problem in the scientific literature is that of "*m*-rankings", in which each of *m* experts formulates his/her own (subjective) ranking of *n* objects, based on the degree of an attribute of these objects; then, rankings have to be aggregated into a collective judgment (Kendall and Smith, 1939; Kendall, 1963; Keeney and Raiffa, 1993; Agresti, 2010).

This problem of wide cross-cutting nature is debated in various scientific fields, ranging from decision science to social choice theory, psychometrics, voting theory, multi-criteria decision making, etc. (Kelly, 1991; Tideman and Plassmann, 2012; Coaley, 2014). This problem is also debated in the manufacturing field: let us consider, for example, the aggregation of (i) the opinions of different designers on some alternative design concepts, or (ii) the opinions of maintenance/reliability experts on the more critical failures of a production process, or (iii) the opinions of a panel of service/product customers on the degree of importance of some customer needs (Nahm et al., 2013; Franceschini et al., 2015; Franceschini and Maisano, 2019c; Geramian et al., 2019).

A range of aggregation techniques have been produced in the scientific literature over the years (Frey et al., 2010; Katsikopoulos, 2012; Franceschini and Maisano, 2017). These techniques differ in various aspects, including the type of model (e.g., mathematical, statistical, fuzzy techniques, verbal rules of thumb, etc.) and the type of collective judgment (e.g., object rankings, ordinal/interval/ratio scale values, etc.) (Coaley, 2014; De Vellis, 2016; Wang et al. 2017; Çakır, 2018). The different aggregation techniques, although sophisticated, rational and practical, inexorably clash with Arrows' theorem, which determines their "imperfection" with respect to several properties, also known as *fairness* criteria (Franssen, 2005; Hunt, 2007; Reich, 2010; Arrow, 2012; Jacobs et al., 2014; Franceschini and Maisano, 2019a).

This paper takes a step back with respect to the ranking-aggregation problem, dealing with the evaluation of the degree of agreement/conflict of the rankings themselves. This evaluation has relevant practical implications, such as (Franceschini and Maisano, 2015):

- it may lead to the detection of anomalies in the selection of a panel of supposedly homogeneous experts;
- it may indicate possible intrinsic difficulties in reaching a consensus, depicting in some ways the plausibility of the collective judgment;
- with some adjustments, it allows to assess the goodness of alternative aggregation techniques (Franceschini and Maisano, 2019a).

The scientific literature includes several indicators to estimate the degree of agreement of a pair of rankings, based on two steps: (1) decomposing the two rankings into paired-comparison relationships and (2) measuring the association between the above-mentioned relationships. Some of the most popular indicators are the Pearson product-moment correlation coefficient (ρ), the Kendall's Tau coefficient (τ), and the Spearman's coefficient of rank correlation (R) (Gibbons and Chakraborti, 2010; Franceschini and Maisano, 2015). All these indicators allow to estimate the degree of agreement between two rankings and to test the null hypothesis of independence between them, in a simple and rigorous way (Gibbons and Chakraborti, 2010; Agresti, 2010). Unfortunately, for problems characterized by more than two rankings, these indicators are not practical and do not allow simple and effective significance test.

Kendall and Smith (1939) were the first researchers who developed a single coefficient of overall association for more than two rankings are. Their coefficient, called *concordance coefficient* (W), is related to the dispersion in the ranks associated with each object, depending on the rankings formulated by the experts. Besides being relatively simple and effective, the *W* indicator allows to test the null hypothesis of overall independence between rankings, with a specified significance level (Gibbons and Chakraborti, 2010; Agresti, 2010).

An important limitation of *W* is that it is only applicable to *complete* rankings, i.e., rankings including all the objects of interest, according to a hierarchical sequence characterized by relationships of *strict dominance* and/or *indifference* (Nederpelt and Kamareddine, 2004). Unfortunately, the formulation of complete rankings is unsuitable to some practical contexts, such as (Chen and Cheng, 2010):

- those characterized by a relatively large number of objects, which make the formulation of complete rankings potentially uncertain and time-consuming;
- those in which experts have the possibility to exclude some objects from their rankings, because they do not know enough about them and/or are not able to compare them with other objects;
- those in which experts have limited attention (e.g., in the case of telephone or street interviews).

For the above reasons, it seems reasonable to envisage a less rigid response mode, in which experts do not necessarily have to formulate complete rankings but can also formulate incomplete rankings including only the objects with the higher and/or lower degree of the attribute. In addition, experts can choose whether or not to order these objects, or, if they are not familiar with other objects, they can decide to exclude them from their rankings.

Although the information content of incomplete rankings is unequivocally lower than that of complete ones, incomplete rankings are not necessarily less accurate or less useful for practical purposes: it has been observed that experts tend to focus on the objects at the extremes of a ranking, providing more reliable judgments about them to the detriment of the remaining objects (Harzing et al., 2009; Amodio et al., 2016; Lagerspetz, 2016; Vetschera, 2017; Aledo et al., 2018).

The aim of this paper is to adapt the traditional *W* to problems characterized by incomplete rankings, such as the afore-described ones. The scientific literature embraces some attempts to extend the use of *W* to incomplete rankings, although they rely on assumptions that somehow limit the "degree of incompleteness" of rankings. For example, Durbin (1951) proposed a model relying on the assumption that rankings are incomplete in the same symmetrical way as in an incomplete Latin square, or Youden array (Alvo and Cabilio, 1991). Grzegorzewski (2006) presented a fuzzy generalization of *W*, assuming a "non-degenerated" set of rankings, which means that all objects under study have been univocally ranked by at least one of the experts. Lewis and Johnson (1971) proposed an extension of *W* when experts and objects coincide, and each expert evaluates all the others ones except him/herself.

Our intention is instead to avoid imposing constraints such as those mentioned, leaving some freedom to experts in their response mode. From a methodological point of view, we will develop a procedure to associate a so-called "midrank" to each object, even to those that cannot be compared with at least a portion of the other ones (e.g., those omitted or deliberately excluded by the experts). The revisited W, which will then be calculated considering these midranks, can be interpreted as a generalization of the traditional one.

The rest of this paper is organized into four sections. Section 2 gives some hints on the meaning and calculation of the traditional W. Section 3 illustrates the new approach, providing several application examples to incomplete rankings. Section 4 summarizes the original contributions of this paper and its practical implications, limitations and suggestions for future research. Further information is contained in the Appendix section.

2. BASICS ON THE TRADITIONAL CONCORDANCE COEFFICIENT

2.1 Definition and meaning

Let us assume that each of m experts formulates a *complete* ranking of n objects, that is, an ordered sequence characterized by relationships of *strict dominance* (e.g., "$o_i > o_j$") among the possible pairs of objects (e.g., "$o_1 > o_3 > o_2 > \ldots$"). For the sake of simplicity, relationships of *indifference* (e.g., "$o_i \sim o_j$") are – at least for the moment – neglected; however, these relationships will be introduced later on.

For each of these rankings, it is possible to univocally associate rank data to the various objects; these data could be visualized in the form of a so-called *rank table*, i.e., a two-way layout of dimension $m \times n$, with row and column labels designating experts and objects. We denote the ranked observations by R_{ij}, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, so that R_{ij} is the rank of object number j when considering the ranking by expert i (see Table 1).

		Objects				
		o_1	o_2	…	o_n	Row totals
Experts	e_1	R_{11}	R_{12}	…	R_{1n}	$n \cdot (n+1)/2$
	e_2	R_{21}	R_{22}	…	R_{2n}	$n \cdot (n+1)/2$
	⋮	⋮	⋮	⋱	⋮	⋮
	e_m	R_{m1}	R_{m2}	…	R_{mn}	$n \cdot (n+1)/2$
	Col. totals	R_1	R_2	…	R_n	$m \cdot n \cdot (n+1)/2$

Table 1. Visualization of data related to m (complete) rankings, in the form of a rank table containing the ranks (R_{ij}) assigned by each i-th expert to each j-th object.

The i-th row is a permutation of the numbers 1, 2, ..., n – therefore the row totals are $n \cdot (n + 1)/2$ – and the j-th column is the collection of ranks given to the j-th object by all experts. The ranks in each column are then indicative of the agreement between experts: if the j-th object has the same preference relative to all other objects in the opinion of each of the m experts, all ranks in the j-th column will be identical. If this is true for every column, the experts agree perfectly and the respective column totals ($R_1, R_2, \ldots, R_n$) will be some permutation of the numbers:

$$1\cdot m,\ 2\cdot m,\ 3\cdot m,\ \ldots,\ n\cdot m. \tag{1}$$

Since the average column total is $\overline{R} = m\cdot(n + 1)/2$, for perfect agreement between rankings, the sum of squares of deviations of column totals from the average column total will be a constant:

$$\sum_{j=1}^{n}\left[j\cdot m-\frac{m\cdot(n+1)}{2}\right]^2 = m^2\cdot\sum_{j=1}^{n}\left(j-\frac{n+1}{2}\right)^2 = \left[m^2\cdot n\cdot\left(n^2-1\right)\right]/12\,. \tag{2}$$

The actual observed sum of squares of these deviations is:

$$S=\sum_{j=1}^{n}\left[R_j-\frac{m\cdot(n+1)}{2}\right]^2=\sum_{j=1}^{n}\left[R_j-\overline{R}\right]^2. \tag{3}$$

Therefore the value of S for any set of m rankings ranges between zero and $[m^2\cdot n\cdot(n^2 - 1)]/12$, with the maximum value attained in the case of perfect agreement and the minimum value attained when $R_j = \overline{R} = m\cdot(n + 1)/2$ for all j, that is, when each expert's rankings are assigned completely at random so that there is no agreement between experts.

The ratio of S to its maximum value

$$W=\frac{\text{actual observed } S}{S \text{ for perfect agreement}}=\frac{\sum_{j=1}^{n}\left[R_j-\overline{R}\right]^2}{\left[m^2\cdot n\cdot\left(n^2-1\right)\right]/12} \tag{4}$$

therefore provides a measure of agreement between experts or concordance between rankings. This measure, called *Kendall's coefficient of concordance*, ranges between 0 and 1, with 1 designating perfect agreement/concordance and 0 indicating no agreement or independence of rankings (Kendall, 1963; Legendre, 2010).

2.2 Calculation example

Let us assume that experts are able to formulate complete rankings of the eight objects, including *strict dominance* (">") relationships (see Table 2(a)). Then, the W coefficient of these rankings is calculated, with the aim of estimating the degree of agreement of the experts. Table 2(b) contains the corresponding rank table from which W can be determined (through Eq. 4) as:

$$W=\frac{2586}{4200}=61.6\%\,, \tag{5}$$

denoting a certain degree of agreement between the experts.

2.3 Tied objects

Up to now we have assumed that each row of our $m \times n$ rank table is a permutation of the first n integers (see Table 1). If an expert cannot express any preference (or dominance) between two or more objects, or if the objects are actually indistinguishable, we may wish to allow the expert to assign equal ranks, in fact introducing also some relationships of *indifference* (e.g., "$o_i \sim o_j$") among the possible pairs of objects (Fabbris, 2013). If these numbers are the average ranks that each set of tied objects would occupy if a strict dominance relationship could be expressed, the average column total is not changed with respect to the case seen in Sect. 2.1 (i.e., $\overline{R} = m \cdot (n + 1)/2$).

Experts	(a) Complete rankings	(b) Rank table o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	Row totals
e_1	$o_6 > o_3 > o_1 > o_5 > o_4 > o_2 > o_7 > o_8$	3	6	2	5	4	1	7	8	36
e_2	$o_6 > o_3 > o_5 > o_7 > o_1 > o_8 > o_4 > o_2$	5	8	2	7	3	1	4	6	36
e_3	$o_6 > o_3 > o_5 > o_4 > o_8 > o_2 > o_1 > o_7$	7	6	2	4	3	1	8	5	36
e_4	$o_6 > o_5 > o_2 > o_3 > o_1 > o_8 > o_4 > o_7$	5	3	4	7	2	1	8	6	36
e_5	$o_3 > o_6 > o_5 > o_7 > o_1 > o_4 > o_2 > o_8$	5	7	1	6	3	2	4	8	36
e_6	$o_3 > o_5 > o_6 > o_7 > o_2 > o_1 > o_8 > o_4$	6	5	1	8	2	3	4	7	36
e_7	$o_5 > o_6 > o_3 > o_4 > o_1 > o_8 > o_2 > o_7$	5	7	3	4	1	2	8	6	36
e_8	$o_5 > o_6 > o_3 > o_1 > o_2 > o_8 > o_7 > o_4$	4	5	3	8	1	2	7	6	36
e_9	$o_6 > o_2 > o_3 > o_7 > o_8 > o_1 > o_5 > o_4$	6	2	3	8	7	1	4	5	36
e_{10}	$o_3 > o_5 > o_1 > o_6 > o_8 > o_4 > o_2 > o_7$	3	7	1	6	2	4	8	5	36
	Col. totals	49	56	22	63	28	18	62	62	360

Table 2. (a) Complete rankings of eights objects formulated by ten experts and (b) corresponding rank table.

