

Aggregating multiple ordinal rankings in engineering design: the best model according to the Kendall's coefficient of concordance

*Original*

Aggregating multiple ordinal rankings in engineering design: the best model according to the Kendall's coefficient of concordance / Franceschini, Fiorenzo; Maisano, DOMENICO AUGUSTO FRANCESCO. - In: RESEARCH IN ENGINEERING DESIGN. - ISSN 0934-9839. - STAMPA. - 32:1(2021), pp. 91-103. [10.1007/s00163-020-00348-3]

*Availability:*

This version is available at: 11583/2870247 since: 2021-02-09T14:25:10Z

*Publisher:*

Springer

*Published*

DOI:10.1007/s00163-020-00348-3

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

Springer postprint/Author's Accepted Manuscript

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: <http://dx.doi.org/10.1007/s00163-020-00348-3>

(Article begins on next page)



# Aggregating multiple ordinal rankings in engineering design: the best model according to the Kendall's coefficient of concordance

Fiorenzo Franceschini<sup>1</sup> · Domenico Maisano<sup>1</sup>

Received: 1 April 2020 / Revised: 5 October 2020 / Accepted: 7 October 2020 / Published online: 27 October 2020  
© The Author(s) 2020

## Abstract

Aggregating the preferences of a group of experts is a recurring problem in several fields, including engineering design; in a nutshell, each expert formulates an ordinal ranking of a set of alternatives and the resulting rankings should be aggregated into a collective one. Many aggregation models have been proposed in the literature, showing strengths and weaknesses, in line with the implications of Arrow's impossibility theorem. Furthermore, the coherence of the collective ranking with respect to the expert rankings may change depending on: (i) the expert rankings themselves and (ii) the aggregation model adopted. This paper assesses this coherence for a variety of aggregation models, through a recent test based on the Kendall's coefficient of concordance ( $W$ ), and studies the characteristics of those models that are most likely to achieve higher coherence. Interestingly, the so-called Borda count model often provides best coherence, with some exceptions in the case of collective rankings with ties. The description is supported by practical examples.

**Keywords** Engineering design · Expert group · Multiple ordinal rankings · Collective decision-making · Borda count · Kendall's coefficient of concordance ( $W$ ) · Degree of concordance

## 1 Introduction

A problem that is common to a number of fields, including engineering design, is that of aggregating multiple ordinal rankings of a set of alternatives into a collective ranking. This problem may concern the early-design stage, in which  $m$  experts (or decision-making agents:  $D_1$  to  $D_m$ ) formulate their individual rankings of  $n$  design alternatives (or objects:  $O_1$  to  $O_n$ ) (Fu et al. 2010; Frey et al. 2009; Hoyle and Chen 2011; Keeney 2009). In the simplest case, these rankings are complete, i.e.:

- (i) each expert is able to rank all the alternatives of interest, without omitting any of them;

- (ii) each ranking can be decomposed into paired-comparison relationships of strict preference (e.g.,  $O_1 > O_2$  or  $O_1 < O_2$ ) and/or indifference (e.g.,  $O_1 \sim O_2$ ).

The objective of the problem is to aggregate the expert ordinal rankings into a collective one, which is supposed to reflect them as much as possible, even in the presence of diverging preferences (Weingart et al. 2005; See and Lewis 2006). For this reason, the collective ranking is often defined as social, consensus or compromise ranking (Cook 2006; Herrera-Viedma et al. 2014; Franceschini et al. 2015, 2016).

Returning to the context of the early-design stage, design alternatives are often not very well defined and there are doubts about how) to prioritize them (Weingart 2005; Kaldate et al. 2006; McComb et al. 2017. Although there is a substantial agreement on the design criteria, the selection of design alternatives is generally driven by the different personal experience of designers (Dwarakanath and Wallace 1995). Thus arises the need to aggregate preference rankings of design alternatives that reflect the opinions of individual experts, using appropriate aggregation models (Fishburn 1973b; Franssen 2005; Cook 2006; Hazelrigg 1999; Frey et al. 2010; Katsikopoulos 2009; Ladha et al. 2003; Reich 2010; Nurmi 2012).

---

✉ Fiorenzo Franceschini  
fiorenzo.franceschini@polito.it  
Domenico Maisano  
domenico.maisano@polito.it

<sup>1</sup> Department of Management and Production Engineering, Politecnico di Torino, DIGEP, Corso Duca degli Abruzzi 24, 10129 Torino, Italy

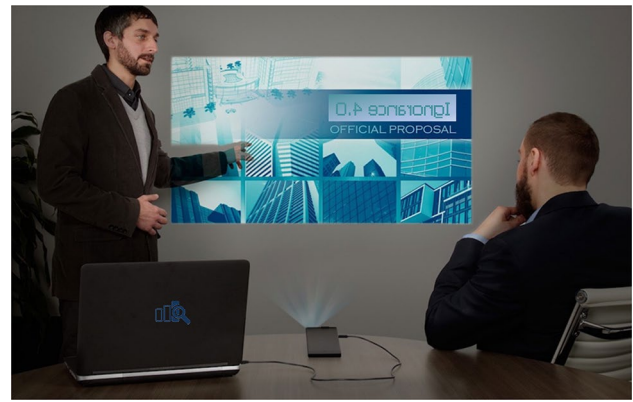
Alongside this, a passionate debate on the effects of the Arrow's impossibility theorem in engineering design is still going on (Arrow 2012; Reich 2010; Hazelrigg 1996, 1999, 2010; Scott and Antonsson 1999; Franssen 2005; Yeo et al. 2004; McComb et al. 2017). In short, this theorem establishes the impossibility of a generic aggregation model to provide a collective ranking that always satisfies several desirable properties, also known as fairness criteria, i.e., *unrestricted domain*, *non-dictatorship*, *independence of irrelevant alternatives* (IIA), *weak monotonicity*, and *Pareto efficiency* (Arrow 2012; Fishburn 1973a; Nisan et al. 2007; Saari 2011; Saari and Sieberg 2004; Franssen 2005; Jacobs et al. 2014).

For a given set of  $m$  expert rankings concerning  $n$  alternatives, different aggregation models may obviously lead to different collective rankings (Saari 2011; McComb et al. 2017). Identifying the model that best reflects the  $m$  rankings is not easy, also because it may change from case to case. Some researchers showed the effectiveness of specific aggregation models, even though they cannot always satisfy all of the Arrow's fairness criteria (Dym, Wood and Scott 2002). Yet, the Arrow's theorem does not close the doors to the possibility of comparing different aggregation models, identifying the best one(s) on the basis of certain tests. For example, several authors attempted to measure the coherence (or consistency) between the expert rankings and the collective one (Chiclana 2002; Franceschini and Maisano 2015, 2017; Franceschini and Garcia-Lapresta 2019). Other authors hypothesized a relationship between the so-called implicit agreement of the expert rankings and the Arrow's fairness (McComb et al. 2017). Moreover, Katsikopoulos (2009) expressed the need for greater clarity in the discussion of engineering design methods to support decision making.

In general, the choice of the best aggregation model may depend on: (1) the specific objective(s) of the expert group and/or (2) the rationale of the test used (Dong et al. 2004; Li et al. 2007; Paulus et al. 2011; Cagan and Vogel 2012; Franceschini et al. 2019; Franceschini and Maisano 2019b).

The aim of this article is to make a comparison between four relatively popular aggregation models—i.e., the so-called Best of the best, Best two, Best three, and Borda count model—trying to answer the research question: “Which is the model producing the collective ranking that best reflects the expert rankings?”. The comparison will be performed by measuring the coherence of the models, through a recent test based on the so-called Kendall's coefficient of concordance ( $W$ ) (Kendall 1962; Legendre 2010). This test quantitatively evaluates the coherence of the collective ranking provided by any aggregation model, for a specific ranking-aggregation problem.

