Clarifying the debate on selection methods for engineering: Arrow's impossibility theorem, design performances, and information basis

Johannes F. Jacobs · Ibo van de Poel · Patricia Osseweijer

Abstract To demystify the debate about the validity of selection methods that utilize aggregation procedures, it is necessary that contributors to the debate are explicit about (a) their personal goals and (b) their methodological aims. We introduce three additional points of clarification: (1) the need to differentiate between the aggregation of preferences and of performances, (2) the application of Arrow's theorem to performance measures rather than to preferences, and (3) the assumptions made about the information that is available in applying selection methods. The debate about decision methods in engineering design would be improved if all contributors were more explicit about these issues.

1 Introduction

In his editorial "My method is better!", Reich (2010) has pointed out that implicit presuppositions and arguments need to be laid out in the open to enable a serious and reflective debate on the validity of methods for dealing with multi-criteria problems in engineering design. In his

J. F. Jacobs (✉) · P. Osseweijer
Biotechnology and Society, Department of Biotechnology, Delft University of Technology, Julianalaan 67, 2628 BC Delft, The Netherlands
e-mail: j.f.jacobs@tudelft.nl

I. van de Poel
Department of Philosophy, School of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands

P. Osseweijer
Kluyver Centre for Genomics of Industrial Fermentation, PO Box 5057, 2600 GA Delft, The Netherlands

analysis of the debate, Reich suggested that misinterpretations could be avoided if authors clearly state which kind of goals they strive for when presenting certain arguments. He makes a distinction between the goal to improve design practice and the goal to obtain theoretical rigor. Katsikopoulos (2009, 2012) also analyzed the debate and indicates an important distinction between methodological aims in order to achieve coherence or correspondence.[1] He argues that the debate can be improved if authors would acknowledge this distinction.

Several authors have thus investigated the debate and their suggestions will contribute to the development of a more reflective debate. However, we want to point out some further unclarities and implicit assumptions that also need to be addressed to engage in a fully open debate on matrix-based selection methods.

This paper is structured as follows. In section two, the selection methods with a matrix-based structure, which are used in engineering design practice to deal with multi-criteria problems, are described. We argue that Arrow's impossibility theorem affects decision-making in engineering design in two distinct ways, knowingly the aggregation of preferences and of performances. In the following section, we will modify the theorem to design performances as utilized in engineering design rather than to preferences as in Arrow's original formulation. In section four, the presumed amount of information, which is available for the multi-criteria selection problem, is discussed in terms of uncertainty, comparability, and measurability. In the final section, we will draw conclusions by

---

[1] Coherence refers to the internal consistency of the decision-making process (e.g., the method contains no logical contradictions), while correspondence refers to the external consistency (e.g., success of the made decision in the real world).

suggesting several ways in which the current debate on the validity of matrix selection methods can be improved.

2 Preferences and design performances

Selecting a single or a few promising design concepts from a larger set is an important part of engineering design. The selection of the "best" design concept(s) for further design and development is a decision-making activity that is based on various criteria, which are, in turn, based on the design specification and requirements. Various methodologies have been developed to support this type of multi-criteria selection. Methods that use a matrix structure are frequently applied in engineering design practices, for example, the analytic hierarchy process, Pahl and Beitz method, Pugh's concept selection,[2] quality function deployment, and weighted-product method (Akao 2004; Hauser and Clausing 1988; Pahl and Beitz 2005; Saaty 1980). These matrix methods visually indicate how the alternative designs are scored on various criteria by tabulating the performances of the design concepts on each criterion in a chart (see Table 1 for an example). A global performance structure (e.g., the scores on the three alternatives in Table 1) is generated by means of an aggregation procedure based on the performances on the various criteria. The global performance structure is used to guide the selection of the best alternative design for further development.

This kind of decision methods is very similar to a multi-criteria decision analysis in which the diverse performances are aggregated into one overall performance score. The aggregation of diverse measures into a global measure is, however, not straightforward. A similar aggregation problem has been studied within the field of voting theory and welfare economics that later formed the field of social choice theory. A huge literature was initiated, most importantly by the works of Arrow (1950) and May (1952), in which the aggregation of individual preferences into a collective preference of a group is analyzed. Various authors have claimed that the results of these theories are applicable to multi-criteria decision analysis, because the problems of social choice and multi-criteria decision analysis are structurally identical (Franssen 2005; Bouyssou et al. 2009).