Let us consider the (complete) rankings in Table 3(a), which are identical to those of Table 2(a), except for those formulated by the experts e_5 and e_6, in which some relationships of strict dominance are replaced by relationships of indifference. For example, in the new ranking by e_5, the objects o_5, o_7 and o_1 are tied for 3rd, 4th, and 5th place; since the mean of {3, 4, 5} = 4, ranks would be assigned to the raw data values as follows: {4, 7, 1, 6, 4, 2, 4, 8}.

However, the sum of squares of deviations of any set of n ranks is reduced if there are ties. It can be shown that, for any ranking i = 1, 2, …, m, the corrected value of the denominator of W is (Kendall, 1963; Legendre, 2010):

$$\left[m^2 \cdot n \cdot \left(n^2 - 1\right) - m \cdot \sum_{i=1}^{m} T_i \right] \Big/ 12 , \qquad (6)$$

where $T_i = \sum_{k=1}^{g_i} \left(t_k^3 - t_k\right)$ is a correction factor for ties[1], in which t_k is the number of tied ranks in the k-th group of tied ranks (where a group is a set of values having constant tied rank) and g_i is the number of groups of ties in the set of ranks (ranging from 1 to n) for expert i. Thus, T_i is the

[1] In this case, "ties" are represented by indifference relationships (e.g., "$o_i \sim o_j$").

correction factor required for the set of ranks for expert i. The W version with the correction for tied objects then becomes:

$$W = \frac{\sum_{j=1}^{n} \left[R_j - \overline{R} \right]^2}{\left[m^2 \cdot n \cdot \left(n^2 - 1 \right) - m \cdot \sum_{i=1}^{m} T_i \right] \Big/ 12}. \tag{7}$$

(a) Complete rankings / (b) Rank table

	(a) Complete rankings	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	Row totals
e_1	$o_6 > o_3 > o_1 > o_5 > o_4 > o_2 > o_7 > o_8$	3	6	2	5	4	1	7	8	36
e_2	$o_6 > o_3 > o_5 > o_7 > o_1 > o_8 > o_4 > o_2$	5	8	2	7	3	1	4	6	36
e_3	$o_6 > o_3 > o_5 > o_4 > o_8 > o_2 > o_1 > o_7$	7	6	2	4	3	1	8	5	36
e_4	$o_6 > o_5 > o_2 > o_3 > o_1 > o_8 > o_4 > o_7$	5	3	4	7	2	1	8	6	36
e_5	$o_3 > o_6 > (o_5 \sim o_7 \sim o_1) > o_4 > o_2 > o_8$	4	7	1	6	4	2	4	8	36
e_6	$(o_3 \sim o_5) > o_6 > o_7 > o_2 > o_1 > (o_8 \sim o_4)$	6	5	1.5	7.5	1.5	3	4	7.5	36
e_7	$o_5 > o_6 > o_3 > o_4 > o_1 > o_8 > o_2 > o_7$	5	7	3	4	1	2	8	6	26
e_8	$o_5 > o_6 > o_3 > o_1 > o_2 > o_8 > o_7 > o_4$	4	5	3	8	1	2	7	6	36
e_9	$o_6 > o_2 > o_3 > o_7 > o_8 > o_1 > o_5 > o_4$	6	2	3	8	7	1	4	5	36
e_{10}	$o_3 > o_5 > o_1 > o_6 > o_8 > o_4 > o_2 > o_7$	3	7	1	6	2	4	8	5	36
	Col. totals	48.0	56.0	22.5	62.5	28.5	18.0	62.0	62.5	360

Table 3. (a) Complete rankings of eight objects (o_1 to o_8) formulated by ten experts (e_1 to e_{10}); these rankings include both indifference ("~") and strict dominance (">") relationships. (b) Corresponding rank table.

Again, the range of W is between 0 (full disagreement) and 1 (full agreement). We observe that, in the case of absence of tied objects, i.e., when $T_i = 0 \;\; \forall i$, the formula in Eq. 7 becomes that in Eq. 4. As an example, considering the rank table in Table 3(b), it is obtained $T_i = 0 \;\; \forall i \in \{1, 2, 3, 4, 7, 8\}$, $T_5 = 24$ and $T_6 = 12$; consequently, $\sum_{i=1}^{m} T_i = 36$ and $W = \frac{2539}{[50400 - 10 \cdot 36]/12} = \frac{2539}{4170} = 60.9\%$.

2.4 Test of significance

The null hypothesis of independence among the rankings means that the relevant ranks are allotted randomly by each expert to the set of objects, so that there is no concordance. The appropriate rejection region is represented by relatively large values of W. The exact sampling distribution of W could be determined only by an extensive enumeration process. Exact tables for relatively small values of m and n are given in (van der Laan and Prakken, 1972). For relatively large m values (i.e., $m \geq 7$), an approximation to the sampling distribution may be used for tests of significance. Precisely, W can be described by a chi-square distribution with $n - 1$ degrees of freedom (Gibbons and Chakraborti, 2010). The rejection region for significance level α then is:

$$Q \in \Re \quad \text{for } Q \geq \chi^2_{n-1,\alpha}, \tag{8}$$

Q being defined as:

$$Q=\frac{(n-1)\cdot\sum_{j=1}^{n}\left[R_j-\overline{R}\right]^2}{\left[m\cdot n\cdot\left(n^2-1\right)-\sum_{i=1}^{m}T_i\right]\Big/12}=W\cdot m\cdot(n-1). \quad (9)$$

Returning to the example in Sect. 2.2 (including complete rankings with strict-dominance relationships only) and that in Sect. 2.3 (including complete rankings with strict-dominance and indifference relationships), the Q-values of 43.100 and 42.621 can be determined respectively. Since they are both $\geq \chi^2_{n-1=7,\,\alpha=5\%}=14.067$, the hypothesis of independence among the rankings is rejected in both cases, with a confidence level of $1-\alpha = 95\%$.

3. ADAPTING *W* TO INCOMPLETE RANKINGS

3.1 Introductory example

Sect. 2 has shown that W can be computed after translating each of the m rankings into a relevant string of object ranks. This "translation" is immediate for a *complete* ranking, where one-and-one-only rank can be associated with each object. On the other hand, the problem is more complicated for *incomplete* rankings. For the purpose of example, let us consider a problem in which three experts (e_1 to e_3) formulate incomplete rankings of five objects (o_1 to o_5), as reported in Figure 1. It can be noticed that:

- e_1 indicates only the two most preferred objects (o_1 and o_4), without ordering them. For this reason, a relationship of *incomparability* – which means "we do not know what kind of (*strict dominance* or *indifference*) relationship exists between them" – ("$o_1 \parallel o_4$") has been included;
- e_2 identifies and orders the three most preferred objects;
- e_3 identifies the most preferred object (o_4), the least preferred one (o_3) and claims not to know o_5. The last object is therefore excluded from his/her evaluation.

(a) Incomplete rankings	(b) Objects (explicit)	(implicit)	(excluded)	(c) Rank table o_1	o_2	o_3	o_4	o_5	Row totals
e_1: $(o_1 \parallel o_4) > \ldots$	o_1 and o_4	o_2, o_3 and o_5	None	?	?	?	?	?	?
e_2: $o_2 > (o_1 \sim o_4) > \ldots$	o_1, o_2 and o_4	o_3 and o_5	None	2.5	1	?	2.5	?	?
e_3: $o_4 > \ldots > o_3$ (o_5 excluded)	o_3 and o_4	o_1 and o_2	o_5	?	?	?	?	?	?
			Col. totals	?	?	?	?	?	?

Figure 1. Example of incomplete rankings of five objects, formulated by three experts.

In general, incomplete rankings cannot be uniquely translated into corresponding object ranks. For example, considering the ranking by e_1, the ranks of o_1 and o_4 – although certainly not higher than 2 – cannot be determined univocally. Or, considering the ranking by e_3, it cannot necessarily be said that o_4 is ranked first, since the excluded object (o_5) – if known – could be placed ahead or at the same level of o_4.

3.2 Descriptive parameters

Before tackling the problem of managing incomplete rankings, let us take a step back, introducing some descriptive parameters of a generic incomplete ranking:

- n is the total number of objects of interest;
- k is the number of so-called *isolated* objects, i.e., objects that the expert believes should be excluded from the evaluation since they are not know well enough (Bruggemann and Carlsen, 2011). Each isolated i-th object is therefore considered incomparable with any other j-th object ("$o_i \parallel o_j$").
- t is the number of so-called "t-objects", where "t" stands for "top", i.e., objects explicitly indicated by the expert, with a higher degree of the attribute than all others.
- b is the number of so-called "b-objects", where "b" stands for "bottom", i.e., objects explicitly indicated by the expert, with a lower degree of the attribute than all others.
- $n - t - b - k$ is the number of *remaining* objects, i.e., objects that – although they were not explicitly mentioned by the expert, being neither t nor b-objects – are not excluded from the expert's evaluation; it should be noted that these objects – which are well known – are radically different from the *isolated* ones and are characterized by mutual relations of incomparability.

Depending on the availability/expertise of the expert, t/b-objects can be ordered ("o") or unordered ("u"), i.e., the expert simply mentions them, without building a hierarchy using relationships of strict dominance or indifference. Returning to the example in Figure 1, o_1 and o_4 are unordered t-objects for the incomplete ranking by e_1, while they are ordered t-objects for that by e_2.
The combination of the parameters n, t, b, k and the fact that the t/b-objects are ordered or not determine different types of incomplete rankings, which can be translated into corresponding paired-comparison relationships. E.g., for unordered t-objects, it is not possible to deduce relationships of strict dominance or indifference between them, although it can be argued that the degree of the attribute of these objects is certainly higher than that of the b-objects (see the example in Figure 2).
We now introduce the so-called *degree of completeness* of a generic ranking:

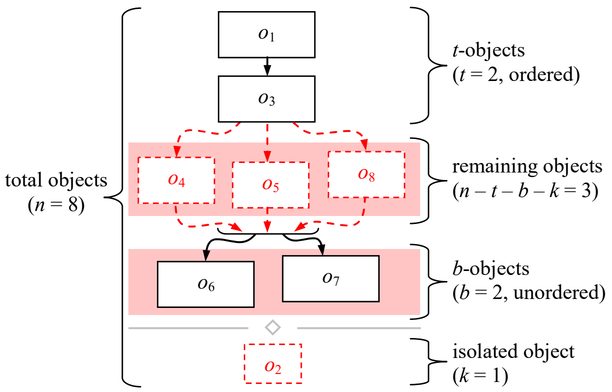
$$c = \frac{\text{No. of "usable" paired comparison relations in the ranking}}{C_2^n}, \tag{10}$$

which expresses the fraction of "usable" paired-comparison relationships – i.e. relationships of strict dominance ("$o_i > o_j$") or indifference ("$o_i \sim o_j$") – with respect to the total ones: $C_2^n = n \cdot (n-1)/2$, where n is the total number of objects of the problem. By way of example, Figure 2 shows the determination of the c values related to a specific incomplete ranking. It can be seen that the more complete the ranking, the higher $c \in [0,1]$; the unit value is reached for complete rankings.

(a) Verbal description by the expert:

"*The top objects are o_1 and o_3, which are already sorted in descending order.*
The bottom objects are o_6 and o_7, which I prefer not to order.
I do not know the object o_2 well enough, so I refrain from evaluating it and exclude it from the ranking."

(b) Incomplete-ranking graph



Analytic form:

$\{(o_1 > o_3 > \{o_4 \| o_5 \| o_8\} > \{o_6 \| o_7\}) \| o_2\}$

(c) Paired-comparison relationships

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8
o_1	-							
o_2	‖	-						
o_3	<	‖	-					
o_4	<	‖	<	-				
o_5	<	‖	<	‖	-			
o_6	<	‖	<	<	<	-		
o_7	<	‖	<	<	<	‖	-	
o_8	<	‖	<	‖	‖	>	>	-

">" and "<" strict dominance;
"‖" incomparability.

Degree of completeness:

c = 17 / 28 = 60.7%

(d) Determination of rank intervals

	L. bound	U. bound	Midrank
o_1	1	2	1.5
o_2	1	8	4.5
o_3	2	3	2.5
o_4	3	6	4.5
o_5	3	6	4.5
o_6	6	8	7.0
o_7	6	8	7.0
o_8	3	6	4.5

Sum = 36.0

Figure 2. Example of incomplete ranking: (a) verbal description of the ranking as it was collected from the expert; (b) graph and description through parameters n, t, b and k; (c) translation of the ranking into paired-comparison relationships and calculation of the degree of completeness c; (d) assignment of rank intervals to each object, according to the rules in Table 4.