A previous research (Franceschini and Maisano 2019a) illustrated the test in general terms, regardless of the



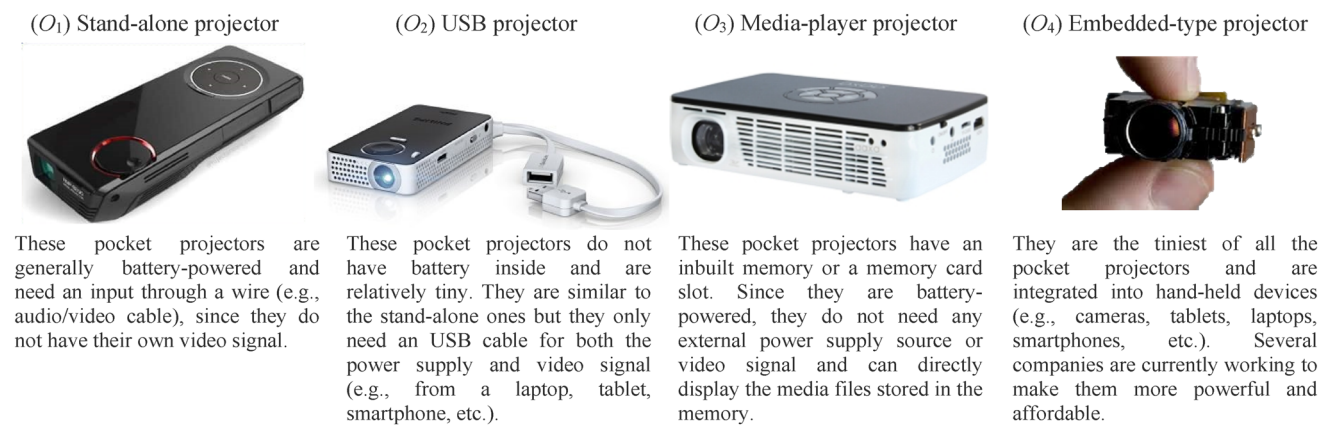
**Fig. 1** Example of pocket projector, i.e., small hardware device designed to project content from a smartphone, camera, tablet, notebook or memory device onto a wall or other flat surface

characteristics of the specific aggregation models. This work significantly extends the previous one, investigating the characteristics of the aggregation models that are most likely to achieve higher coherence, according to the above test. The new investigation generalizes earlier results, including a mathematical optimization of a specific coherence indicator. Thanks to the outcomes of this study, the engineering-design management will have extra support for choosing the most “promising” aggregation models.

The remainder of this article is organized into three sections. Section 2 illustrates a case study that will accompany the description of the proposed methodology. Section 3 is divided into two parts: the first part formalizes the concept of coherence of the collective ranking with respect to the expert rankings, recalling the coherence test proposed in (Franceschini and Maisano 2019a); the second part analyzes the test itself thoroughly, showing its close link with the Borda count model. Section 4 provides a discussion of the practical implications and limitations of this research for the engineering-design field, summarizing original contributions and suggestions for future research. Further details are contained in the Appendix section.

## 2 Case study

This section contains an application example that will be used to illustrate the proposed methodology. An important hi-tech company—which is kept anonymous for reasons of confidentiality—operates predominantly in the sector of video projectors. Recent advances in imaging technology have led the company to increasingly invest in the development of hand-held projectors, also known as a pocket projectors, mobile projectors, pico-projectors or mini beamers (see Fig. 1) (Borisov et al. 2018).



**Fig. 2** Schematic representation and short description of four alternative design concepts of pocket projectors

Four design concepts of pocket projectors ( $O_1$  to  $O_4$ , i.e., objects) have been generated by a team of ten engineering designers (i.e., the experts of the problem:  $D_1$  to  $D_{10}$ ), during the conceptual design phase (see also the description in Fig. 2):

- ( $O_1$ ) stand-alone projector;
- ( $O_2$ ) USB projector;
- ( $O_3$ ) media-player projector;
- ( $O_4$ ) embedded-type projector.

The objective is to evaluate the aforementioned design concepts in terms of user friendliness, i.e., a measure of the ease of use of a pocket projector. Some of the factors that can positively influence this attribute are: (i) quick set-up time, (ii) intuitive controls, and (iii) good user interface.

Given the great difficulty in bringing together all the experts and making them interact to reach shared decisions, management leaned towards a different solution: a collective ranking of the four design concepts can be obtained by merging the individual rankings formulated by the ten engineering designers (Table 1 shows these rankings).

Before focusing on the possible aggregation models, let us take a step back dealing with the evaluation of the experts' degree of concordance (Franceschini and Maisano 2019a). The scientific literature includes an important indicator to evaluate the overall association for more than two rankings, i.e., the so-called Kendall's coefficient of concordance, which is defined as (Kendall and Smith, 1939; Kendall 1962; Fishburn 1973b; Legendre 2005, 2010):

$$W^{(m)} = \frac{12(\sum_{i=1}^n R_i^2) - 3m^2n(n+1)^2}{m^2n(n^2 - 1) - m(\sum_{j=1}^m T_j)}, \quad (1)$$

where:  $R_i = \sum_{j=1}^m r_{ij}$  is the sum of the rank positions for the  $i$ -th object,  $r_{ij}$  being the rank position of the object  $O_i$  according to the  $j$ -th expert;  $n$  is the total

number of objects;  $m$  is the total number of ordinal rankings;  $T_j = \sum_{i=1}^{g_j} (t_i^3 - t_i), \forall j = 1, \dots, m$ , being  $t_i$  the number of objects in the  $i$ -th group of ties (a group is a set of tied objects), and  $g_j$  is the number of groups of ties in the ranking by the  $j$ -th expert. If there are no ties in the  $j$ -th ranking, then  $T_j = 0$ .

Regarding the rank positions ( $r_{ij}$ ) of the tied objects, a convention is adopted whereby they should be the average rank positions that each set of tied objects would occupy if a strict dominance relationship could be expressed (Gibbons and Chakraborti 2010). This convention guarantees that—for a certain  $j$ -th ranking and regardless of the presence of ties—the sum of the objects' rank positions is an invariant equal to:

$$\sum_{i=1}^n r_{ij} = \frac{n \cdot (n + 1)}{2}. \quad (2)$$

In terms of range,  $W^{(m)} \in [0,1]$ .  $W^{(m)} = 0$  indicates the absence of concordance, while  $W^{(m)} = 1$  indicates the complete concordance (or unanimity). The superscript " $(m)$ "

**Table 1** Ordinal rankings of four design concepts (i.e.,  $O_1$  to  $O_4$ ) formulated by ten engineering designers (i.e.,  $D_1$  to  $D_{10}$ )

Engineering designer	Ranking
$D_1$	$O_1 > O_2 > O_3 > O_4$
$D_2$	$O_1 > O_2 > O_3 > O_4$
$D_3$	$O_1 > O_3 > O_4 > O_2$
$D_4$	$O_1 > O_4 > O_3 > O_2$
$D_5$	$O_1 > O_4 > O_3 > O_2$
$D_6$	$O_3 > O_2 > O_4 > O_1$
$D_7$	$O_3 > O_2 > O_4 > O_1$
$D_8$	$O_4 > O_2 > O_3 > O_1$
$D_9$	$O_4 > O_2 > O_3 > O_1$
$D_{10}$	$O_4 > O_2 > O_3 > O_1$

**Table 2** Ranks positions ( $r_{ij}$ ) for the four design concepts ( $O_1, O_2, O_3$  and  $O_4$ ), deduced from the expert rankings in Table 1

Engineering designer	Rank positions ( $r_{ij}$ )			
	$O_1$	$O_2$	$O_3$	$O_4$
$D_1$	1	2	3	4
$D_2$	1	2	3	4
$D_3$	1	4	2	3
$D_4$	1	4	3	2
$D_5$	1	4	3	2
$D_6$	4	2	1	3
$D_7$	4	2	1	3
$D_8$	4	2	3	1
$D_9$	4	2	3	1
$D_{10}$	4	2	3	1
$R_i = \sum_{j=1}^m r_{ij}$	25	26	25	24

was added by the authors to underline that the coefficient of concordance is applied to the  $m$  expert rankings and to distinguish it from another indicator—referred to as  $W^{(m+1)}$ —which will be applied to  $m + 1$  rankings.

Returning to the problem in Table 1, which does not include any ranking with ties (i.e.,  $T_j = 0$ ), the formula in Eq. 1 can be applied, obtaining  $W^{(m)} = 0.004 = 0.4\%$ . This result denotes a relatively low degree of concordance among experts.

Table 2 shows the calculation of the rank positions ( $r_{ij}$ ) of the four objects, for each of the ten expert rankings in Table 1.

Inspired by different design strategies, the team of engineering designers decides to consider four popular aggregation models from the scientific literature (Saari 2011; McComb et al. 2017; Franceschini and Maisano 2019a). A brief description of these models follows:

- (i) *Best of the best* model (BoB or standard plurality vote). For each ranking, the most preferred design concept obtains one point. According to the data in Table 2, the resulting collective ranking is  $O_1 > O_4 \sim O_3 > O_2$  and the “winning” design concept is  $O_1$ . Table 3(i) contains the intermediate calculations.
- For example, this model is used to designate the winner of important competitions, such as the “Red Dot Award”, awarded by eminent design associations/centres to the best design concept of the year (see <https://www.red-dot.org>, last accessed on September 2020).
- (ii) *Best two* model (BTW or vote for two). For each ranking, the two most preferred design concepts obtain one point each. According to the data in Table 1, the resulting collective ranking is  $O_1 > O_4 > O_3 > O_2$

**Table 3** Scores related to the design concepts (i.e.,  $O_1$  to  $O_4$ ) and corresponding collective ranking, produced by each of the four aggregation models (i.e., BoB, BoB, BTW, BTW and BC)

Aggregation model	Scores				Collective ranking
	$O_1$	$O_2$	$O_3$	$O_4$	
(i) BoB	5	0	2	3	$\rightarrow O_1 > O_4 > O_3 > O_2$
(ii) BTW	5	7	3	5	$\rightarrow O_2 > O_1 \sim O_4 > O_3$
(iii) BTH	5	7	10	8	$\rightarrow O_3 > O_4 > O_2 > O_1$
(iv) BC	25	26	25	24	$\rightarrow O_4 > O_1 \sim O_3 > O_2$

and the “winning” design concept is  $O_2$ . Table 3(ii) contains the intermediate calculations.