One of the best-known results from social choice theory is the famous impossibility theorem that was proved by the

Table 1 Weighted-sum method with the performance measure that is based on five-point ranks

Criteria	Weight	Concept A	Concept B	Concept C
Yield	1	4	5	2
Safety	1	5	1	3
Controllability	1	4	5	2
Revenues	1	1	2	4
Score	–	14	13	11

Performance measure: 1 = inadequate, 2 = weak, 3 = satisfactory, 4 = good, 5 = excellent

economist Kenneth Arrow in 1950. Arrow's impossibility theorem states that if there are a finite number of individuals and there are at least three options to choose from,[3] no aggregation method can simultaneously satisfy five general conditions [see Sen (1995) for a short argumented proof or see Blackorby et al. (1984) for a geometric proof]. These conditions are as follows:

- *Collective rationality* The collective preference ordering[4] must be complete and transitive. The preference ordering is complete if it accounts for all considered options. Transitivity demands that if option A is preferred to B and B to C, A is also preferred to C.
- *Independence of irrelevant alternatives* The collective preference ordering has the same ranking of preferences among a subset of options as it would for a complete set of options.
- *Non-dictatorship* The collective preference ordering may not be determined by a single preference order.
- *Unrestricted domain* The profile of single preference orders is only restricted with respect to transitivity and completeness.
- *Weak Pareto principle* Providing that one option is preferred to another option in all single preference orders, then so must be the resulting collective preference order.

The theorem prevents the construction of a generally acceptable aggregation method that combines the single preferences into a collective preference structure, such that it represents the preferences of the group as a whole and satisfies the five conditions mentioned. Arrow's impossibility theorem only bars general procedures, because aggregation may still be possible in specific cases in which the single preferences align in very specific ways; for

---

[2] The earlier Pugh's controlled conversion method (Pugh 1991, 1996) is mainly a tool to enhance design creativity. Several others have adapted the method into a selection tool by calculating net scores (Ulrich and Eppinger 2008) or overall ratings (Eggert 2005). These Pugh-like methods are also known as Pugh's concept selection methods.

[3] The impossibility result does not hold in the situation that there are only two alternatives to choose from; see for example May's theorem.

[4] A single preference ordering refers to the set of ordinal measures between different options, while the collective preference ordering refers to the aggregated single preference orderings over these same options.

example, when all agents prefer the same outcome over all others or when all agents have single-peaked preferences over the considered outcomes.

In engineering design, the choice between design concepts is usually a team effort, involving designers, clients, and managers which all may have different opinions on what the "best" design is. The combination of all these various single preferences into a final decision that selects the "best" design among various alternatives is very similar to the type of problem that is considered in social choice theory. Various authors have identified these similarities of selections made by groups and have claimed that Arrow's impossibility theorem affects engineering design decisions in this manner (Hazelrigg 1996; Lowe and Ridgway 2000; Van de Poel 2007).

This is, however, not the only way in which Arrow's impossibility theorem affects engineering design. The aggregation problem of design performances into a global performance structure is also similar in nature to the aggregation problem of Arrow. Nonetheless, there is a difference. The most commonly applied aggregation procedures are indeed quite similar to those studied in social choice theory, but the object of aggregation is not. In the case of social choice theory, as well as in the case of group selection in engineering design, the object of aggregation is individual preferences. In contrast, in the selection of design alternatives, the object of aggregation is performances on design criteria.

To summarize, there are two distinct ways Arrow's impossibility theorem can affect decision-making in engineering design, knowingly the aggregation problem of (a) preferences and (b) design performances.

The distinction between preferences and design performances could help to clarify the existing debate. An example is the discussion in this journal about the way Arrow's impossibility theorem affects the Pugh controlled convergence method. Franssen challenges the Pugh controlled convergence method by arguing that the method "explicitly aims to arrive at a global judgment on the relative worth of several design concepts" (Franssen 2005, p 54) and that the "preference order is not explicitly part of Pugh's method, but its existence must be presumed in order to understand how the designer arrives at the comparative judgments" (Franssen 2005, p 55). Franssen, therefore, concludes that the Pugh controlled convergence method will "run into the kind of difficulties associated with Arrow's theorem" (Franssen 2005, p 55). In essence, he claims that the method presumes the formation of preference, even though the method uses a set of design criteria that would indicate the use of design performances. Furthermore, Franssen claims that these presumed preference structures on various criteria are aggregated to achieve a

selection, which implies that the method can be challenged with the impossibility theorem.