Interestingly, even rankings that are significantly incomplete may contain a relevant portion of usable paired-comparison relationships. E.g., consider the incomplete ranking in Figure 2, in which only half of the objects are explicitly indicated by the expert (in fact, o_4, o_5 and o_6 are not among the t/b-objects, while o_2 was intentionally excluded) and some among them (i.e., the b-objects) have not even been ordered; despite the apparently large incompleteness, Figure 2(c) shows that more that sixty percent of the usable paired-comparison relationships are still preserved (c = 60.7%).

The indicator c can be extended from a single ranking to sets of m rankings – such as those characterizing a decision-making problem with m experts. We thus define a new aggregated indicator ($\bar{c}$), depicting the overall degree of completeness:

$$\bar{c} = \frac{\sum_{i=1}^{m} \text{No. of "usable" paired comparison relations in the } i^{\text{th}} \text{ ranking}}{\sum_{i=1}^{m} \text{Total no. of "usable" paired comparison relations in the } i^{\text{th}} \text{ ranking}} = \frac{\sum_{i=1}^{m} c_i \cdot C_2^n}{m \cdot C_2^n} = \frac{\sum_{i=1}^{m} c_i}{m}, \quad (11)$$

c_i being the c value related to the ranking by a generic i-th expert.

Eq. 11 shows that $\bar{c} \in [0,1]$ can also be interpreted as the arithmetic average of the c values related to the set of rankings under consideration.

Figure 4(a) collects the descriptive parameters n, t, b, k, c, $\bar{c}$ related to the three incomplete rankings in Figure 1(a).

3.3 Rationale of the revisited W

Considering a generic i-th incomplete ranking, we can associate a so-called *rank interval* to each j-th object: [R_{ij}^{L} , R_{ij}^{U}], being

- R_{ij}^{L} a lower bound (superscript "L" stands for "lower") corresponding to the minimum possible object rank, in one of the possible complete rankings that are *compatible* with the incomplete one. Any complete ranking is defined as "compatible" if, when translated into paired-comparison relationships, it is characterized by the same usable relationships (of strict dominance or indifference) that characterize the incomplete ranking, excluding those of incomparability. Using the terminology of the Mathematics' Poset Theory, such a complete ranking is said to be a "linear extension" of the incomplete ranking from which it derives (Caperna and Boccuzzo, 2018).
- R_{ij}^{U} an upper bound (superscript "U" stands for "upper") corresponding to the maximum possible object rank, in one of the possible complete rankings that are *compatible* with the incomplete one.

For example, let us consider the incomplete ranking exemplified in Figure 2(b). If, for a moment, we forget the isolated object o_2, the object o_3 would have rank 2 for any compatible complete ranking. However, since the (unknown) rank of o_2 could be higher, lower or tied with respect to that of o_3, the maximum possible rank of o_3 will be 3. The rank interval of o_3 would therefore be [2, 3] (see Figure 2(d)). On the other hand, the rank interval of o_2 will be [1, 8], since this object – which is by definition incomparable to the others – could be in any position of a complete compatible ranking. This reasoning can be extended to all other objects, allowing to determine their corresponding rank intervals.

Table 4 shows a reference scheme to determine the rank intervals of objects, depending on the parameters n, t, b and k. Mathematical arguments supporting the determination of this scheme are contained in section A.1 (in the Appendix).

Using this scheme, it is not necessary to consider all possible complete rankings that are compatible with the incomplete one of interest, which would be extremely complex and computationally expensive, especially for large n values. In that regard, Sect. A.2 (in the Appendix) includes a "mind-expanding" example showing that even for simple incomplete rankings, the number of incomplete rankings compatible with them can be relatively large (*NP-hard* problem) (De Loof et al., 2011; Bruggemann and Carlsen, 2011).

	No. of objects	Rank interval R_{ij}^L	R_{ij}^U	R_{ij}	Midrank sum
t-objects (ordered)	t	r	$r+k$	$r+\frac{k}{2}$	$t\cdot\left(\frac{1+t+k}{2}\right)$
(unordered)	*idem*	1	$t+k$	$\frac{1+t+k}{2}$	*idem*
b-objects (ordered)	b	$s+n-b-k$	$s+n-b$	$s+n-b-\frac{k}{2}$	$b\cdot\left(n+\frac{1-b-k}{2}\right)$
(unordered)	*idem*	$1+n-b-k$	n	$n+\frac{1-b-k}{2}$	*idem*
Isolated objects	k	1	n	$\frac{1+n}{2}$	$k\cdot\left(\frac{1+n}{2}\right)$
Remaining objects	$n-t-b-k$	$t+1$	$n-b$	$\frac{(t+1)+(n-b)}{2}$	$(n-t-b-k)\cdot\frac{(t+1)+(n-b)}{2}$

Note: n is the total number of objects;
t is the number of *t*-objects;
b is the number of *b*-objects;
k is the number of *isolated* objects;
$r \in [1, t]$ is the rank of a generic *t*-object in a (complete) sub-ranking including only the (ordered) *t*-objects;
$s \in [1, b]$ is the rank of a generic *b*-object in a (complete) sub-ranking including only the (ordered) *b*-objects;
R_{ij}^L is the rank-interval lower bound;
R_{ij}^U is the rank-interval upper bound;
R_{ij} is the midrank of the rank interval of interest.

Table 4. Reference scheme to determine the R_{ij}^L, R_{ij}^U, and R_{ij} values of each object, depending on the parameters t, b, k and n.

Returning to the incomplete ranking in Figure 2, we would have the following parameters $n = 8$, $t = 2$ (ordered), $b = 2$ (unordered) and $k = 1$. For the purpose of example, the rank intervals of o_3 and o_6 can be calculated by applying the formulae in Table 4, obtaining:

$$\begin{aligned} &o_3\,(t\text{ - object, ordered}): && \begin{cases} R_{ij}^L = r = 2 \\ R_{ij}^U = r + k = 2 + 1 = 3 \end{cases} \\ &o_6\,(b\text{ - object, unordered}): && \begin{cases} R_{ij}^L = 1 + n - b - k = 1 + 8 - 2 - 1 = 6 \\ R_{ij}^U = n = 8 \end{cases}, \end{aligned} \tag{12}$$

where $r \in [1, t]$ is the rank of a generic *t*-object in a (complete) sub-ranking including only the (ordered) *t*-objects.

In the hypothesis that the rank of each object follows a symmetrical distribution with respect to the relevant rank interval, it can be said that the average value of the ranking coincides with the so-called "midrank", defined as:

$$R_{ij} = \frac{R_{ij}^L + R_{ij}^U}{2}. \quad (13)$$

The "mind-expanding" example in Sect. A.3 (in the Appendix) concerns a specific decision-making problem where the above hypothesis of symmetry seems to be respected. However, the general legitimacy of the hypothesis will be rigorously investigated in future studies.

Following the scheme in Table 4, each object can be uniquely associated with one-and-only-one midrank, based on the parameters t, b, k and n. For example, considering the ranking in Figure 2, the midranks of o_3 and o_6 would be:

$$\begin{aligned} &o_3\,(t\text{-object, ordered}): && r + \frac{k}{2} = 2 + \frac{1}{2} = 2.5 \\ &o_6\,(b\text{-object, unordered}): && n + \frac{1-b-k}{2} = 8 + \frac{1-2-1}{2} = 7 \end{aligned}. \quad (14)$$

Rank data (R_{ij}^L, R_{ij}^U and R_{ij}) related to the objects of a generic incomplete ranking can therefore be collected in a new rank table (Table 5), which is more general than that one seen for complete rankings (Table 1). In the case of complete rankings – i.e. when there are neither isolated objects ($k = 0$) nor remaining ones ($t + b = n$), and the t/b-objects are ordered – it is trivial to demonstrate that $R_{ij}^L = R_{ij}^U = R_{ij}$ for each (i, j) combination. In other words, the rank table for complete rankings (Table 1) can be interpreted as a special case of the new rank table with rank intervals and midranks (Table 5).

Additionally, we notice that the midranks defined in Table 4 are compatible with the convention adopted for the calculation of W in the case of tied object (in Sect. 2.3), i.e., each row total of the rank table should be equal to $n \cdot (n+1)/2$ (see also Table 1). A formal demonstration is contained in Sect. A.1 (in the Appendix).

		Objects				
		o_1	o_2	…	o_n	Row totals
Experts	e_1	R_{11}^L, R_{11}^U $\rightarrow R_{11}$	R_{12}^L, R_{12}^U $\rightarrow R_{12}$	…	R_{1n}^L, R_{1n}^U $\rightarrow R_{1n}$	$n\cdot(n+1)/2$
	e_2	R_{21}^L, R_{21}^U $\rightarrow R_{21}$	R_{22}^L, R_{22}^U $\rightarrow R_{22}$	…	R_{2n}^L, R_{2n}^U $\rightarrow R_{2n}$	$n\cdot(n+1)/2$
	⋮	⋮	⋮	⋱	⋮	⋮
	e_m	R_{m1}^L, R_{m1}^U $\rightarrow R_{m1}$	R_{m2}^L, R_{m2}^U $\rightarrow R_{m2}$	…	R_{mn}^L, R_{mn}^U $\rightarrow R_{mn}$	$n\cdot(n+1)/2$
Col. totals		$R_1 = \Sigma_i(R_{i1})$	$R_2 = \Sigma_i(R_{i2})$	…	$R_n = \Sigma_i(R_{in})$	$m\cdot n\cdot(n+1)/2$

Table 5. Scheme of the new rank table containing the rank intervals [R_{ij}^L, R_{ij}^U] and midranks (R_{ij}) of the objects, for incomplete rankings.

The next step is the application of W to the data contained in the new rank table. Since for the same ranking it is possible to have multiple objects with the same midrank, the formula in Eq. 7, which allows tied (mid)ranks, can be used. Then the significance can be tested as shown in Sect. 2.4.

The flow chart in Figure 3 summarizes the fundamental phases of the proposed approach.

Figure 4(b) shows an example of the construction of the rank table, midranks and W for the incomplete rankings in Figure 1(a).

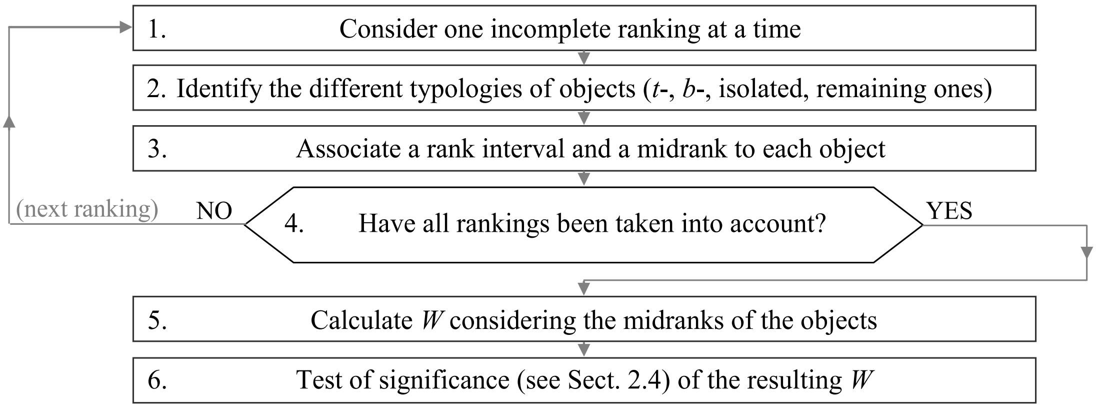


Figure 3. Flow-chart of the proposed procedure for the calculation of W, in the case of incomplete rankings.

(a) Descriptive params.

	n	k	t	b	c
e_1:	5	0	2 (u)	0	60.0%
e_2:	5	0	3 (o)	0	90.0%
e_3:	5	1	1	1	50.0%
					$\bar{c}$ = 66.7%

(b) Rank table

	o_1	o_2	o_3	o_4	o_8	Row totals	T
e_1	1.0, 2.0 → 1.5	3.0, 5.0 → 4.0	3.0, 5.0 → 4.0	1.0, 2.0 → 1.5	3.0, 5.0 → 4.0	15	30
e_2	2.5, 2.5 → 2.5	1.0, 1.0 → 1.0	4.0, 5.0 → 4.5	2.5, 2.5 → 2.5	4.0, 5.0 → 4.5	15	12
e_3	2.0, 4.0 → 3.0	2.0, 4.0 → 3.0	4.0, 5.0 → 4.5	1.0, 2.0 → 1.5	1.0, 2.0 → 1.5	15	24
Col. totals	7.0	8.0	13.0	5.5	11.5	45	ΣT = 66

W = 53.7%

Note: n is the total number of objects;
k is the number of *isolated* objects;
t is the number of t-objects;
b is the number of b-objects;
"(o)" means *ordered* t/b-objects;
"(u)" means *unordered* t/b-objects;
T is a correction factor for ties of each ranking;
c is the degree of incompleteness of each ranking;
$\bar{c}$ depicts the overall degree of incompleteness.