For example, this model is used for municipal elections of City Commissioner in some major U.S. cities (Boyd and Markman 1983).

- (iii) *Best three* model (BTH or vote for three). For each ranking, the three most preferred design concepts obtain one point each (i.e., this is equivalent to neglecting the worst design concept). The resulting collective ranking is  $O_3 > O_4 > O_2 > O_1$  and the “winning” design concept is  $O_3$ . Table 3(iii) contains the intermediate calculations. Whilst this model is less common than the above models, it is occasionally used for municipal elections in several city councils (Stark 2008).
- (iv) *Borda count* model (BC). For each expert ranking, the first design concept accumulates one point, the second two points, and so on (Borda 1781). According to this model, the cumulative scores of the four design concepts are calculated as:

$$\begin{aligned}
 BC(O_1) &= 1 + 1 + 1 + 1 + 1 + 4 + 4 + 4 + 4 + 4 = 25 \\
 BC(O_2) &= 2 + 2 + 4 + 4 + 4 + 2 + 2 + 2 + 2 + 2 = 26 \\
 BC(O_3) &= 3 + 3 + 2 + 3 + 3 + 1 + 1 + 3 + 3 + 3 = 25 \\
 BC(O_4) &= 4 + 4 + 3 + 2 + 2 + 3 + 3 + 1 + 1 + 1 = 24
 \end{aligned}
 \tag{3}$$

being  $BC(O_1), BC(O_2), BC(O_3)$  and  $BC(O_4)$  the so-called Borda counts related to the four design concepts. Of course, the degree of preference of an  $i$ -th design concept decreases as the corresponding  $BC(O_i)$  increases. In this specific case, the collective ranking is  $O_4 > O_1 \sim O_3 > O_2$  and the most preferred alternative is  $O_4$ .

In addition to being used for engineering design (Dym et al. 2002; McComb et al. 2017), it is also used for: (1) political elections in several countries, (2) internal elections in some professional and technical societies (e.g., board of governors in the International Society for Cryobiology, board of directors in the X.Org Foundation, research area committees in the U.S. Wheat and Barley Scab Initiative, etc.), and (3) a variety of other contexts (e.g., world

champion of “Public Speaking” contest by Toastmasters International, “RoboCup” autonomous robot soccer competition at the University of Bremen in Germany, ranking of NCAA college teams, etc.) (Emerson 2013).

Reflecting different design strategies, the four aggregation models produce four different collective rankings in this case (see overview in Table 3). Even more surprising is that the best pocket projector design concept (i.e. the object at the top of each collective ranking) is different for each of the four aggregation models. Although this plurality of results may at first glance confuse the reader, it is in some measure justified by the low degree of concordance of the expert rankings (i.e.,  $W^{(m)} = 0.004$ , as seen before). Additionally, this plurality of results raises the question: “Which is the model producing the collective ranking that best reflects the expert rankings?”. To answer this question, a test can be used to measure the coherence between (1) the expert rankings and (2) the collective ranking obtained through each model.

### 3 Testing and maximizing the coherence

This section is divided into two parts: the first one recalls the concept of *coherence* and the so-called  $W^{(m+1)}$  test, while the second one analytically studies the maximization of the coherence itself.

#### 3.1 The $W^{(m+1)}$ test

The basic idea of the  $W^{(m+1)}$  test, recently proposed by the authors (Franceschini and Maisano 2019a), is to analyse the level of coherence between the expert rankings and the collective ranking resulting from the application of the ( $k$ -th) aggregation model. The test is based on the construction of an indicator, denominated  $W_k^{(m+1)}$ , which is nothing more than the Kendall’s concordance coefficient (see Eq. 1), applied to the ( $m + 1$ ) rankings consisting of:

- The  $m$  expert rankings, involved in a engineering-design decision problem;
- The collective ranking obtained by applying the ( $k$ -th) aggregation model to the previous  $m$  rankings. The collective ranking is actually treated as an additional ( $m + 1$ )-th ranking.

The formula of the indicator  $W_k^{(m+1)}$  follows:

$$W_k^{(m+1)} = \frac{12 \left[ \sum_{i=1}^n (R_i + r_i)^2 \right] - 3(m + 1)^2 n(n + 1)^2}{(m + 1)^2 n(n^2 - 1) - (m + 1)(\sum_{j=1}^m T_j) - (m + 1)T_{m+1}}, \tag{4}$$

**Table 4**  $W^{(m)}$ ,  $W_k^{(m+1)}$  and  $b_k^{(m)}$  values related to the  $m$  rankings in Table 1, when applying the four different aggregation models ( $k$ : BoB, BTW, BTW, BC)

Aggregation model ( $k$ )	Collective ranking	$W^{(m)}$	$W_k^{(m+1)}$	$b_k^{(m)}$
(i) BoB	$O_1 > O_4 > O_3 > O_2$	0.4%	1.82%	4.55
(ii) BTW	$O_2 > O_1 \sim O_4 > O_3$	idem	0.58%	1.45
(iii) BTH	$O_3 > O_4 > O_2 > O_1$	idem	1.48%	3.70
(iv) BC	$O_4 > O_1 \sim O_3 > O_2$	idem	2.00%	5.00

where  $r_i$  is the rank position of the  $i$ -th object in the collective ranking; obviously  $r_i \in [1, n]$ . In case of tied objects, the same convention described in Sect. 2 is adopted.

Going back to the case study, the indicator  $W_k^{(m+1)}$  can be determined by applying the formula in Eq. (4) to the ten rankings in Table 1 plus the collective ranking resulting from the application of each aggregation model. Table 4 reports the resulting  $W_k^{(m+1)}$  values; subscript “ $k$ ” denotes a generic aggregation model,  $k$ : BoB, BTW, BTH, BC. For this specific problem, the BC model is the one with the highest coherence ( $W_{BC}^{(m+1)} \approx 2.00\%$ ).

In this specific case, the condition  $W_k^{(m+1)} \geq W^{(m)}$  holds for each  $k$ -th aggregation model, depicting a certain coherence (or positive coherence) between the corresponding collective ranking and the  $m$  rankings. An opposite result (i.e.,  $W_k^{(m+1)} < W^{(m)}$ ) would depict incoherence (or negative coherence). Even though the latter situation is in some ways paradoxical, it can occur when a collective ranking is somehow conflicting with the  $m$  rankings (Franceschini and Maisano 2019a).

To quantitatively measure the degree of coherence of an aggregation model, the following synthetic indicator can be used (Franceschini and Maisano 2019a):

$$b_k^{(m)} = \frac{W_k^{(m+1)}}{W^{(m)}}. \tag{5}$$

For a given set of alternative aggregation models, the most coherent can be considered the one that maximizes  $b_k^{(m)}$ ; in formal terms, the model for which:

$$b_*^{(m)} = \max_k [b_k^{(m)}]. \tag{6}$$

The last column of Table 4 reports the  $b_k^{(m)}$  values related to the four aggregation models of interest. It is worth noticing that this indicator allows a quick and practical quantitative comparison.

#### 3.2 Maximization of $b_k^{(m)}$

Let us now focus on the synthetic indicator  $b_k^{(m)}$ . Replacing Eqs. (1) and (4) into Eq. (5),  $b_k^{(m)}$  can be expressed as:

$$b_k^{(m)} = \frac{W_k^{(m+1)}}{W^{(m)}} = \frac{12 \left[ \sum_{i=1}^n (R_i + r_i)^2 \right] - 3(m+1)^2 n(n+1)^2}{(m+1)^2 n(n^2-1) - (m+1) \left( \sum_{j=1}^m T_j \right) - (m+1) T_{m+1}} \quad (7)$$

$$\frac{12 \left( \sum_{i=1}^n R_i^2 \right) - 3m^2 n(n+1)^2}{m^2 n(n^2-1) - m \left( \sum_{j=1}^m T_j \right)}$$

The previous expression is deliberately general, as it contemplates the possibility of:

- ties—i.e., relationships of indifference (“~”) and not only of strict preference (“>” or “<”)—among the objects within the  $m$ -expert rankings;
- ties among the objects within the collective ranking, or  $(m + 1)$ -th ranking.