Hazelrigg and Frey et al. also discuss the validity of the Pugh controlled convergence method. These authors disagree with each other about the question whether the Pugh method entails voting, because voting would entail the aggregation of preferences of individual persons, and as a result, the method could be challenged along the lines of Arrow's impossibility theorem. Frey and coauthors "note that there is no voting in Pugh's method" (Frey et al. 2009, p 43). Hazelrigg responds by pointing out that "voting is used in the Pugh method, firstly to obtain consensus on the relative merits of candidate designs compared to the datum design and second to aggregate the symbols … in the Pugh matrix" (Hazelrigg 2010, p 143). In a reply, Frey et al. argue against the position of Hazelrigg and hold that for building of consensus "there is not voting in the Pugh method" (Frey et al. 2010, p 147). The discussion is about the need of voting in the practical utilization of the Pugh controlled convergence method, since the method does not explicitly state how consensus should be reached (Pugh 1991, 1996). Pugh controlled convergence method does explicitly prevent any general attempt to aggregate the performance measures in the matrix, because it is stated that "[t]he scores or numbers must not, in any sense, be treated as absolute; they are for guidance only and must not be summed algebraically" (Pugh 1991, p 77). Hence, Arrow's impossibility theorem cannot be used to challenge the method on this specific point.

Another example is the debate of Hazelrigg, Franssen, Scott, and Antonsson about the importance of Arrow's impossibility theorem for engineering design in which the issue seems to be the switch between the two different measures. Hazelrigg (1996) claims that the theorem "holds great importance to the theory of engineering design" and that the methods "Total Quality Management (TQM) and Quality Function Deployment (QFD), are logically inconsistent and can lead to highly erroneous results" (Hazelrigg 1996, p 161). Hazelrigg claims that "customer satisfaction, taken in the aggregate for the group of customers, does not exist" (Hazelrigg 1996, p 163). Hence, the arguments of Hazelrigg are based on the aggregation of single preferences into a collective preference structure. Scott and Antonsson (1999) have challenged the conclusion of Hazelrigg. In the view of Scott and Antonsson, the "engineering design decision problem is a problem of decision with multiple criteria" (Scott and Antonsson 1999, p 218). The authors claim that in "engineering design, there may be many people involved, but decisions still depend upon the aggregation of engineering criteria" (Scott and Antonsson 1999, p 220). Hence, the authors are considering design performances instead of preferences.

Scott and Antonsson reject the conclusions of Hazelrigg on the grounds that the conditions of Arrow's theorem are unreasonable for design performances. In turn, Franssen (2005) criticizes the conclusion of Scott and Antonsson. Franssen argues that "Arrow's theorem applies to multi-criteria decision problems in engineering design as well" (Franssen 2005, p 43). He claims that the aggregation problem of preferences is structurally identical to that of design performances and that the conditions are reasonable for engineering design, if interpreted in the right way.

This leads us to the following point of clarification. Although many authors have discussed the implications of Arrow's impossibility theorem on engineering design with regard to the aggregation problem of performance measures, no one has to our knowledge explicitly adapted the theorem and its conditions from the context of preferences to the context of design performances.

3 Impossibility theorem for performance aggregation

We have adapted the five conditions of Arrow's impossibility theorem on preference aggregation into an Arrowian type of impossibility theorem for performance aggregation in engineering design[5] as follows:

- *Independence of irrelevant concepts* The global performance structure[6] between two design concepts depends only on their performances and not on the performances of other design concepts.
- *Non-dominance* The global performance may not be determined by a single performance structure.
- *Unrestricted scope* The performance structures on a single criterion as well as the global performance structure are only restricted with respect to transitivity, reflexivity, and completeness. Completeness requires that the performance structure accounts for all design concepts. A performance structure is reflective if the separate performance measure can be compared to themselves. Transitivity requires that if performance A is related to performance B and performance B in turn is related in the same way to

performance C, then performance A is likewise related to performance C.[7]

- *Weak Pareto principle* Providing that one design concept is strictly better than another design concept in all the single performance structures, then so must be the resulting global performance structure for these two concepts.