Figure 4. Determination of the descriptive parameters, rank table, midranks and W for the three incomplete rankings in Figure 1(a).

3.4 Application examples

This sub-section contains some application examples of the revisited version of W to six different problems. All these problems include m = 10 experts formulating their rankings of n = 8 objects. These problems are intended to reflect different degrees of completeness.

The first problem (0) is the one already exemplified in Sect. 2.2, characterized by complete rankings; it will be classified as "complete problem". This problem represents a situation in which experts have the competence and the possibility to formulate complete rankings, without omitting/excluding any object.

The next four problems (1 to 4) include incomplete rankings *compatible* with those of the problem (0); for this reason, they will be labelled as "incomplete problems". This compatibility indicates that experts are consistent with themselves, although they formulate rankings with very different degrees of incompleteness in the various problems. These four incomplete problems represent other situations, in which – due to their inability or practical impossibility – not all experts are able to formulate complete rankings; nevertheless, they still give their contribution to the problem through the formulation of incomplete rankings. The sixth problem (5) represents a deliberately extreme situation of incompleteness, in which all the experts – not knowing the totality of the objects – exclude them all from their evaluations!

Tables A.3 to A.8 (in Sect. A.3, in the Appendix) contains the rankings and the parameters characterizing these problems; it can be noticed that the degree of completeness – expressed using c and $\bar{c}$ (Eqs. 10 and 11) – tends to decrease gradually.

Focusing on the incomplete problem (1), the corresponding rank data are reported in Table A.9 (in Sect. A.3, in the Appendix) and the resulting W = 65.0%. This value is very close and even slightly higher than that related to the complete problem (0) in Sect. 2.2 (W = 61.6%). The reason is that – despite the numerator of W

$$Num = \sum_{j=1}^{n} \left[R_j - \bar{R} \right]^2 \tag{15}$$

decreases due to the decreased degree of completeness ($\bar{c}$ = 73.2%) – also the denominator

$$Den = \left[m^2 \cdot n \cdot \left(n^2 - 1 \right) - m \cdot \sum_{i=1}^{m} T_i \right] \Big/ 12 \tag{16}$$

decreases due to the effect of the corrective term for tied ranks, i.e., $\sum_{i=1}^{m} T_i$ in Eq. 7 (see the synthetic results in Table 6(1)).

Problem	Description	$\bar{c}$	$\sum_{i=1}^{m} T_i$	*Num*	*Den*	*W*	*Q*	$\chi^2_{n-1,\alpha}$	Reject H_0?
(0)	complete	100.0%	0	2586	4200	61.6%	43.100	14.067	Yes
(1)	incomplete	73.2%	834	2277	3505	65.0%	45.475	14.067	Yes
(2)	incomplete	57.9%	480	1269	3800	33.4%	23.376	14.067	Yes
(3)	incomplete	42.9%	300	733	3950	18.5%	12.981	14.067	No
(4)	incomplete	25.7%	1392	337	3040	11.1%	7.760	14.067	No
(5)	incomplete	0.0%	5040	0	0	indeterminate	indeterminate	14.067	indeterminate

Table 6. Synthetic results of the application of W to the six problems in Tables A.3 to A.8 (see Sect. A.3 in the Appendix), according to the proposed approach.

In the present case

$$Q = W \cdot m \cdot (n-1) = 45.475 \geq \chi^2_{n-1=7,\, \alpha=5\%} = 14.067\,, \tag{17}$$

leading to reject the null hypothesis (H_0) of independence between rankings, as already done for the complete problem (0), in Sect. 2.4.

Extending the analysis to the next four incomplete problems (2 to 4), the results in Table 6 are obtained (see also the relevant rank tables in Sect. A.3: Tables A.10 to A.14). Interestingly, *W* tends to decrease as the degree of completeness decreases (see also Figure 5). This reduction is mainly determined by the reduction of the numerator of *W* (*Num*), which depicts the variability between the sums of the object ranks related to different experts (see numerical data in Table 6).

Additionally, in the "exaggeratedly" incomplete problem (5), *W* is indeterminate because:

- Numerator (*Num*) is zero, denoting the lowest possible variability;
- Denominator (*Den*) is zero, due to the effect of the corrective term for tied objects, i.e., $\sum_{i=1}^{m} T_i = n \cdot \left(n^2 - 1\right)$ in this case.

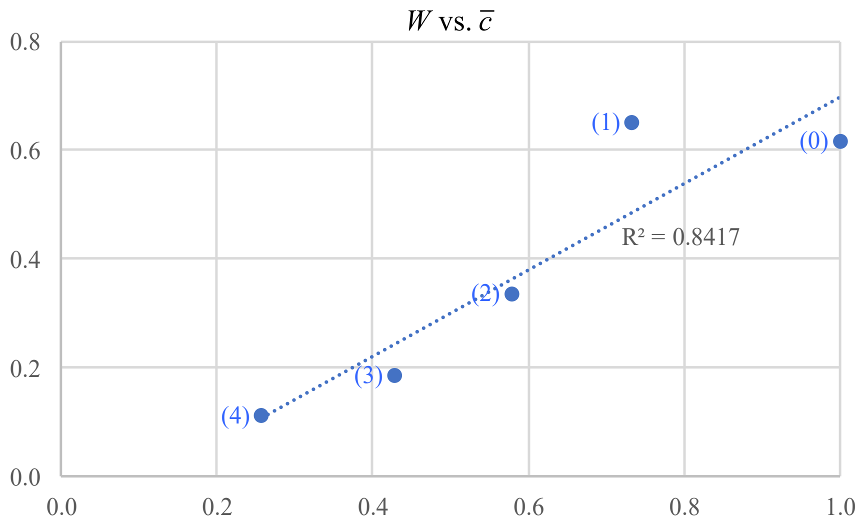


Figure 5. Relationship between the concordance coefficient (*W*) and the degree of completeness ($\bar{c}$), for the exemplified problems (numerical data in Table 6).

Finally, *Q* tends to decrease as *c* decreases, leading not to reject the hypothesis of independence between (incomplete) rankings for the problems (3) and (4).

The authors acknowledge that the results presented in this sub-section – in particular the presumed proportionality between *W* and $\bar{c}$ for problems characterized by compatible rankings (see Figure 5) – are specific to the above six problems and may not necessarily have a general validity. Nevertheless, these results are corroborated by further studies that the authors are currently undertaking on a much larger sample of incomplete problems.

4. CONCLUSIONS

This article revisited the traditional coefficient of concordance *W*, extending its use to incomplete rankings. Although the proposed procedure does not affect the formulation, the practical meaning and the test of significance of the traditional *W*, it requires an additional preliminary phase for the calculation of the object rank intervals and the corresponding midranks. Reversing the perspective, the traditional procedure can be interpreted as a special case of the revisited one, in which rank intervals have coincident bounds and the rank of each object is one-and-only-one.

Preliminary application examples showed that the coefficient of concordance tends to decrease as the degree of completeness of rankings decreases – reflecting situations where experts cannot handle all objects. Additionally, the ranking incompleteness favours the variability in the object midranks, which in turn determines a "polarization" of the rank totals (R_j) towards the average value $\overline{R}$. This in turn produces a reduction of the numerator of W, which is only partially compensated by a reduction of the denominator (see Table 6).

The proposed procedure, although simple, versatile and automatable, has some limitations:

1. The determination of a midrank assumes that the rank of an object is distributed symmetrically with respect to the rank interval. A more rigorous alternative would be to study the actual rank distribution, generating all the complete rankings compatible with the incomplete one (Bruggemann and Carlsen, 2011; De Loof et al., 2011; Caperna and Boccuzzo, 2018).
2. The proposed methodology for calculating midranks only covers incomplete rankings that can be described using the parameters t, b, k and n. For example, it does not cover incomplete rankings with multiple branches (Bruggemann and Carlsen, 2011) or with "anchor" objects (De Loof et al., 2011; Franceschini and Maisano, 2019b).
3. The synthesis of rank intervals into midranks is in some ways questionable because it leads to equalizing objects with very different rank intervals. To clarify this concept, let us consider two different rankings of the objects o_1 and o_2: the first ranking – i.e., "$o_1 \sim o_2$" – is complete, while the second one is assumed to be incomplete since the expert, not knowing both objects, refrains from evaluating them (so both objects can be classified as *isolated*: $n = k = 2$). For the first ranking, the rank intervals of o_1 and o_2 are therefore coincident and equal to [1.5, 1.5]; for the second ranking, they would be again coincident and equal to [1, 2]. It should be noted that, while the width of [1.5, 1.5] is zero, that of [1, 2] is the maximum possible for an object in a ranking with $n = 2$ objects (see Table 4); nevertheless, these different rank intervals are synthesized in the same midrank value: i.e., 1.5. A question now arises: *Is this result acceptable/reasonable?* The authors believe that this concern can be overcome by associating W with an uncertainty indicator – currently under development – which takes into account the width/dispersion of rank intervals before being synthesized into midranks. More details about this issue are contained in Sect. A.2, in the Appendix.

Regarding the future, we plan to extend the testing of *W*, in order to organically investigate the influence of the characteristic parameters of the rankings (e.g., degree of completeness, number of objects, number of experts, etc.).

ACKNOWLEDGMENTS

This research was partially supported by the award "TESUN-83486178370409 finanziamento dipartimenti di eccellenza CAP. 1694 TIT. 232 ART. 6", which was conferred by "Ministero dell'Istruzione, dell'Università e della Ricerca".

REFERENCES

Agresti, A. (2010). Analysis of Ordinal Categorical Data (2nd ed.), New York: John Wiley & Sons. ISBN 978-0-470-08289-8.

Aledo, J. A., Gámez, J. A., Rosete, A. (2018). Approaching rank aggregation problems by using evolution strategies: the case of the optimal bucket order problem. European Journal of Operational Research, 270(3), 982-998.

Alvo, M., Cabilio, P. (1991). On the balanced incomplete block design for rankings. *The Annals of Statistics*, 1597-1613.

Amodio, S., D'Ambrosio, A., Siciliano, R. (2016). Accurate algorithms for identifying the median ranking when dealing with weak and partial rankings under the Kemeny axiomatic approach. European Journal of Operational Research, 249(2), 667-676.

Arrow K.J. (2012) *Social choice and individual values*, 3rd edn. Yale University Press, New Haven.

Bruggemann, R., Carlsen, L. (2011). An improved estimation of averaged ranks of partial orders. *MATCH Commun. Math. Comput. Chem*, 65, 383-414.

Çakır, S. (2018). An integrated approach to machine selection problem using fuzzy SMART-fuzzy weighted axiomatic design. *Journal of Intelligent Manufacturing*, 29(7): 1433-1445.

Caperna, G., Boccuzzo, G. (2018). Use of poset theory with big datasets: A new proposal applied to the analysis of life satisfaction in Italy. *Social Indicators Research*, 136(3), 1071-1088.

Chen, Y.L., Cheng, L.C. (2010). An approach to group ranking decisions in a dynamic environment. *Decision support systems*, 48(4), 622-634.

Coaley, K. (2014), *An introduction to psychological assessment and psychometrics* (2nd edn.), Sage, London.

De Loof, K., De Baets, B., De Meyer, H. (2011). Approximation of average ranks in posets. *Match Commun Math Comput Chem*, 66, 219-229.

DeVellis, R.F. (2016), *Scale development: Theory and applications* (4th edn.), Sage, London.

Durbin, J. (1951). Incomplete blocks in ranking experiments. *British Journal of Statistical Psychology*, 4(2), 85-90.

Fabbris, L. (2013). Measurement scales for scoring or ranking sets of interrelated items. In Survey data collection and integration (pp. 21-43). Springer, Berlin, Heidelberg.

Franceschini, F., Maisano, D. (2015). Checking the consistency of the solution in ordinal semi-democratic decision-making problems. *Omega*, 57, 188-195.

Franceschini, F., Maisano, D., Mastrogiacomo, L. (2015). Customer requirement prioritization on QFD: a new proposal based on the generalized Yager's algorithm. *Research in Engineering Design*, 26(2), 171-187.