By grouping some terms, Eq. (7) can be reformulated in a more compact form as follows:

$$b_k^{(m)} = \frac{N_1 + 24 \left[ \sum_{i=1}^n (R_i r_i) \right] + 12 \left( \sum_{i=1}^n r_i^2 \right)}{D_1 - W^{(m)}(m + 1) T_{m+1}} \quad (8)$$

It can be seen that the indicator  $b_k^{(m)}$  includes four types of contributions:

- $N_1 = 12 \left( \sum_{i=1}^n R_i^2 \right) - 3(m + 1)^2 n(n + 1)^2$  and  $D_1 = W^{(m)} \left\{ (m + 1)^2 n(n^2 - 1) - (m + 1) \left( \sum_{j=1}^m T_j \right) \right\}$ , concerning the  $m$  rankings (and therefore the  $r_{ij}$  and  $R_i$  values, also known as experts’ preference profile) and the parameters related to the “size” of the problem (i.e.,  $n$  and  $m$ );
- $24 \left[ \sum_{i=1}^n (R_i r_i) \right]$ , concerning a mixture of the experts’ preference profile (through the  $R_i$  values) and the ranks of the collective ranking (through the  $r_i$  values);
- $12 \left( \sum_{i=1}^n r_i^2 \right)$ , concerning the  $r_i$  values of the collective ranking;
- $W^{(m)}(m + 1) T_{m+1}$ , concerning a mixture of the experts’ preference profile (through  $W^{(m)}(m + 1)$ ) and the  $T_{m+1}$  value related to the collective ranking.

Note that the first contribution (i) is not related to the results of the collective ranking. Instead, the remaining three contributions—(ii), (iii) and (iv)—are all related to the collective ranking. In line with the research question behind this study, let us try to identify the aggregation model that produces the most coherent collective ranking, through the maximization of  $b_k^{(m)}$ , operating on the three terms concerning the aggregation model: (ii)  $\sum_{i=1}^n (R_i r_i)$ , (iii)  $\sum_{i=1}^n r_i^2$  and (iv)  $T_{m+1}$ .

Additionally, note that the analytic maximization of Eq. (8) as a function of the  $r_i$  values is relatively complex for two reasons: (i) the  $r_i$  values are variables defined on a discrete domain; (ii) the  $r_i$  values are explicit in some terms (i.e.,  $24 \left[ \sum_{i=1}^n (R_i r_i) \right]$  and  $12 \left( \sum_{i=1}^n r_i^2 \right)$ ) and implicit in others

(i.e.,  $W^{(m)}(m + 1) T_{m+1}$ , where possible ties in the  $r_i$  collective ranks affect the  $T_{m+1}$  term).

In the following subsections, the major terms of Eq. (8) will be analysed separately, although they are closely related. A more rigorous, though laborious, alternative could be performing numerical maximization through Monte Carlo simulations.

### 3.2.1 Analysis of $\sum_{i=1}^n (R_i r_i)$

This term can be interpreted as a scalar product between two vectors:  $\mathbf{R} = (R_1, \dots, R_n)$  and  $\mathbf{r} = (r_1, \dots, r_n)$ . In general, the scalar product of two vectors ( $\mathbf{r}$  and  $\mathbf{R}$ ) with predetermined modules is maximized if these vectors are completely aligned, i.e., when direct proportionality between the relevant components occurs:  $r_i \propto R_i$ . With reference to the problem of interest, this perfect alignment can hardly be achieved in practice, due to the fact that the  $r_i$  components are rank positions  $\in [1, n]$ , with constant sum equal to  $n \cdot (n + 1) / 2$ . Compatibly with the previous constraint, it can be demonstrated that the Borda model provides a collective ranking that maximises the term  $\sum_{i=1}^n (R_i r_i)$  (see the proof in Appendix A.1).

### 3.2.2 Analysis of $\sum_{i=1}^n r_i^2 = |\mathbf{r}|^2$

We note that this term corresponds to the squared module of the collective-rank vector  $\mathbf{r} = (r_1, \dots, r_n)$ . It can be demonstrated that  $\mathbf{r}$  has the maximum-possible module when its components are a permutation of natural numbers included between 1 and  $n$ , in the absence of ties (see the proof in Appendix A.2). Precisely, the maximum-possible value of the term of interest is  $n \cdot (n + 1) \cdot (2 \cdot n + 1) / 6$  (Gibbons and Chakraborti 2010). If there is a tie, however, this term tends to decrease. The most disadvantageous case would be the one with an overall tie of all the alternatives (i.e.,  $r_1 = r_2 = \dots = r_n = \frac{n+1}{2}$ ), with a consequent value of the term of  $\left( \frac{n+1}{2} \right)^2 \cdot n$ ; therefore:

$$\sum_{i=1}^n r_i^2 \in \left[ \left( \frac{n + 1}{2} \right)^2 \cdot n, \frac{n \cdot (n + 1) \cdot (2 \cdot n + 1)}{6} \right]. \quad (9)$$

### 3.2.3 Analysis of $T_{m+1}$

This term is maximized in the case of an all-all tie of all the alternatives in the collective ranking, i.e.,  $r_1 = r_2 = \dots = r_n$ , with the value of  $n^3 - n$  (see Sect. 2). In the case of no ties, it is obviously equal to zero. Therefore, the range of  $T_{m+1}$  is:

$$T_{m+1} \in [0, n^3 - n] \quad (10)$$

### 3.3 Close link between $b_k^{(m)}$ and the Borda count model

In light of the previous analyses, it is possible to outline two potential situations, as described in the following subsections.

#### 3.3.1 Absence of ties in the collective ranking

In the absence of ties in the collective ranking, the two terms  $\sum_{i=1}^n r_i^2$  and  $T_{m+1}$  become constants, i.e., respectively:  $\sum_{i=1}^n r_i^2 = \frac{n \cdot (n+1) \cdot (2 \cdot n+1)}{6}$  and  $T_{m+1} = 0$ . It can be proven that the BC model is the one maximizing the term  $\sum_{i=1}^n (R_i r_i)$  (see demonstration in Appendix A.1); given the constancy of the aforementioned two other terms, the BC model will also maximize  $b_k^{(m)}$  in this situation.

To better focalize this result, it may be appropriate to consider the general expression of  $W^{(m)}$  (in Eq. 1) more closely. It can be observed that the  $R_i$  values (e.g., those reported at the bottom of Table 2) are the same as the Borda scores assigned to the individual objects of interest (i.e.,  $BC(O_i)$ ,  $\forall i = 1, \dots, n$ , as exemplified in Eq. 3) (Cook and Seiford 1982). In this situation, the expression of  $b_k^{(m)}$  in Eq. (7) can be reformulated as:

$$b_k^{(m)} = \frac{W_k^{(m+1)}}{W^{(m)}} = \frac{12 \left\{ \left[ \sum_{i=1}^n BC(O_i)^2 \right] + 2 \cdot \left\{ \sum_{i=1}^n [BC(O_i) \cdot r_i] + \frac{n \cdot (n+1) \cdot (2 \cdot n+1)}{6} \right\} - 3 \cdot (m+1)^2 \cdot n \cdot (n+1)^2 \right.}{(m+1)^2 n(n^2-1) - (m+1) \left( \sum_{j=1}^m T_j \right)} \cdot \frac{12 \cdot \left[ \sum_{i=1}^n BC(O_i)^2 \right] - 3m^2 n(n+1)^2}{m^2 n(n^2-1) - m \left( \sum_{j=1}^m T_j \right)} \tag{11}$$

Pooling the terms that do not depend on the collective ranking (i.e., all except for the  $r_i$  terms), the following compact expression can be obtained:

$$b_k^{(m)} = a + b \cdot \sum_{i=1}^n [BC(O_i) \cdot r_i], \tag{12}$$

being  $a$  and  $b$  two terms that—for a given problem—can be treated as constants, as they depend exclusively on  $n$ ,  $m$ ,  $R_i = BC(O_i)$  ( $\forall i = 1, \dots, n$ ), and  $T_j$  ( $\forall j = 1, \dots, m$ ). Equation (12) highlights the close link between the collective ranks ( $r_i$ ) and the Borda scores related to the  $m$ -expert rankings: in the absence of ties in the collective ranking, the BC model maximizes the term  $\sum_{i=1}^n [BC(O_i) \cdot r_i]$ , and therefore also  $b_k^{(m)}$ , determining the maximum alignment (or projection) of the two vectors  $\mathbf{r} = (r_1, \dots, r_n)$  and  $\mathbf{BC} = [BC(O_1), BC(O_2), \dots, BC(O_n)]$  (cf. Sect. 3.2.1 and Appendix A.1).

#### 3.3.2 Presence of ties in the collective ranking

The presence of ties in the collective ranking can affect the maximisation of  $b_k^{(m)}$  in a somewhat unpredictable way: although it contributes to reduce the term  $\sum_{i=1}^n r_i^2$ , it also contributes to increase the term  $(m+1)T_{m+1}$  (cf. Eq. 8). Thus, the overall effect on  $b_k^{(m)}$  is not simply predictable and should be considered on a case-by-case basis; this also emerges from the additional asymptotic analysis of  $b_k^{(m)}$ , contained in Appendix A.4.

Reversing the perspective, in the presence of ties, the maximization of  $\sum_{i=1}^n [BC(O_i) \cdot r_i]$  by the BC model does not guarantee the maximization of  $b_k^{(m)}$  (see the example in Appendix A.3).

## 4 Discussion

Leaving the mathematical issues, this section focuses on (i) practical implications and limitations of this research for the engineering-design field, and (ii) original contributions and ideas for future research. These topics are covered in the following two sub-sections respectively.

### 4.1 Implications and limitations for engineering design

In early-design stages, initial decisions often should be made when information is incomplete and many goals are contradictory, leading to situations of conflict between (co-)designers. Managing the conflict that emerges from multi-design interaction is therefore a critical element of collaborative design (Grebici et al. 2006). According to some authors, conflict itself is the process through which ideas are validated and developed: “the engine of design” (Brown 2013).

Considering the problem of interest, the engineering-design conflict finds its shape in the (discordant) object rankings, which are formulated by the individual designers; for example, this conflict is quite evident from the  $m$  rankings in the case study (see Table 1). The collective ranking represents a way to solve this conflict and the aggregation

**Table 5** Spearman’s  $\rho$  correlation coefficients with relevant  $p$ -values for the BC-model collective ranking (in Table 3) and each of the ten expert rankings in Table 1

Expert	$e_1$	$e_2$	$e_3$	$e_4$	$e_5$	$e_6$	$e_7$	$e_8$	$e_9$	$e_{10}$
Spearman’s $\rho$	− 0.632	− 0.632	0.316	0.632	0.632	− 0.316	− 0.316	0.316	0.316	0.316
$p$ -value	0.368	0.368	0.684	0.368	0.368	0.684	0.684	0.684	0.684	0.684

model therefore represents a sort of conflict-management tool. However, the plurality of aggregation models makes their selection non-trivial: any aggregation model is by definition imperfect and may provide more or less sound results, depending on the specific ranking-aggregation problem (Arrow 2012). In this research, the coherence between the collective ranking and the corresponding  $m$  rankings was considered as a selection criterion; in fact, the indicator  $b_k^{(m)}$  allows to identify the most coherent aggregation model(s) in different practical situations.

In cases where a certain conflict between collective ranking and expert rankings is observed, decision makers can deepen the analysis, identifying those expert rankings that represent the main sources of incoherence. A possible in-depth analysis could be based on the calculation of the Spearman’s rank correlation coefficient<sup>1</sup> ( $\rho$ ) between the collective ranking and each of the expert rankings. Intuitively, the Spearman correlation between the rankings will be high (i.e., tending towards + 1) when objects have a similar rank between the two rankings, and low (i.e., tending towards − 1) when objects have a dissimilar (or even opposed) rank between the two rankings. Of course, the rankings that will produce the highest incoherence are those with negative  $\rho$  values.

For the purpose of example, let us return to the case study; Table 5 reports the  $\rho$  values between the collective ranking related to the application of the BC model (in Table 3) and the corresponding expert rankings (in Table 1). The expert rankings most in contrast with the collective ranking are those by experts  $e_1$  and  $e_2$  (both with  $\rho$  values of − 0.632), followed by

those by experts  $e_6$  and  $e_7$  (both with  $\rho$  values of − 0.316). It could be interesting for the engineering-design management to identify the reasons for the misalignment of these experts (e.g., different view or poor understanding of some design concepts, errors in the ranking formulation, etc.).

Among the jungle of possible aggregation models existing in the scientific literature, this researcher has considered exclusively aggregation models characterised by simplicity and easy understanding. It is well known that simple and intuitive models are more easily “digested” and implemented by management than obscure and complicated ones (Franceschini et al. 2019). To quote a phrase by Leonardo da Vinci, “Simplicity is the ultimate sophistication”. In line with that, our analysis was limited to four simple, intuitive and popular aggregation models, showing that the traditional BC generally provides very coherent results. The authors believe that this is a relevant indication for the engineering-design management when selecting the most “promising” aggregation models for a certain ranking-aggregation problem.

The proposed study has several limitations, summarised in the following points:

- Since the present analysis makes extensive use of the  $W^{(m+1)}$  test, it “inherits” the limitations associated with it, i.e., (Franceschini and Maisano 2019a): (i) the test does not consider the (possible) uncertainty in expert rankings, and (ii) the test allows only an ex post (i.e., case-by-case) analysis of the impact of aggregation models.
- The study revealed that the BC model often provides the best coherence. Over and above the merits of the BC model, this result is also due to the structural characteristics of the  $W^{(m+1)}$  test; in fact, being based on the Kendall’s coefficient of concordance, this test is somehow related to the BC model (cf. Section 3.3) (Cook and Seiford 1978, 1982). This aspect in some ways limits the proposed coherence analysis: measuring coherence through another indicator would not necessarily lead to the same results. As an example, one could use Cronbach’s alpha ( $\alpha_{Ck}$ ), applied respectively to the ranks related to the expert rankings and the collective ranking. A new indicator, similar to  $b_k^{(m)}$ , could then be defined as:

$$c_k^{(m)} = \frac{\alpha_{Ck}^{(m+1)}}{\alpha_C^{(m)}}, \tag{13}$$

<sup>1</sup> This coefficient can be seen as a special case of the Pearson correlation coefficient of two sets of variables (e.g.,  $X$  and  $Y$ ), where the values of the variables are replaced with rank values (e.g.,  $r_X$  and  $r_Y$ ) before calculating the coefficient itself (Spearman 1904; Ross 2009; Myers et al. 2010):

$$\rho = \frac{\sum_{i=1}^n [r_X(O_i) - \bar{r}_X] \cdot [r_Y(O_i) - \bar{r}_Y]}{\sqrt{\sum_{i=1}^n [r_X(O_i) - \bar{r}_X]^2} \cdot \sqrt{\sum_{i=1}^n [r_Y(O_i) - \bar{r}_Y]^2}}$$

where:

$X$  and  $Y$  are subscripts referred to two generic experts ( $e_X$  and  $e_Y$ ), who have formulated their respective rankings of  $n$  objects ( $O_1, O_2, \dots, O_n$ );

$r_X(O_i)$  and  $r_Y(O_i)$  are the rank values of the  $i$ -th object, formulated by  $e_X$  and  $e_Y$  respectively;

$\bar{r}_X = \frac{\sum_{i=1}^n r_X(O_i)}{n}$  and  $\bar{r}_Y = \frac{\sum_{i=1}^n r_Y(O_i)}{n}$  are the mean rank values of the objects, considering the rankings<sup>n</sup> formulated by  $e_X$  and  $e_Y$ .

being  $\alpha_C^{(m)}$  and  $\alpha_{Ck}^{(m+1)}$  respectively the Cronbach’s alpha related to the  $m$ -expert rankings and the  $(m + 1)$  rankings, when adding the collective ranking obtained through the  $(k$ -th) aggregation model (cf. Sect. 3.1).

However, the choice fell on  $W$  for many reasons: (i) it is specific for judgments expressed in the form of rankings and not in other forms, such as on cardinal scales (Hammond et al. 2015); (ii) the  $W$ -distributional properties are well known (Kendall 1962; Gibbons and Chakraborti 2010); (iii) it is intuitive and relatively easy to implement (Franceschini and Maisano 2019a).

- The proposed analysis considers only complete expert rankings, where all objects are ranked through strict preference and/or indifference relationships only. Nevertheless, some practical contexts may make it difficult to formulate complete rankings, e.g., problems with many alternatives, where experts can face practical impediment or do not have the concentration to formulate complete rankings. In such cases, experts may prefer to formulate incomplete rankings, which include the most/least relevant objects only and/or deliberately exclude some other objects (Franceschini and Maisano 2019b, 2020).

#### 4.2 Original contributions and ideas for future research

This paper analysed the coherence of alternative aggregation models, trying to answer the research question: “Which is the model producing the collective ranking that best reflects the expert rankings?”. The coherence between the  $m$ -expert rankings and the collective ranking (which is obtained through a certain aggregation model) was assessed using the  $W^{(m+1)}$  test and the corresponding synthetic indicator  $b_k^{(m)}$  (Franceschini and Maisano 2019a).

It was found that the BC model offers, with some exceptions, the best coherence. Precisely, when no ties appear among the objects of the collective ranking, it was analytically shown that the BC model maximizes both the indicators  $W^{(m+1)}$  and  $b_k^{(m)}$ . Admitting ties, instead, the BC model’s collective ranking is not necessarily the best one, although it is generally close to it (cf. Appendix A.3).

The above result confirms the versatility and practicality of the BC model, which—in spite of some inevitable imperfections (Dym et al. 2002; Arrow 2012)—remains intuitive, easy to implement, computationally light and coherent as results.