Franceschini, F., Maisano, D. (2017). Consistency analysis in quality classification problems with multiple rank-ordered agents. *Quality Engineering*, 29(4):672-689.

Franceschini, F., Maisano, D. (2019a). Design decisions: concordance of designers and effects of the Arrow's theorem on the collective preference ranking. *Research in Engineering Design*, 30(3), 425-434.

Franceschini, F. Maisano, D. (2019b). Fusing incomplete preference rankings in design for manufacturing applications through the ZM_{II}-technique. *International Journal of Advanced Manufacturing Technology*, 103(9-12), 3307-3322.

Franceschini, F., Galetto, M., Maisano, D. (2019c) *Designing Performance Measurement Systems: Theory and Practice of Key Performance Indicators*, Springer International Publishing. Cham, Switzerland, ISBN: 978-3-030-01191-8.

Franssen, M. (2005). Arrow's theorem, multi-criteria decision problems and multi-attribute preferences in engineering design. *Research in Engineering Design*, 16(1-2): 42-56.

Frey, D.D., Herder, P.M., Wijnia, Y., Subrahmanian, E., Katsikopoulos, K., de Neufville, R., Oye, K., Clausing, D.P. (2010). Research in engineering design: the role of mathematical theory and empirical evidence. Research in Engineering Design, 21(3): 145-151.

Geramian A., Abraham A., Ahmadi Nozari M. (2019) Fuzzy logic-based FMEA robust design: a quantitative approach for robustness against groupthink in group/team decision-making. *International Journal of Production Research*; 57(5):1331-1344.

Gibbons, J.D., Chakraborti, S. (2010). *Nonparametric statistical inference* (5th ed.). CRC Press, Boca Raton, ISBN 978-1420077612.

Grzegorzewski, P. (2006). The coefficient of concordance for vague data. *Computational Statistics & Data Analysis*, 51(1), 314-322.

Harzing, A.W., Baldueza, J., Barner-Rasmussen, W., Barzantny, C., Canabal, A., Davila, A., Espejo, A., Ferreira, R., Giroud, A., Koester, K., Liang, Y.K., Mockaitis, A., Morley, M.J., Myloni, B., Odusanya, J.O.T., O'Sullivan., S.L., Palaniappan, A.K., Prochno, P., Roy Choudhury, S., Saka-Helmhout, A., Siengthai, S., Viswat, L., Uzuncarsili Soydas, A., Zander, L. (2009), Rating versus ranking: What is the best way to reduce response and language bias in cross-national research?, *International Business Review*, 18(4):417-432.

Hunt, E. (2007) The Mathematics of Behavior. Cambridge University Press. ISBN 9780521850124. The chapter "*Defining Rationality: Personal and Group Decision Making*" has a detailed discussion of the Arrow Theorem, with proof.

Katsikopoulos, K.V. (2012). Decision methods for design: insights from psychology. *Journal of Mechanical Design*, 134(8):084504.

Keeney, R.L., Raiffa, H. (1993) *Decisions with multiple objectives: preferences and value trade-offs*, Cambridge University Press, Cambridge.

Kelly, J.S. (1991) Social choice bibliography, *Social Choice and Welfare*, 8, 97–169

Kendall M.G., Smith B.B. (1939), The problem of *m*-rankings, *Annals of Math. Statist*, 10:275-287.

Kendall, M.G. (1945), The treatment of ties in ranking problems. *Biometrika*, 239-251.

Kendall, M.G. (1963), Rank Correlation Methods (3rd ed.), New York, Hafner, chap. 6.

Jacobs, J.F., van de Poel, I., Osseweijer, P. (2014). Clarifying the debate on selection methods for engineering: Arrow's impossibility theorem, design performances, and information basis. Research in Engineering Design, 25(1), 3-10.

Laan, P. van der and J. Prakken (1972), Exact distribution of Durbin's distribution-free test statistic for balanced incomplete block designs, and comparison with the chi-square and F approximation, Statistica Neerlandica, 26, 155–164.

Lagerspetz, E. (2016), *Social choice and democratic values*, Springer, Heidelberg.

Legendre, P. (2010). Coefficient of concordance. Encyclopedia of research design, 1, 164-169.

Lewis, G.H., Johnson, R.G. (1971). Kendall's Coefficient of Concordance for sociometric rankings with self excluded. *Sociometry*, 496-503.

Nahm, YE., 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.

Nederpelt, R., Kamareddine, F. (2004), *Logical Reasoning: A First Course*, King's College Publications, London.

Reich Y. (2010), My method is better! *Research in Engineering Design*, 21(3): 137-142.

Tideman, T.N., Plassmann, F. (2012). Modeling the outcomes of vote-casting in actual elections. In Electoral Systems (pp. 217-251). Springer, Berlin, Heidelberg.

Vetschera, R. (2017). Deriving rankings from incomplete preference information: A comparison of different approaches. European Journal of Operational Research, 258(1), 244-253.

Wang, G.X., Huang, S.H., Yan, Y., Du J.J. (2017), Reconfiguration schemes evaluation based on preference ranking of key characteristics of reconfigurable manufacturing systems. *International Journal of Advanced Manufacturing Technology*, 89(5-8):2231-2249.

APPENDIX

A.1 Details on the reference scheme to determine midranks

This section provides some mathematical arguments to test the formulae contained in the scheme in Table 4. The first four sub-sections are dedicated respectively to (1) *t*-objects (ordered or not), (2) *b*-objects (ordered or not), (3) *isolated* objects and (4) *remaining* objects. The last sub-section provides a mathematical demonstration concerning the sum of the midranks of all the objects of a generic incomplete ranking.

A.1.1 t-objects

In case the *t*-objects are ordered, let us assume that r is the rank of a generic *t*-object within a sub-ranking consisting of the *t*-objects only. If there are no isolated objects (i.e. objects intentionally excluded from expert evaluation), the rank of each *t*-object can only be r for any compatible complete ranking. In the presence of (k) isolated objects, the maximum possible rank of the generic *t*-object can be ($r + k$), in the hypothesis that all the isolated objects are placed ahead of it. In conclusion, we would have $R_{ij}^L = r$ and $R_{ij}^U = r + k$, with a corresponding midrank of $R_{ij} = \left(R_{ij}^L + R_{ij}^U\right)/2 = r + k/2$.

Let us now consider the case in which the *t*-objects are not ordered. In the best case, a generic *t*-object could be alone at the top of the complete compatible ranking, therefore $R_{ij}^L = 1$. In the worst case, in presence of (*k*) isolated objects, a generic *t*-object will be alone at the bottom of the sub-ranking consisting of the (*t*) *t*-objects and below the isolated objects themselves; it will therefore be in the position: $R_{ij}^U = t + k$. The corresponding midrank will therefore be: $R_{ij} = \frac{1+(t+k)}{2}$. In this case, the sum of the midranks of all the (*t*) *t*-objects would be: $t \cdot \left(\frac{1+t+k}{2}\right)$.

This result would be obtained also for ordered *t*-objects, in fact:

$$\sum_{r=1}^{t}\left(r + \frac{k}{2}\right) = (1+2+...+t) + t \cdot \frac{k}{2} = \frac{t \cdot (t+1)}{2} + t \cdot \frac{k}{2} = t \cdot \left(\frac{1+t+k}{2}\right).$$

A.1.2 b-objects

In case the *b*-objects are ordered, let us assume that *s* is the rank of a generic *b*-object within a sub-ranking consisting of the *b*-objects only. If there are no isolated objects, in any possible compatible complete ranking, the rank of each *b*-object can only be *s* plus the rank of the objects ahead of it, i.e., the (*t*) *t*-objects and the ($n - t - b - k$) remaining objects (see Sect. 3.2); the resulting rank will therefore be: $s + t + (n - t - b - k) = s + n - b - k$.

In the presence of (*k*) isolated objects, the maximum possible rank of the generic *b*-object can be $(s + n - b - k) + k = s + n - b$, in the hypothesis that all the (*k*) isolated objects are placed ahead of it. In conclusion, we would have $R_{ij}^L = s + n - b - k$ and $R_{ij}^U = s + n - b$, with a corresponding midrank of $R_{ij} = \left(R_{ij}^L + R_{ij}^U\right)/2 = s + n - b - k/2$.

Let us now consider the case in which the *b*-objects are not ordered. In the best case, a generic *b*-object could be alone at the top of the *b*-objects, therefore immediately behind the *t*-objects and the remaining objects; therefore $R_{ij}^L = 1 + t + (n - t - b - k) = 1 + n - b - k$. In the worst case, in presence of (*k*) isolated objects, a generic *b*-object will be alone at the bottom of a compatible complete ranking; it will therefore be in the position: $R_{ij}^U = n$. The corresponding midrank will therefore be: $R_{ij} = n + \frac{1-b-k}{2}$. In this case, the sum of the midranks of all the (*b*) *b*-objects would be: $b \cdot \left(n + \frac{1-b-k}{2}\right)$.

The last result would also be obtained for ordered b-objects, in fact: $\sum_{s=1}^{b}\left(s+n-b-\frac{k}{2}\right)=(1+2+...+t)+b\cdot\left(n-b-\frac{k}{2}\right)=\frac{b\cdot(b+1)}{2}+b\cdot\left(n-b-\frac{k}{2}\right)=b\cdot\left(n+\frac{1-b-k}{2}\right)$.

A.1.3 Isolated objects

Being deliberately excluded as not sufficiently known, these objects could be placed in any position of the complete compatible rankings. In extreme cases, each of these isolated objects could then be placed:

- alone at the top, ahead of the other objects ($R_{ij}^{L}=1$);
- alone in the bottom, below the other objects ($R_{ij}^{U}=n$).

The corresponding midrank will therefore be: $R_{ij}=\frac{R_{ij}^{L}+R_{ij}^{U}}{2}=\frac{1+n}{2}$. The sum of the midranks related to all the (k) isolated objects is: $k\cdot\left(\frac{1+n}{2}\right)$.

A.1.4 Remaining objects

These objects – which are not explicitly mentioned by the expert in the incomplete ranking – will be in an intermediate zone between the t-objects and the b-objects. Since these ($n – t – b – k$) objects are not ordered, the R_{ij}^{L} and R_{ij}^{U} values will be the same for all; the lowest possible rank will be the one immediately after the t-objects, i.e., $R_{ij}^{L}=t+1$, while the highest possible rank will be the one immediately before the b-objects, i.e., $R_{ij}^{L}=n-b$. The corresponding midrank is: $R_{ij}=\frac{R_{ij}^{L}+R_{ij}^{U}}{2}=\frac{(t+1)+(n-b)}{2}$, while the sum of the midranks of all remaining objects is: $(n-t-b-k)\cdot\frac{(t+1)+(n-b)}{2}$.

A.1.5 Proof concerning the sum of all object midranks

This sub-section contains a proof that the midranks in Table 4 are compatible with the convention adopted by Kendall for the calculation of W (in Sect. 2.1), i.e., that each row total of the rank table is equal to $n\cdot(n+1)/2$. The proof is that adding the elements contained in the last column of Table 4 ("Midrank sum"), it can be obtained:

$$t\cdot\left(\frac{1+t+k}{2}\right)+b\cdot\left(n+\frac{1-b-k}{2}\right)+k\cdot\left(\frac{1+n}{2}\right)+(n-t-b-k)\cdot\frac{(t+1)+(n-b)}{2}=$$
$$\frac{1}{2}\cdot(t+t^2+t\cdot k+2\cdot b\cdot n+b-b^2-b\cdot k+k+n\cdot k+n\cdot t+n+n^2-n\cdot b-t^2-t+ \quad \text{(A.1)}$$
$$-n\cdot t+b\cdot t-b\cdot t-b-b\cdot n+b^2-t\cdot k-k-n\cdot k+b\cdot k)=\frac{1}{2}\cdot n\cdot(n+1)$$

A.2 Mind-expanding example

This section contains a further example with a double purpose:

1. Providing a preliminary assessment of whether the proposed procedure is based on legitimate assumptions and provides plausible results;
2. Justifying the need to integrate W with an additional uncertainty indicator (which will be developed in a future research).

Let us consider two decisional problems, both characterized by four experts (e_1 to e_4) formulating individual rankings of four objects (o_1 to o_4). In the first case, rankings are *complete*, while in the second case are *incomplete*. Figure A.1 shows (a) these two sets of rankings and (b) their mutual compatibility (cf. definition of *compatibility* in Sect. 3.3). Additionally, Figure A.1(c) shows that both problems produce identical rank tables and, consequently, identical W values.