Regarding the future, we plan to explore in greater depth the gap between (i) the collective ranking resulting from the BC model and (ii) the one maximizing coherence, based on a large number of tests and experimental simulations.

**Acknowledgements** 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—ITALY”.

**Funding** Open access funding provided by Politecnico di Torino within the CRUI-CARE Agreement.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### Appendix

#### A.1. Rearrangement inequality

The problem of maximizing the term  $\sum_{i=1}^n (R_i r_i)$  has already been addressed in Mathematics in the form of the so-called Rearrangement Inequality (Hardy et al. 1952).

*Statement.* For every choice of real numbers  $R_1 \leq \dots \leq R_n$  and  $r_1 \leq \dots \leq r_n$ , and every permutation  $r_{p(1)} \leq \dots \leq r_{p(n)}$  of  $r_1, \dots, r_n$ , the following inequality applies:

$$R_n \cdot r_1 + \dots + R_1 \cdot r_n \leq R_{p(1)} \cdot r_1 + \dots + R_{p(n)} \cdot r_n \leq R_1 \cdot r_1 + \dots + R_n \cdot r_n. \tag{14}$$

If the numbers are different, meaning that  $R_1 < \dots < R_n$  and  $r_1 < \dots < r_n$ , then the lower bound is attained only for the permutation which reverses the order, i.e.,  $p(i) = n - i + 1$  for all  $i = 1, \dots, n$ , and the upper bound is attained only for the identity, i.e.,  $p(i) = i$  for all  $i = 1, \dots, n$ .

*Proof (by induction).* Observe first that the condition

$$R_1 > R_2 \text{ and } r_1 > r_2 \tag{15}$$

implies

$$(R_1 - R_2) \cdot (r_1 - r_2) > 0 \text{ or } R_1 \cdot r_1 + R_2 \cdot r_2 > R_1 \cdot r_2 + R_2 \cdot r_1 \tag{16}$$

hence the result is true if  $n = 2$ . Assuming that it is true at the  $(n - 1)$ -th rank position, and let

$$R_1 > \dots > R_n \text{ and } r_1 > \dots > r_n \tag{17}$$

choose a permutation  $p$  for which the arrangement gives rise a maximal result. If  $p(n)$  were different from  $n$ , say  $p(n) = k$ , there would exist  $j < n$  such that  $p(j) = n$ . But  $r_j > r_n$  and  $R_k > R_n$  and, hence  $R_n \cdot r_n + R_j \cdot r_k > R_n \cdot r_k + R_j \cdot r_n$ , by what has just been proved. Consequently, it would follow that the permutation  $q$  coinciding with  $p$ , except at  $j$  and  $n$ , where  $q(j) = k$  and  $q(n) = n$ , gives rise a better result. This

contradicts the choice of  $p$ . Hence  $p(n) = n$ , and from the induction hypothesis,  $p(i) = i \forall i < n$  (Hardy et al. 1952).

The same proof holds if one replaces strict inequalities (“ $<$ ”) by non-strict ones (“ $\leq$ ”) (Hardy et al. 1952).

Since, for a certain sequence of  $R_i$  values, the BC model provides a sequence of  $r_i$  values that—in addition to complying with the constraint that the sum is constant and equal to  $n \cdot (n + 1)/2$ —also maintains the order, it can be inferred that it maximizes the term of interest. The same consideration also applies if the BC model results in some ties between the objects in the collective ranking.

### A.2. Ties in the BC collective ranking

The BC model determines a collective ranking (i.e., a set of  $r_i$  values), which is similarly sorted with respect to the ranking of the  $R_i$  values (i.e., if  $R_k$  is greater than or equal to exactly  $R_j$ , then  $r_k$  will also be greater than or equal to exactly  $r_j$ ).

Considering a ranking of  $R_i$  values of  $n$  objects, we assume that two of these values, respectively in positions  $j$  and  $j + 1$ , coincide. The BC model (cf. Sect. 2) will then produce a corresponding collective ranking, containing a tie between two corresponding objects in positions  $j$  and  $j + 1$ .

In this case, the sum of the component-by-component products between the two rankings is:

$$\sum_{i=1}^n (R_i r_i) = \left[ \sum_{i \in (1, \dots, n) \setminus \{j, j+1\}} (R_i r_i) \right] + R_j r_j + R_{j+1} r_{j+1}. \tag{18}$$

Focusing on the last two terms on the second member of Eq. 18 and imposing (1)  $R_j = R_{j+1} = R$ , which reflects the initial hypothesis of a tie between the two  $R_i$  values, and (2)  $r_j = r_{j+1} = [j + (j + 1)]/2 = j + 1/2$ , which reflects the average rank of the tied objects in the collective ranking, Eq. (18) results in:

$$\sum_{i=1}^n (R_i r_i) = \left[ \sum_{i \in (1, \dots, n) \setminus \{j, j+1\}} (R_i r_i) \right] + 2 \left[ R \left( j + \frac{1}{2} \right) \right]. \tag{19}$$

Additionally, it can be demonstrated that the sum of the squares of the  $r_j$  values is:

$$\begin{aligned} \sum_{i=1}^n r_i^2 &= \left[ \sum_{i \in (1, \dots, n) \setminus \{j, j+1\}} r_i^2 \right] + 2 \left( j + \frac{1}{2} \right)^2 \\ &= \left[ \sum_{i \in (1, \dots, n) \setminus \{j, j+1\}} r_i^2 \right] + 2j^2 + 2j + \frac{1}{2}. \end{aligned} \tag{20}$$

If we eliminate the tie between  $r_j = j + 1/2$  and  $r_{j+1} = j + 1/2$ , replacing the respective values with other values compatible with the condition  $\sum_{i=1}^n r_i = \frac{n \cdot (n+1)}{2}$ , we will obtain  $r_j = j$  and  $r_{j+1} = j + 1$  (or, vice versa:  $r_j = j + 1$  and  $r_{j+1} = j$ ). In this

**Table 6** Rankings of five object (i.e.,  $O_1$  to  $O_5$ ), formulated by five experts (i.e.,  $D_1$  to  $D_5$ )

Engineering designer	Ranking
$D_1$	$O_4 > O_1 > O_2 > O_5 > O_3$
$D_2$	$O_1 > O_4 > O_2 > O_3 > O_5$
$D_3$	$O_4 > O_1 > O_5 > O_2 > O_3$
$D_4$	$O_1 > O_4 > O_3 > O_2 > O_5$
$D_5$	$O_4 > O_2 > O_1 > O_5 > O_3$

other case, the sum of the component-by-component products between the  $R_i$  and  $r_i$  rankings is:

$$\sum_{i=1}^n (R_i r_i) = \left[ \sum_{i \in (1, \dots, n) \setminus \{j, j+1\}} (R_i r_i) \right] + R_j + R_{j+1} (j + 1), \tag{21}$$

which is equivalent to Eq. 19. Consequently, it can be said that the new  $r_i$  ranking, without ties, maximizes the term  $\sum_{i=1}^n (R_i r_i)$  as the one resulting from the BC model.

Additionally, in this other case the sum of the squares of the  $r_j$  values is:

$$\begin{aligned} \sum_{i=1}^n r_i^2 &= \left[ \sum_{i \in (1, \dots, n) \setminus \{j, j+1\}} r_i^2 \right] + j^2 + (j + 1)^2 \\ &= \left[ \sum_{i \in (1, \dots, n) \setminus \{j, j+1\}} r_i^2 \right] + 2j^2 + 2j + 1, \end{aligned} \tag{22}$$

which is greater than that in Eq. 20, due to the last term: i.e., 1 instead of  $1/2$ . It can therefore be deduced that the term  $\sum_{i=1}^n r_i^2$  is maximized in the case of absence of ties in the collective ranking.

Obviously, the above considerations can be extended to rankings with ties between more than two objects and/or rankings with different groups of ties.

### A.3. Additional example

Table 6 exemplifies a fictitious problem consisting of  $m = 5$  expert rankings, concerning  $n = 5$  alternatives ( $O_1$  to  $O_5$ ); for the sake of simplicity, the above rankings do not include ties.

Table 7 reports the collective rankings obtained through the four different aggregation models in Sect. 2 (i.e., BoB, BTW, BTH and BC) and several related indicators:

- $W^{(m)} = 74.4\%$ , which depicts a relatively high degree of concordance among the expert rankings;
- The  $W_k^{(m+1)}$  values, specifying the contributions of the three terms discussed in Sect. 3.2:  $\sum_{i=1}^n (R_i r_i)$ ,  $\sum_{i=1}^n r_i^2$  and  $T_{m+1}$ ;
- The  $b_k^{(m)}$  values.

**Table 7**  $W_k^{(m)}$ ,  $W_k^{(m+1)}$  and  $b_k^{(m)}$  values related to the  $m$  rankings in Table 6, when applying four different aggregation models ( $k$ : BoB, BTW, BTH, BC)

Aggregation model	Collective ranking	$W^{(m)}$	$W_k^{(m+1)}$	$\sum_{i=1}^n (R_i r_i)$	$\sum_{i=1}^n r_i^2$	$T_{m+1}$	$b_k^{(m)}$
(i) Bob	$O_4 > O_1 > (O_2 \sim O_3 \sim O_5)$	74.40%	76.44%	261.0	53.0	24	1.0274
(ii) BTW	$O_4 > O_1 > O_2 > (O_3 \sim O_5)$	Idem	78.01%	266.5	54.5	6	1.0485
(iii) BTH	$(O_1 \sim O_4) > O_2 > (O_3 \sim O_5)$	Idem	77.97%	265.5	54.0	12	1.0479
(iv) BC	$O_4 > O_1 > O_2 > O_5 > O_3$	Idem	77.78%	267.0	55.0	0	1.0454

The presence of ties for some collective rankings does not guarantee the achievement of the maximum value for the indicator  $b_k^{(m)}$  through the BC model (Sect. 3.3.2).

It can be seen that the  $W_k^{(m+1)}$  values are very close to each other, as are the respective  $b_k^{(m)}$  values. Consistently with what illustrated in Sects. A.1 and A.2, the BC model—which results in a collective ranking without ties (see Table 7 (iv)) maximizes both the first two terms,  $\sum_{i=1}^n (R_i r_i)$  and  $\sum_{i=1}^n r_i^2$ , in the numerator of Eq. 8, but minimizes the third term, in the denominator of Eq. 8.

As mentioned, the aggregation model that maximizes  $W_k^{(m+1)}$  (and therefore also  $b_k^{(m)}$ ) is not the BC model, as in the case study in Sect. 2, but the BTW model ( $b_{BoB}^{(m)} = 1.0485$ ), followed by the BTH model ( $b_{BoB}^{(m)} = 1.0479$ ) as it appears in the last column of Table 7. For this specific problem, the aggregation model obtained through the BC is only the third one in terms of coherence. This apparently negative result is at least partially mitigated, considering that:

- the  $b_k^{(m)}$  values of the four models are extremely close to each other, in particular those of the first three ones (which differ by a few thousandths of a unit), reflecting collective rankings that are very close to each other (see the second column of Table 7).
- The term  $m$ , which is one of the two “size” terms of the specific problem (together with  $n$ ), is relatively low in this case. It can be proven that the gap between  $b_{BC}^{(m)}$  and  $b_*^{(m)} = b_{BoB}^{(m)}$  would tend to narrow as  $m$  increases.

#### A.4. Asymptotic analysis of $b_k^{(m)}$

This section quantitatively analyses the contributions of the three terms of  $b_k^{(m)}$ , which depend on the aggregation model (cf. Sect. 3): (i)  $\sum_{i=1}^n (R_i r_i)$ , (ii)  $\sum_{i=1}^n r_i^2$ , and (iii)  $T_{m+1}$ . Precisely, an asymptotic analysis of  $b_k^{(m)}$  is performed as the  $n$  and  $m$  parameters grow. Although we are aware that these parameters will never grow too much in realistic decision-making problems (e.g., reaching at most the order of magnitude of tens), the proposed analysis can nevertheless provide useful indications.

Considering Eq. (7),  $b_k^{(m)}$  can be expressed as:

$$b_k^{(m)} = \frac{W_k^{(m+1)}}{W^{(m)}} = \frac{1}{W^{(m)}} \cdot \frac{12 \left( \sum_{i=1}^n R_i^2 \right) - 3(m+1)^2 n(n+1)^2 + 24 \left[ \sum_{i=1}^n (R_i r_i) \right] + 12 \left( \sum_{i=1}^n r_i^2 \right)}{(m+1)^2 n(n^2 - 1) - (m+1) \left( \sum_{j=1}^m T_j \right) - (m+1) T_{m+1}} \tag{23}$$

The asymptotic order of magnitude of the various addenda in the numerator and denominator can then be determined using the so-called asymptotic notation by Bachmann-Landau, with respect to  $n$  and  $m$  (Knuth 1976). The numerator consists of four addenda:

$$12 \left( \sum_{i=1}^n R_i^2 \right) \in O(n^3 \cdot m^2), \tag{24}$$

$$-3(m+1)^2 n(n+1)^2 \in O(n^3 \cdot m^2), \tag{25}$$

$$+24 \left[ \sum_{i=1}^n (R_i r_i) \right] \in O(n^3 \cdot m), \tag{26}$$

$$+12 \left( \sum_{i=1}^n r_i^2 \right) \in O(n^3), \tag{27}$$

where the operator “ $O()$ ” (to be read as “big- $O$ ”) describes the asymptotic order of magnitude of a generic function (Knuth 1976). We specify that, being the  $R_i$  terms given by the sum of  $m$  rankings that  $\in [1, n]$ , they will  $\in O(n \cdot m)$ .

Considering the  $n$  parameter alone, we note that all four addenda in the numerator are asymptotically of the same order of magnitude— $O(n^3)$ —so no one will tend to predominate over the others as  $n$  increases.

On the other hand, considering only the  $m$  parameter, we note that the fourth addend in the numerator, which contains the term  $\sum_{i=1}^n r_i^2$ , will tend to be asymptotically negligible compared to the third addend, which contains the term  $\sum_{i=1}^n (R_i r_i)$ . It can therefore be argued that as  $m$  increases, the contribution of the term  $\sum_{i=1}^n r_i^2$  tends to be negligible with respect to that of the term  $\sum_{i=1}^n (R_i r_i)$ , i.e., using the Bachmann-Landau notation (Knuth 1976):

$$\sum_{i=1}^n r_i^2 = o \left[ \sum_{i=1}^n (R_i r_i) \right], \tag{28}$$

where the operator “ $o()$ ” (to be read as “little- $o$ ”) describes the asymptotic negligibility of the first function compared to the second one (Knuth 1976). We further note that, although Eq. (28) exists, both terms  $\sum_{i=1}^n r_i^2$  and  $\sum_{i=1}^n (R_i r_i)$ , in the denominator of the fraction in Eq. (23), will tend to be asymptotically negligible compared to the first two addenda (cf. Eqs. (24) and (25), being both  $\in O(m^2)$ ).

With reference to the denominator, it consists of three addenda:

$$(m+1)^2 n(n^2-1) \in O(n^3 \cdot m^2), \quad (29)$$

$$-(m+1) \left( \sum_{j=1}^m T_j \right) \in O(n^3 \cdot m^2), \quad (30)$$

$$-(m+1) T_{m+1} \in O(n^3 \cdot m). \quad (31)$$

We specify that the  $T_j$  values and  $T_{m+1}$  are  $O(n^3)$  (see also the considerations about  $T_j$  and  $T_{m+1}$  in Sects. 2 and 3.2). Again, we note that the three addenda are asymptotically of the same order of magnitude with respect to parameter  $n$ . With respect to  $m$ , on the other hand, it emerges that the third addend (containing the term  $T_{m+1}$ ) is asymptotically negligible compared to the other terms, that is:

$$-(m+1) T_{m+1} = o \left[ (m+1)^2 n(n^2-1) - (m+1) \left( \sum_{j=1}^m T_j \right) \right]. \quad (32)$$

In conclusion, this asymptotic analysis has not shown any clear trend, if not a certain predominance of the contribution of the term  $\sum_{i=1}^n (R_i r_i)$  compared to  $\sum_{i=1}^n r_i^2$  and  $T_{m+1}$ , as  $m$  grows. Therefore, aggregation models that maximise  $\sum_{i=1}^n (R_i r_i)$ , such as the BC model, will have a better chance of maximising the  $b_k^{(m)}$  indicator.