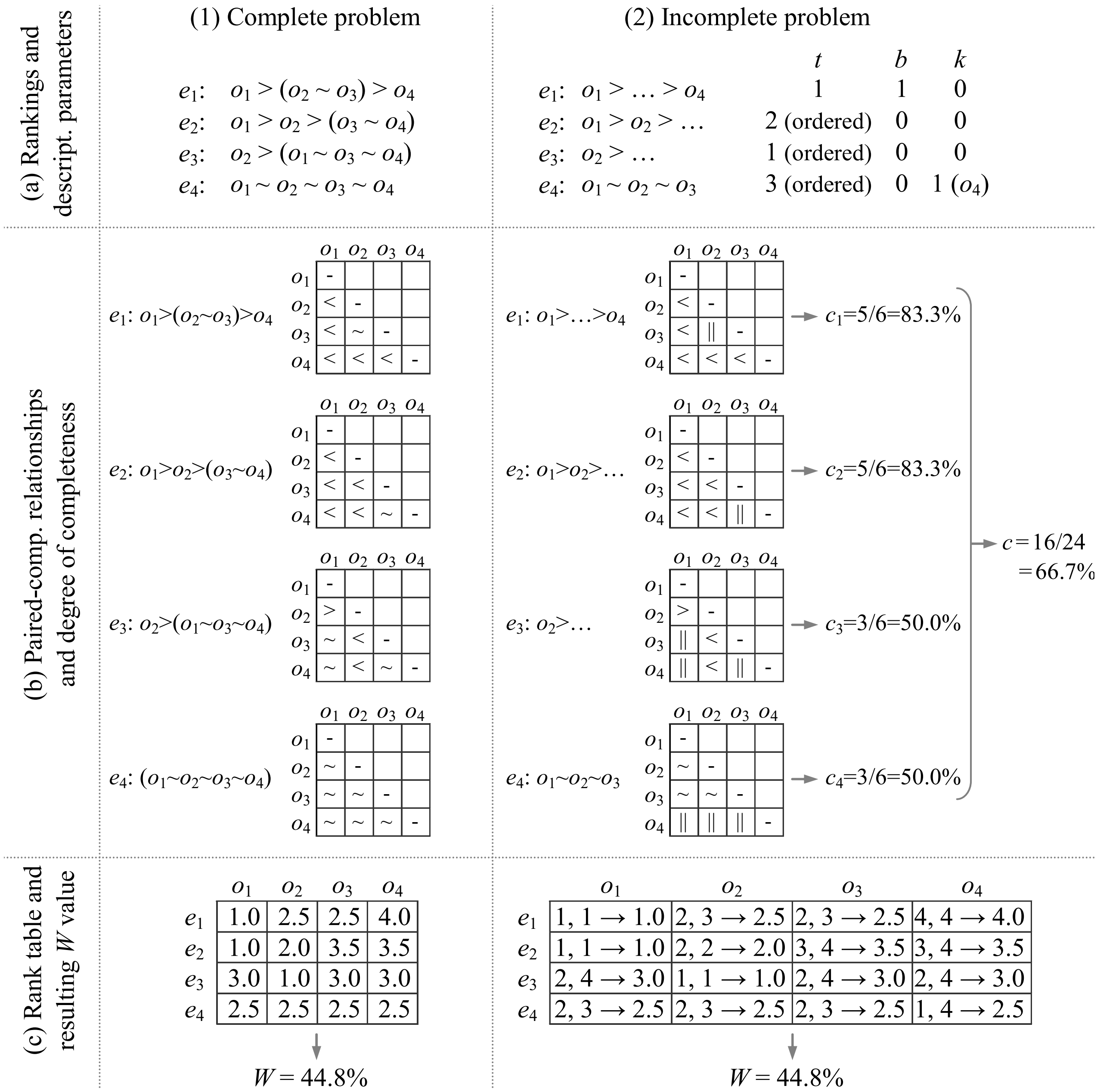


Figure A.1. Example of two decisional problems (one complete and one incomplete) characterized by identical rank tables and W values. The complete problem is compatible with the incomplete one.

The suggested example brings out a somewhat questionable aspect: i.e., the procedure for calculating midranks actually equals objects with very different rank intervals. To clarify this concept, let us consider the (complete and incomplete) rankings by expert e_4 in Figure A.1. Regarding the complete ranking, the rank interval of o_4 is [2.5, 2.5] (absence of dispersion); regarding the incomplete ranking, it is [1, 4] (maximum possible dispersion for a rank interval related to a ranking with $n = 4$ objects). So, the proposed procedure synthesizes these two radically different rank intervals into the same midrank: i.e., 2.5. *Is this result acceptable/reasonable?*

Reflecting on the proposed procedure, the first part is a conventional transformation of each rank interval into a single *equivalent* rank (i.e., midrank). Of course, different conventional constructions could be adopted, resulting in different W values. This aspect raises a further question: *Of all the possible ways of determining W for incomplete rankings, does the proposed one rely on reasonable hypotheses and provide plausible results?* To provide a comprehensive answer to the above question, it would be necessary to carry out a structured study, as we plan to do in the future. Nevertheless, some preliminary arguments to support the proposed technique are presented below, in the form of comments about the results of the previous example.

Let us consider a rigorous but also very laborious way of determining W for problems with incomplete rankings; this method is based on three steps (De Loof et al., 2011; Bruggemann and Carlsen, 2011):

- For each incomplete ranking, all the possible compatible complete rankings are generated.
- Combining the above complete rankings, all the possible complete problems that are compatible with the initial incomplete one are identified.
- The W value related to each complete problem is determined. Then, the distribution of the resulting W values is constructed and studied.

Table A.1 exemplifies this exercise for the four incomplete rankings in Figure A.1(2). Even very simple incomplete rankings may generate a relatively large number of compatible complete rankings, e.g., the one formulated by expert e_3 generates thirteen complete rankings. In fact, this problem can be classified as *NP-hard*, as its complexity increases exponentially with the number of objects and experts (Bruggemann and Carlsen, 2011).

Considering the set of rankings compatible with a certain incomplete ranking, it is interesting to examine the rank distributions. For example, Figure A.2(a) shows the rank distributions related to the thirteen rankings compatible with the incomplete ranking by expert e_3. Analyzing these distributions reveals some interesting aspects:

1. All four distributions, respectively referred to each of the four objects (o_1 to o_4), are symmetrical. This symmetry is probably related to the structure of the incomplete rankings considered in this specific case (Caperna and Boccuzzo, 2018). In the future, we plan to assess the legitimacy of the symmetry hypothesis in more rigorous and general terms.
2. Due to the aforesaid symmetry, the average value of each distribution coincides with the midrank (cf. Figure A.2 and Figure A.1(c)).
3. The distributions of the object ranks are generally correlated with each other, as also exemplified in the Pearson correlation matrix in Table A.2.

Experts:	e_1	e_2	e_3	e_4
Incompl. rankings:	$o_1 > \dots > o_4$	$o_1 > o_2 > \dots$	$o_2 > \dots$	$o_1 \sim o_2 \sim o_3$ (o_2 excluded)
Compatible compl. rankings:	(1) $o_1 > (o_2 \sim o_3) > o_4$	(1) $o_1 > o_2 > (o_3 \sim o_4)$	(1) $o_2 > o_1 > o_3 > o_4$	(1) $(o_1 \sim o_2 \sim o_3) > o_4$
	(2) $o_1 > o_2 > o_3 > o_4$	(2) $o_1 > o_2 > o_3 > o_4$	(2) $o_2 > o_1 > o_4 > o_3$	(2) $o_4 > (o_1 \sim o_2 \sim o_3)$
	(3) $o_1 > o_3 > o_2 > o_4$	(3) $o_1 > o_2 > o_4 > o_3$	(3) $o_2 > o_1 > (o_3 \sim o_4)$	(3) $o_1 \sim o_2 \sim o_3 \sim o_4$
			(4) $o_2 > o_3 > o_1 > o_4$	
			(5) $o_2 > o_3 > o_4 > o_1$	
			(6) $o_2 > o_3 > (o_1 \sim o_4)$	
			(7) $o_2 > o_4 > o_1 > o_3$	
			(8) $o_2 > o_4 > o_3 > o_1$	
			(9) $o_2 > o_4 > (o_1 \sim o_3)$	
			(10) $o_2 > (o_1 \sim o_3) > o_4$	
			(11) $o_2 > (o_1 \sim o_4) > o_3$	
			(12) $o_2 > (o_3 \sim o_4) > o_1$	
			(13) $o_2 > (o_1 \sim o_3 \sim o_4)$	

Table A.1. Possible complete rankings, which are compatible with the four incomplete rankings at the top.

(a) Ranks

	o_1	o_2	o_3	o_4
	2.0	1.0	3.0	4.0
	2.0	1.0	4.0	3.0
	2.0	1.0	3.5	3.5
	3.0	1.0	2.0	4.0
	4.0	1.0	2.0	3.0
	3.5	1.0	2.0	3.5
	3.0	1.0	4.0	2.0
	4.0	1.0	3.0	2.0
	3.5	1.0	3.5	2.0
	2.5	1.0	2.5	4.0
	2.5	1.0	4.0	2.5
	4.0	1.0	2.5	2.5
	3.0	1.0	3.0	3.0
Mean	3.0	1.0	3.0	3.0
St. dev.	0.76	0.00	0.76	0.76

(b) Distributions

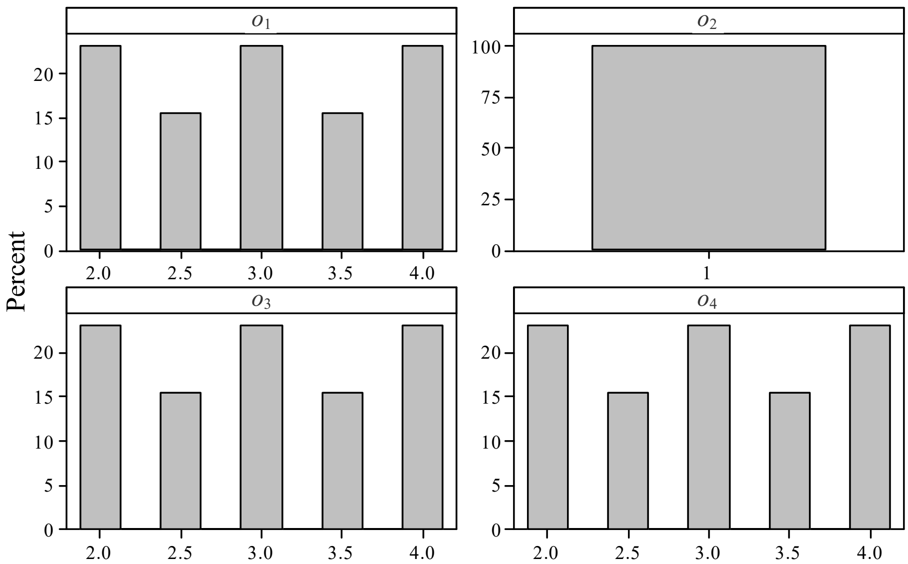


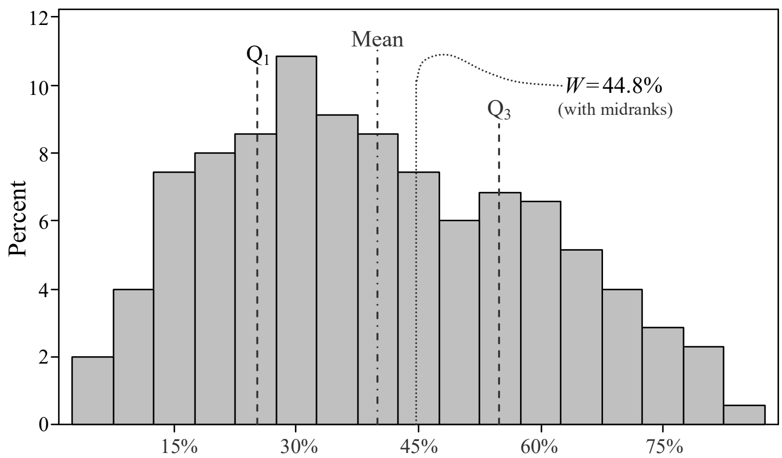
Figure A.2. Ranks and corresponding distributions related to the four objects (o_1 to o_4), considering the thirteen complete rankings compatible with the incomplete ranking by expert e_3 (see Table A.1).

Subsequently, we consider all the possible combinations between the sets of compatible complete rankings in Table A.1, i.e., (i) the three complete rankings related to the incomplete rankings by e_1, (ii) the three ones related to the incomplete rankings by e_2, (iii) the thirteen ones related to the incomplete rankings by e_3, and (iv) three ones related to the incomplete rankings by e_4. Consequently, $3 \cdot 3 \cdot 13 \cdot 3 = 351$ complete decisional problems, which are compatible with the initial (incomplete) one, can be identified. For each of these problems, it is then possible to determine a corresponding value of W (through the traditional procedure in Sect. 2) and then to study the corresponding W distribution. We point out that these 351 complete problems arise from an incomplete problem with a relatively small number of experts and objects. This denotes the low sustainability of the proposed construction, for realistic problems characterized by a large number of objects and/or experts.

Object	o_1	o_2	o_3	o_4
o_1	1			
o_2	0	1		
o_3	-0.5	0	1	
o_4	-0.5	0	-0.5	1

Table A.2. Pearson correlation table of the ranks related to the four objects (o_1 to o_4), considering the thirteen complete rankings compatible with the incomplete ranking by expert e_3 (see Table A.1).

The histogram in Figure A.3 represents the distribution of the W values resulting from the example. We note that this distribution is slightly right-skewed (median below mean value) and has a relatively high dispersion (standard deviation of 19.3%). The mean value of W is equal to 39.7%. Interestingly, the W value determined through the procedure based on midranks (i.e., 44.8%, as shown in Figure A.1) is relatively close to the previous mean value, denoting a certain plausibility of the procedure itself.



Descriptive statistics	
Total count	351
Mean	39.7%
St. dev.	19.3%
Minimum	5.6%
Maximum	86.1%
Q_1	25.0%
Median	37.5%
Q_3	54.3%

Figure A.3. Histogram and descriptive statistics of the W distribution related to the $3 \cdot 3 \cdot 13 \cdot 3 = 351$ possible complete problems examined.

The authors are currently developing a new technique to determine an indicator that expresses the uncertainty of the *W* values, calculated using midranks; this indicator should somehow take into account the dispersion of the initial rank-intervals, without neglecting any correlations between them. In addition, this new indicator could be interpreted as a proxy for the dispersion of the above-exemplified *W* distribution, which avoids to carry out such a laborious construction.

A.3 Further data concerning the six problems exemplified

Tables A.3 to A.8 illustrate the rankings and descriptive parameters (n, t, b, k, c and $\bar{c}$) related to the six problems exemplified in Sect. 3.4. For each ranking, the (k) isolated objects are specified in brackets; "(o)" denotes t/b-objects in case they are ordered, while "(u)" in case they are unordered.

Expert	(0) Complete problem Rankings	t	b	k	c
e_1	$o_6>o_3>o_1>o_5>o_4>o_2>o_7>o_8$	4 (o)	3 (o)	0	100.0%
e_2	$o_6>o_3>o_5>o_7>o_1>o_8>o_4>o_2$	4 (o)	3 (o)	0	100.0%
e_3	$o_6>o_3>o_5>o_4>o_8>o_2>o_1>o_7$	4 (o)	3 (o)	0	100.0%
e_4	$o_6>o_5>o_2>o_3>o_1>o_8>o_4>o_7$	4 (o)	3 (o)	0	100.0%
e_5	$o_3>o_6>o_5>o_7>o_1>o_4>o_2>o_8$	4 (o)	3 (o)	0	100.0%
e_6	$o_3>o_5>o_6>o_7>o_2>o_1>o_8>o_4$	4 (o)	3 (o)	0	100.0%
e_7	$o_5>o_6>o_3>o_4>o_1>o_8>o_2>o_7$	4 (o)	3 (o)	0	100.0%
e_8	$o_5>o_6>o_3>o_1>o_2>o_8>o_7>o_4$	4 (o)	3 (o)	0	100.0%
e_9	$o_6>o_2>o_3>o_7>o_8>o_1>o_5>o_4$	4 (o)	3 (o)	0	100.0%
e_{10}	$o_3>o_5>o_1>o_6>o_8>o_4>o_2>o_7$	4 (o)	3 (o)	0	100.0%
					$\bar{c}$ = 100.0%

Note: "(o)" stands for *ordered t*/*b*-objects.

Table A.3. Complete problem (0) and respective descriptive parameters.

Expert	(1) First incomplete problem Rankings	t	b	k	c
e_1	$o_6>o_3>o_1>o_5>o_4>o_2>o_7>o_8$	4 (o)	3 (o)	0	100.0%
e_2	$o_6>o_3>o_5>\{o_7\|\|o_1\}>o_8>o_4>o_2$	3 (o)	3 (o)	0	96.4%
e_3	$o_6>o_3>o_5>\{o_4\|\|o_8\}>o_2>o_1>o_7$	3 (o)	3 (o)	0	96.4%
e_4	$\{o_6\|\|o_5\|\|o_2\}>\{o_3\|\|o_1\}>\{o_8\|\|o_4\|\|o_7\}$	3 (u)	3 (u)	0	75.0%
e_5	$o_3>o_6>o_5>\{o_7\|\|o_1\|\|o_4\|\|o_2\|\|o_8\}$	3 (o)	0	0	64.3%
e_6	$o_3>o_5>o_6>\{o_7\|\|o_2\|\|o_1\|\|o_8\|\|o_4\}$	3 (o)	0	0	64.3%
e_7	$o_5>o_6>o_3>\{o_4\|\|o_1\|\|o_8\|\|o_2\|\|o_7\}$	3 (o)	0	0	64.3%
e_8	$\{o_5\|\|o_6\|\|o_3\}>\{o_1\|\|o_2\|\|o_8\|\|o_7\|\|o_4\}$	3 (u)	0	0	53.6%
e_9	$o_6>o_2>o_3>\{o_7\|\|o_8\|\|o_1\|\|o_5\|\|o_4\}$	3 (o)	0	0	64.3%
e_{10}	$\{o_3\|\|o_5\|\|o_1\}>\{o_6\|\|o_8\|\|o_4\|\|o_2\|\|o_7\}$	3 (u)	0	0	53.6%
					$\bar{c}$ = 73.2%

Note: "(o)" stands for *ordered t*/*b*-objects;
"(u)" stands for *unordered t*/*b*-objects.

Table A.4. First incomplete problem (1) and respective descriptive parameters.

Expert	(2) Second incomplete problem Rankings	t	b	k	c
e_1	$o_6>o_3>o_1>o_5>o_4>o_2>o_7$	4 (o)	3 (o)	1 (o_8)	75.0%
e_2	$o_6>o_3>o_5>o_7>o_1>o_4>o_2$	3 (o)	3 (o)	1 (o_8)	75.0%
e_3	$o_3>o_5>o_4>o_8>o_2>o_1>o_7$	3 (o)	3 (o)	1 (o_6)	75.0%
e_4	$\{o_6\|\|o_5\|\|o_2\}>o_3>\{o_8\|\|o_4\|\|o_7\}$	3 (u)	3 (u)	1 (o_1)	53.6%
e_5	$o_3>o_6>o_5>\{o_7\|\|o_1\|\|o_4\|\|o_2\}$	3 (o)	0	1 (o_8)	53.6%
e_6	$o_3>o_5>o_7>\{o_2\|\|o_1\|\|o_8\|\|o_4\}$	3 (o)	0	1 (o_6)	53.6%

e_7	$o_5>o_6>o_3>\{o_4\|\|o_1\|\|o_2\|\|o_7\}$	3 (o)	0	1 (o_8)	53.6%
e_8	$\{o_5\|\|o_6\|\|o_3\}>\{o_1\|\|o_8\|\|o_7\|\|o_4\}$	3 (u)	0	1 (o_2)	42.9%
e_9	$o_6>o_2>o_3>\{o_8\|\|o_1\|\|o_5\|\|o_4\}$	3 (o)	0	1 (o_7)	53.6%
e_{10}	$\{o_3\|\|o_5\|\|o_1\}>\{o_6\|\|o_8\|\|o_4\|\|o_2\}$	3 (u)	0	1 (o_7)	42.9%
					$\overline{c}$ = 57.9%

Note: "(o)" stands for *ordered t*/*b*-objects;
"(u)" stands for *unordered t*/*b*-objects.

Table A.5. Second incomplete problem (2) and respective descriptive parameters.

(3) Third incomplete problem					
Expert	Rankings	*t*	*b*	*k*	*c*
e_1	$o_6>o_1>o_5>o_2>o_7>o_8$	3 (o)	3 (o)	2 (o_4, o_3)	53.6%
e_2	$o_6>o_3>o_7>o_1>o_8>o_4$	3 (o)	3 (o)	2 (o_2, o_5)	53.6%
e_3	$o_6>o_5>o_4>o_8>o_1>o_7$	3 (o)	3 (o)	2 (o_3, o_2)	53.6%
e_4	$\{o_6\|\|o_5\|\|o_3\}>\{o_1\|\|o_8\|\|o_7\}$	3 (u)	3 (u)	2 (o_2, o_4)	32.1%
e_5	$o_3>o_6>o_5>\{o_4\|\|o_2\|\|o_8\}$	3 (o)	0	2 (o_1, o_7)	42.9%
e_6	$o_3>o_6>o_7>\{o_2\|\|o_8\|\|o_4\}$	3 (o)	0	2 (o_5, o_1)	42.9%
e_7	$o_5>o_3>o_1>\{o_8\|\|o_2\|\|o_7\}$	3 (o)	0	2 (o_4, o_6)	42.9%
e_8	$\{o_5\|\|o3\|\|o_2\}>\{o_8\|\|o_7\|\|o_4\}$	3 (u)	0	2 (o_1, o_6)	32.1%
e_9	$o_2>o_3>o_7>\{o_1\|\|o_5\|\|o_4\}$	3 (o)	0	2 (o_8, o_6)	42.9%
e_{10}	$\{o_5\|\|o_1\|\|o_6\}>\{o_8\|\|o_4\|\|o_2\}$	3 (u)	0	2 (o_7, o_3)	32.1%
					$\overline{c}$ = 42.9%

Note: "(o)" stands for *ordered t*/*b*-objects;
"(u)" stands for *unordered t*/*b*-objects.

Table A.6. Third incomplete problem (3) and respective descriptive parameters.

(4) Fourth incomplete problem					
Expert	Rankings	*t*	*b*	*k*	*c*
e_1	$o_6>o_3>o_5>o_4>o_2>o_7$	3 (o)	3 (o)	2 (o_8, o_1)	53.6%
e_2	$o_6>\{o_3\|\|o_5\|\|o_1\|\|o_4\}>o_2$	1	1	2 (o_7, o_8)	32.1%
e_3	$o_6>\{o_3\|\|o_5\|\|o_8\|\|o_2\}>o_7$	1	1	2 (o_1 o_4)	32.1%
e_4	$o_6>\{o_5\|\|o_3\|\|o_1\|\|o_4\}>o_7$	1	1	2 (o_8, o_2)	32.1%
e_5	$o_3>\{o_6\|\|o_5\|\|o_7\|\|o_4\|\|o_2\}$	1	0	2 (o_8, o_1)	17.9%
e_6	$o_3>\{o_5\|\|o_7\|\|o_1\|\|o_8\|\|o_4\}$	1	0	2 (o_2, o_6)	17.9%
e_7	$o_5>\{o_6\|\|o_3\|\|o_4\|\|o_1\|\|o_8\}$	1	0	2 (o_2, o_7)	17.9%
e_8	$o_5>\{o_3\|\|o_1\|\|o_2\|\|o_8\|\|o_7\}$	1	0	2 (o_6, o_4)	17.9%
e_9	$o_6>\{o_3\|\|o_7\|\|o_8\|\|o_1\|\|o_4\}$	1	0	2 (o_2, o_5)	17.9%
e_{10}	$o_3>\{o_5\|\|o_1\|\|o_6\|\|o_8\|\|o_7\}$	1	0	2 (o_4, o_2)	17.9%
					$\overline{c}$ = 25.7%

Note: "(o)" stands for *ordered t*/*b*-objects.

Table A.7. Fourth incomplete problem (4) and respective descriptive parameters.

(5) Fifth incomplete problem					
Expert	Rankings	*t*	*b*	*k*	*c*
e_1	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_2	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_3	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_4	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_5	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_6	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_7	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_8	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_9	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
e_{10}	$\{o_1\|\|o_2\|\|o_3\|\|o_4\|\|o_5\|\|o_6\|\|o_7\|\|o_8\}$	0	0	All	0.0%
					$\overline{c}$ = 0.0%

Table A.8. Fifth incomplete problem (5) and respective descriptive parameters.

Tables A.9 to A.14 contain the rank tables relating to the six problems (0 to 5) shown in Tables A.3 to A.8 respectively.