## References

- Arrow KJ (2012) Social choice and individual values, 3rd edn. Yale University Press, New Haven
- Borda JC (1781) *Mémoire sur les élections au scrutin*, *Comptes Rendus de l'Académie des Sciences*. Translated by Alfred de Grazia as Mathematical derivation of an election system. *Isis* 44:42
- Borisov N, Weyers B, Kluge A (2018) Designing a human machine interface for quality assurance in car manufacturing: an attempt to address the “Functionality versus User Experience Contradiction” in Professional Production Environments. *Adv Hum Comput Interact*. <https://doi.org/10.1155/2018/9502692>
- Boyd TM, Markman SJ (1983) The 1982 amendments to the Voting Rights Act: a legislative history. *Wash Lee L Rev* 40:1347
- Brown DM (2013) *Designing together: the collaboration and conflict management handbook for creative professionals*. Pearson Education, London
- Cagan J, Vogel CM (2012) *Creating breakthrough products: innovation from product planning to program approval*, 2nd edn. FT Press, Upper Saddle River
- Chen S, Liu J, Wang H, Augusto JC (2012) Ordering based decision making—a survey. *Inf Fusion* 14(4):521–531
- Chiclana F, Herrera F, Herrera-Viedma E (2002) A note on the internal consistency of various preference representations. *Fuzzy Sets Syst* 131(1):75–78
- Cook WD (2006) Distance-based and ad hoc consensus models in ordinal preference ranking. *Eur J Oper Res* 172(2):369–385
- Cook WD, Seiford LM (1978) Priority ranking and consensus formation. *Manag Sci* 24(16):1721–1732
- Cook WD, Seiford LM (1982) On the Borda-Kendall consensus method for priority ranking problems. *Manag Sci* 28(6):621–637
- Dong A, Hill AW, Agogino AM (2004) A document analysis method for characterizing design team performance. *J Mech Des* 126(3):378–385
- Dwarakanath S, Wallace KM (1995) Decision-making in engineering design—observations from design experiments. *J Eng Des* 6(3):191–206
- Dym CL, Wood WH, Scott MJ (2002) Rank ordering engineering designs: pairwise comparison charts and Borda counts. *Res Eng Design* 13:236–242
- Emerson P (2013) The original Borda count and partial voting. *Soc Choice Welfare* 40(2):353–358
- Fishburn PC (1973a) *The theory of social choice*. Princeton University Press, Princeton
- Fishburn PC (1973b) Voter concordance, simple majorities, and group decision methods. *Behav Sci* 18:364–376
- Franceschini F, Maisano D (2015) Checking the consistency of the solution in ordinal semi-democratic decision-making problems. *Omega* 57(1):188–195
- Franceschini F, Maisano D (2017) Consistency analysis in quality classification problems with multiple rank-ordered agents. *Qual Eng* 29(4):672–689
- Franceschini F, Garcia-Lapresta JL (2019) Decision-making in semi-democratic contexts. *Inf Fusion* 52(1):281–289
- Franceschini F, Maisano D (2019a) Design decisions: concordance of designers and effects of the Arrow’s theorem on the collective preference ranking. *Res Eng Design* 30(3):425–434
- Franceschini F, Maisano D (2019b) Fusing incomplete preference rankings in manufacturing decision-making contexts through the ZM<sub>II</sub>-technique. *Int J Adv Manuf Technol* 103(9–12):3307–3322
- Franceschini F, Maisano D (2020) Decision concordance with incomplete expert rankings in manufacturing applications. *Res Eng Design* 31(4):471–490
- Franceschini F, Maisano D, Mastrogiacomo L (2015) Customer requirements prioritization on QFD: a new proposal based on the generalized Yager’s algorithm. *Res Eng Design* 26(2):171–187
- Franceschini F, Maisano D, Mastrogiacomo L (2016) A new proposal for fusing individual preference orderings by rank-ordered agents: A generalization of the Yager’s algorithm. *Eur J Oper Res* 249(1):209–223
- Franceschini F, Galetto M, Maisano D (2019) *Designing performance measurement systems: theory and practice of key performance indicators*. Springer International Publishing, Cham, Switzerland
- Franssen M (2005) Arrow’s theorem, multi-criteria decision problems and multi-attribute preferences in engineering design. *Res Eng Des* 16(1–2):42–56
- Frey DD et al (2009) The Pugh controlled convergence method: model-based evaluation and implications for design theory. *Res Eng Des* 20(1):41–58
- Frey DD et al (2010) Research in engineering design: the role of mathematical theory and empirical evidence. *Res Eng Des* 21(3):145–151

- Fu K, Cagan J, Kotovsky K (2010) Design team convergence: the influence of example solution quality. *J Mech Des* 132(11):111005
- Gibbons JD, Chakraborti S (2010) *Nonparametric statistical inference* (5th ed). CRC Press, Boca Raton, ISBN 978-1420077612
- Grebici K, Ouertani MZ, Blanco E, Gzara-Yesilbas L, Rieu D (2006) Conflict management in design process: focus on changes impact. *Front Artif Intell Appl* 143:161–168
- Hammond JS, Keeney RL, Raiffa H (2015) *Smart choices: a practical guide to making better decisions*. Harvard Business Review Press, Brighton
- Hardy GH, Littlewood JE, Pólya G (1952) *Inequalities*. Cambridge Mathematical Library 2nd ed. (section 10.2, Theorem 368), Cambridge University Press, Cambridge, ISBN 0-521-05206-8
- Hazelrigg GA (1996) The implications of Arrow's impossibility theorem on approaches to optimal engineering design. *J Mech Des* 118(2):161–164
- Hazelrigg GA (1999) An axiomatic framework for engineering design. *J Mech Des* 121(3):342
- Hazelrigg GA (2010) The Pugh controlled convergence method: model-based evaluation and implications for design theory. *Res Eng Des* 21(3):143–144
- Herrera-Viedma E, Cabrerizo FJ, Kacprzyk J, Pedrycz W (2014) A review of soft consensus models in a fuzzy environment. *Inf Fusion* 17:4–13
- Hoyle C, Chen W (2011) Understanding and modelling heterogeneity of human preferences for engineering design. *J Eng Des* 22(8):583–601
- Jacobs JF, van de Poel I, Osseweijer P (2014) Clarifying the debate on selection methods for engineering: arrow's impossibility theorem, design performances, and information basis. *Res Eng Des* 25(1):3–10
- Kaldate A, Thurston D, Emamipour H, Rood M (2006) Engineering parameter selection for design optimization during preliminary design. *J Eng Des* 17(4):291–310. <https://doi.org/10.1080/09544820500274027>
- Katsikopoulos K (2009) Coherence and correspondence in engineering design: informing the conversation and connecting with judgment and decision-making research. *Judgm Decis Making* 4(2):147–153
- Kendall (1962) Rank correlation methods. Griffin & C, London
- Kendall MG, Smith BB (1939) The problem of m-rankings. *Ann Math Stat* 10:275–287
- Keeney RL (2009) The foundations of collaborative group decisions. *Int J Collab Eng* 1:4
- Knuth DE (1976) Big omicron and big omega and big theta. *ACM Sigact News* 8(2):18–24
- Ladha K, Miller G, Oppenheimer J (2003) Information aggregation by majority rule: theory and experiment. <https://www.gvptsites.umd.edu/oppenheimer/research/jury.pdf>
- Legendre P (2005) Species associations: the Kendall coefficient of concordance revisited. *J Agric Biol Environ Stat* 10:226
- Legendre P (2010) Coefficient of concordance. In: Salkind NJ (ed) *Encyclopedia of research design*, vol 1. SAGE Publications Inc, Los Angeles, pp 164–169
- Li H, Bingham JB, Umphress EE (2007) Fairness from the top? Perceived procedural justice and collaborative problem solving in new product development. *Organ Sci* 18(2):200–216
- McComb C, Goucher-Lambert K, Cagan J (2017) Impossible by design? Fairness, strategy and Arrow's impossibility theorem. *Des Sci* 3:1–26
- Myers JL, Well A, Lorch RF (2010) *Research design and statistical analysis*. Routledge, Abingdon
- Nisan N (ed) (2007) *Algorithmic game theory*. Cambridge University Press, Cambridge
- Nurmi H (2012) On the relevance of theoretical results to voting system choice. In: Felsenthal DS, Machover M (eds) *Electoral systems: studies in choice and welfare*. Springer, Berlin, pp 255–274
- Paulus PB, Dzindolet MT, Kohn N (2011) Collaborative creativity, group creativity and team innovation. In: Mumford MD (ed) *Handbook of organizational creativity*. Elsevier, Amsterdam, pp 327–357
- Reich Y (2010) My method is better! *Res Eng Design* 21(3):137–142
- Ross SM (2009) *Introduction to probability and statistics for engineers and scientists*. Academic Press, Cambridge
- Saari DG (1995) *Geometry of voting*. Elsevier B.V, Amsterdam
- Saari DG (2011) *Decision and elections*. Cambridge University Press, Cambridge
- Saari DG, Sieberg KK (2004) Are partwise comparisons reliable? *Res Eng Design* 15(1):62–71
- Scott MJ, Antonsson EK (1999) Arrow's theorem and engineering design decision making. *Res Eng Des* 11:218–228
- See TK, Lewis K (2006) A formal approach to handling conflicts in multiattribute group decision making. *J Mech Des* 128(4):678
- Spearman C (1904) The proof and measurement of association between two things. *Am J Psychol* 15(1):72–101
- Stark PB (2008) Conservative statistical post-election audits. *Ann Appl Stat* 2(2):550–581
- Weingart LR et al (2005) Functional diversity and conflict in cross-functional product development teams: considering representational gaps and task characteristics. In: Neider LL, Schriesheim CA (eds) *Understanding teams*. Information Age Publishing, Charlotte, pp 89–110
- Yeo SH, Mak MW, Balon SAP (2004) Analysis of decision-making methodologies for desirability score of conceptual design. *J Eng Des* 15(2):195–208

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.