(0) Complete problem

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	Row totals
e_1	3, 3 → 3.0	6, 6 → 6.0	2, 2 → 2.0	5, 5 → 5.0	4, 4 → 4.0	1, 1 → 1.0	7, 7 → 7.0	8, 8 → 8.0	36
e_2	5, 5 → 5.0	8, 8 → 8.0	2, 2 → 2.0	7, 7 → 7.0	3, 3 → 3.0	1, 1 → 1.0	4, 4 → 4.0	6, 6 → 6.0	36
e_3	7, 7 → 7.0	6, 6 → 6.0	2, 2 → 2.0	4, 4 → 4.0	3, 3 → 3.0	1, 1 → 1.0	8, 8 → 8.0	5, 5 → 5.0	36
e_4	5, 5 → 5.0	3, 3 → 3.0	4, 4 → 4.0	7, 7 → 7.0	2, 2 → 2.0	1, 1 → 1.0	8, 8 → 8.0	6, 6 → 6.0	36
e_5	5, 5 → 5.0	7, 7 → 7.0	1, 1 → 1.0	6, 6 → 6.0	3, 3 → 3.0	2, 2 → 2.0	4, 4 → 4.0	8, 8 → 8.0	36
e_6	6, 6 → 6.0	5, 5 → 5.0	1, 1 → 1.0	8, 8 → 8.0	2, 2 → 2.0	3, 3 → 3.0	4, 4 → 4.0	7, 7 → 7.0	36
e_7	5, 5 → 5.0	7, 7 → 7.0	3, 3 → 3.0	4, 4 → 4.0	1, 1 → 1.0	2, 2 → 2.0	8, 8 → 8.0	6, 6 → 6.0	36
e_8	4, 4 → 4.0	5, 5 → 5.0	3, 3 → 3.0	8, 8 → 8.0	1, 1 → 1.0	2, 2 → 2.0	7, 7 → 7.0	6, 6 → 6.0	36
e_9	6, 6 → 6.0	2, 2 → 2.0	3, 3 → 3.0	8, 8 → 8.0	7, 7 → 7.0	1, 1 → 1.0	4, 4 → 4.0	5, 5 → 5.0	36
e_{10}	3, 3 → 3.0	7, 7 → 7.0	1, 1 → 1.0	6, 6 → 6.0	2, 2 → 2.0	4, 4 → 4.0	8, 8 → 8.0	5, 5 → 5.0	36
Col. totals	49.0	56.0	22.0	63.0	28.0	18.0	62.0	62.0	360

Table A.9. Rank table concerning the complete problem (0) in Table A.3.

(1) First incomplete problem

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	Row totals
e_1	3, 3 → 3.0	6, 6 → 6.0	2, 2 → 2.0	5, 5 → 5.0	4, 4 → 4.0	1, 1 → 1.0	7, 7 → 7.0	8, 8 → 8.0	36
e_2	4, 5 → 4.5	8, 8 → 8.0	2, 2 → 2.0	7, 7 → 7.0	3, 3 → 3.0	1, 1 → 1.0	4, 5 → 4.5	6, 6 → 6.0	36
e_3	7, 7 → 7.0	6, 6 → 6.0	2, 2 → 2.0	4, 5 → 4.5	3, 3 → 3.0	1, 1 → 1.0	8, 8 → 8.0	4, 5 → 4.5	36
e_4	4, 5 → 4.5	1, 3 → 2.0	4, 5 → 4.5	6, 8 → 7.0	1, 3 → 2.0	1, 3 → 2.0	6, 8 → 7.0	6, 8 → 7.0	36
e_5	4, 8 → 6.0	4, 8 → 6.0	1, 1 → 1.0	4, 8 → 6.0	3, 3 → 3.0	2, 2 → 2.0	4, 8 → 6.0	4, 8 → 6.0	36
e_6	4, 8 → 6.0	4, 8 → 6.0	1, 1 → 1.0	4, 8 → 6.0	2, 2 → 2.0	3, 3 → 3.0	4, 8 → 6.0	4, 8 → 6.0	36
e_7	4, 8 → 6.0	4, 8 → 6.0	3, 3 → 3.0	4, 8 → 6.0	1, 1 → 1.0	2, 2 → 2.0	4, 8 → 6.0	4, 8 → 6.0	36
e_8	4, 8 → 6.0	4, 8 → 6.0	1, 3 → 2.0	4, 8 → 6.0	1, 3 → 2.0	1, 3 → 2.0	4, 8 → 6.0	4, 8 → 6.0	36
e_9	4, 8 → 6.0	2, 2 → 2.0	3, 3 → 3.0	4, 8 → 6.0	4, 8 → 6.0	1, 1 → 1.0	4, 8 → 6.0	4, 8 → 6.0	36
e_{10}	1, 3 → 2.0	4, 8 → 6.0	1, 3 → 2.0	4, 8 → 6.0	1, 3 → 2.0	4, 8 → 6.0	4, 8 → 6.0	4, 8 → 6.0	36
Col. totals	51.0	54.0	22.5	59.5	28.0	21.0	62.5	61.5	360

Table A.10. Rank table concerning the first incomplete problem (1) in Table A.4.

(2) Second incomplete problem

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	Row totals
e_1	3, 4 → 3.5	6, 7 → 6.5	2, 3 → 2.5	5, 6 → 5.5	4, 5 → 4.5	1, 2 → 1.5	7, 8 → 7.5	1, 8 → 4.5	36
e_2	5, 6 → 5.5	7, 8 → 7.5	2, 3 → 2.5	6, 7 → 6.5	3, 4 → 3.5	1, 2 → 1.5	4, 5 → 4.5	1, 8 → 4.5	36
e_3	6, 7 → 6.5	5, 6 → 5.5	1, 2 → 1.5	3, 4 → 3.5	2, 3 → 2.5	1, 8 → 4.5	7, 8 → 7.5	4, 5 → 4.5	36
e_4	1, 8 → 4.5	1, 4 → 2.5	4, 5 → 4.5	5, 8 → 6.5	1, 4 → 2.5	1, 4 → 2.5	5, 8 → 6.5	5, 8 → 6.5	36
e_5	4, 8 → 6.0	4, 8 → 6.0	1, 2 → 1.5	4, 8 → 6.0	3, 4 → 3.5	2, 3 → 2.5	4, 8 → 6.0	1, 8 → 4.5	36
e_6	4, 8 → 6.0	4, 8 → 6.0	1, 2 → 1.5	4, 8 → 6.0	2, 3 → 2.5	1, 8 → 4.5	3, 4 → 3.5	4, 8 → 6.0	36
e_7	4, 8 → 6.0	4, 8 → 6.0	3, 4 → 3.5	4, 8 → 6.0	1, 2 → 1.5	2, 3 → 2.5	4, 8 → 6.0	1, 8 → 4.5	36
e_8	4, 8 → 6.0	1, 8 → 4.5	1, 4 → 2.5	4, 8 → 6.0	1, 4 → 2.5	1, 4 → 2.5	4, 8 → 6.0	4, 8 → 6.0	36
e_9	4, 8 → 6.0	2, 3 → 2.5	3, 4 → 3.5	4, 8 → 6.0	4, 8 → 6.0	1, 2 → 1.5	1, 8 → 4.5	4, 8 → 6.0	36
e_{10}	1, 4 → 2.5	4, 8 → 6.0	1, 4 → 2.5	4, 8 → 6.0	1, 4 → 2.5	4, 8 → 6.0	1, 8 → 4.5	4, 8 → 6.0	36
Col. totals	52.5	53.0	26.0	58.0	31.5	29.5	56.5	53.0	360

Table A.11. Rank table concerning the second incomplete problem (2) in Table A.5.

(3) Third incomplete problem

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	Row totals
e_1	2, 4 → 3.0	4, 6 → 5.0	1, 8 → 4.5	1, 8 → 4.5	3, 5 → 4.0	1, 3 → 2.0	5, 7 → 6.0	6, 8 → 7.0	36
e_2	4, 6 → 5.0	1, 8 → 4.5	2, 4 → 3.0	6, 8 → 7.0	1, 8 → 4.5	1, 3 → 2.0	3, 5 → 4.0	5, 7 → 6.0	36
e_3	5, 7 → 6.0	1, 8 → 4.5	1, 8 → 4.5	3, 5 → 4.0	2, 4 → 3.0	1, 3 → 2.0	6, 8 → 7.0	4, 6 → 5.0	36
e_4	4, 8 → 6.0	1, 8 → 4.5	1, 5 → 3.0	1, 8 → 4.5	1, 5 → 3.0	1, 5 → 3.0	4, 8 → 6.0	4, 8 → 6.0	36
e_5	1, 8 → 4.5	4, 8 → 6.0	1, 3 → 2.0	4, 8 → 6.0	3, 5 → 4.0	2, 4 → 3.0	1, 8 → 4.5	4, 8 → 6.0	36
e_6	1, 8 → 4.5	4, 8 → 6.0	1, 3 → 2.0	4, 8 → 6.0	1, 8 → 4.5	2, 4 → 3.0	3, 5 → 4.0	4, 8 → 6.0	36
e_7	3, 5 → 4.0	4, 8 → 6.0	2, 4 → 3.0	1, 8 → 4.5	1, 3 → 2.0	1, 8 → 4.5	4, 8 → 6.0	4, 8 → 6.0	36
e_8	1, 8 → 4.5	1, 5 → 3.0	1, 5 → 3.0	4, 8 → 6.0	1, 5 → 3.0	1, 8 → 4.5	4, 8 → 6.0	4, 8 → 6.0	36
e_9	4, 8 → 6.0	1, 3 → 2.0	2, 4 → 3.0	4, 8 → 6.0	4, 8 → 6.0	1, 8 → 4.5	3, 5 → 4.0	1, 8 → 4.5	36
e_{10}	1, 5 → 3.0	4, 8 → 6.0	1, 8 → 4.5	4, 8 → 6.0	1, 5 → 3.0	1, 5 → 3.0	1, 8 → 4.5	4, 8 → 6.0	36
Col. totals	46.5	47.5	32.5	54.5	37.0	31.5	52.0	58.5	360

Table A.12. Rank table concerning the third incomplete problem (3) in Table A.6.

(4) Fourth incomplete problem

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	Row totals
e_1	1, 8 → 4.5	5, 7 → 6.0	2, 4 → 3.0	4, 6 → 5.0	3, 5 → 4.0	1, 3 → 2.0	6, 8 → 7.0	1, 8 → 4.5	36
e_2	2, 7 → 4.5	6, 8 → 7.0	2, 7 → 4.5	2, 7 → 4.5	2, 7 → 4.5	1, 3 → 2.0	1, 8 → 4.5	1, 8 → 4.5	36
e_3	1, 8 → 4.5	2, 7 → 4.5	2, 7 → 4.5	1, 8 → 4.5	2, 7 → 4.5	1, 3 → 2.0	6, 8 → 7.0	2, 7 → 4.5	36
e_4	2, 7 → 4.5	1, 8 → 4.5	2, 7 → 4.5	2, 7 → 4.5	2, 7 → 4.5	1, 3 → 2.0	6, 8 → 7.0	1, 8 → 4.5	36
e_5	1, 8 → 4.5	2, 8 → 5.0	1, 3 → 2.0	2, 8 → 5.0	2, 8 → 5.0	2, 8 → 5.0	2, 8 → 5.0	1, 8 → 4.5	36
e_6	2, 8 → 5.0	1, 8 → 4.5	1, 3 → 2.0	2, 8 → 5.0	2, 8 → 5.0	1, 8 → 4.5	2, 8 → 5.0	2, 8 → 5.0	36
e_7	2, 8 → 5.0	1, 8 → 4.5	2, 8 → 5.0	2, 8 → 5.0	1, 3 → 2.0	2, 8 → 5.0	1, 8 → 4.5	2, 8 → 5.0	36
e_8	2, 8 → 5.0	2, 8 → 5.0	2, 8 → 5.0	1, 8 → 4.5	1, 3 → 2.0	1, 8 → 4.5	2, 8 → 5.0	2, 8 → 5.0	36
e_9	2, 8 → 5.0	1, 8 → 4.5	2, 8 → 5.0	2, 8 → 5.0	1, 8 → 4.5	1, 3 → 2.0	2, 8 → 5.0	2, 8 → 5.0	36
e_{10}	2, 8 → 5.0	1, 8 → 4.5	1, 3 → 2.0	1, 8 → 4.5	2, 8 → 5.0	2, 8 → 5.0	2, 8 → 5.0	2, 8 → 5.0	36
Col. totals	47.5	50.0	37.5	47.5	41.0	34.0	55.0	47.5	360

Table A.13. Rank table concerning the fourth incomplete problem (4) in Table A.7.

(5) Fifth incomplete problem

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8	Row totals
e_1	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_2	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_3	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_4	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_6	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_7	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_8	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_9	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
e_{10}	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	1, 8 → 4.5	36
Col. totals	45.0	45.0	45.0	45.0	45.0	45.0	45.0	45.0	360

Table A.14. Rank table concerning the fifth incomplete problem (5) in Table A.8.