The Arrowian impossibility theorem for performance aggregation in conceptual engineering design then states that *if there is a finite number of evaluation criteria and there are at least three alternative design concepts*, *no aggregation method can simultaneously satisfy independence of irrelevant concepts*, *non-dominance*, *unrestricted scope*, *and weak Pareto principle*. Similarly to the original theorem, this theorem only shows that such an aggregation is unattainable in general; so, it may still be possible in very specific cases, for instance, when all criteria give single-peaked performances over the considered design concepts. So, aggregation is possible when one design concept has the best performance on all the listed criteria. Such cases will be the exception rather than the rule in engineering design practice, in which most design decisions involve trade-offs and value-based choices. The four Arrowian conditions are thus already enough to prove the general impossibility result, even though we might wish to impose other conditions on decision-aiding tools for engineering design, such as non-manipulability (preventing unfair influence on the results) and separability (allowing reduction of the amount of design concepts in several subsequent phases).

Prima facie The above-described conditions seem quite reasonable for engineering design practices. If one accepts the conditions, the Arrowian impossibility theorem indicates that the decision matrix methods, which aggregates the performance measures p_{ij} ($i = 1, \ldots, n$; $j = 1, \ldots, m$) over n options on m criteria into a global performance structure s_i ($i = 1, \ldots, n$), can generate misleading conclusions by introducing significant logical errors in the decision process. This point can be illustrated by an example, which is depicted in Tables 1 and 2. Suppose that there are three conceptual designs of which only one will have to be chosen for further development. The design

---

[5] The theorem is formulated in such a way that it is consistent with all three fundamental measures of scale as discussed in Sect. 4.2. The original theorem of Arrow, as presented in Sect. 2, is restricted to ordinal measures of preference.

[6] A performance structure refers to the set of relations between different design concepts. The global performance structure (e.g., the scores on the three alternative concepts in Table 2) is the resulting set of aggregated relations of the performance structures on the various criteria (e.g., the ranks on the three alternative concepts on the four criteria in Table 2).

[7] In a more formal sense, reflexivity means that for all a that are an element of set A, it holds that *aRa*, with R representing a binary relation on set A (e.g., *a* is greater than or equal to *a*). Transitivity means that for all *a*, *b*, and *c* that are an element of set A, it holds that *aRb* and *bRc* imply *aRc* (e.g., *a* is equal to *b* and *b* is equal to *c*, implying that *a* is equal to *c*). The asymmetric part of R is the binary relation P defined by *aPb* that is equivalent to *aRb* and not *bRa*. The symmetric part of R is the relation I defined by *aIb* that is equivalent to *aRb* and *bRa*. In social choice theory, the relation is normally taken as binary; nonetheless, the binariness of choice is unimportant for the impossibility result of Arrow's theorem (see Sen 1986).

Table 2 Weighted-sum method with a different performance measure that is instead based on three-point ranks

Criteria	Weight	Concept A	Concept B	Concept C
Yield	1	2	3	1
Safety	1	3	1	2
Controllability	1	2	3	1
Revenues	1	1	2	3
Score	–	8	9	7

Performance measure: 1 = worst, 2 = neutral, 3 = best

team has made a list of four evaluation criteria, namely production yield, process safety, controllability of the system, and economic revenues. To facilitate the choice, the design team uses the weighted-sum method (sometimes referred to as the Pahl and Beitz method) and scores the performance measure by a five-step point ranking. The resulting decision matrix is given in Table 1 and indicates that design concept A should be selected by the design team. Now, suppose that the team uses a three-step point rank instead, in which the concept with the best performance on a criterion receives three points, the worst scoring concept only one point and the other in-between option two points (see Table 2). Using the same aggregation procedure, the weighted-sum method now indicates that concept B should be selected by the design team instead. The decision procedure of point ranking in combination with weighted-sum aggregation thus gives rationally inconsistent results. In terms of Arrowian impossibility theorem, this distortion is the result of violating the "independence of irrelevant concepts" condition. The three-point rank allows a comparison of three design concepts, while two additional concepts can be evaluated on a five-point rank. Hence, the not shown additional concepts influence the global performance structure and so violate the Arrowian condition.[8]

4 Information availability

A third point of difficulty for a serious reflective debate is the presumed amount of information that is available for the selection of alternative solutions on the basis of multiple criteria in engineering design. We will discuss the uncertainties that plague engineering design work as well as the information basis of the performance measures in relation to preference and performance aggregation.

---

[8] It should be noted that the point ranking as used in the example can, at best, be interpreted as ordinal scale measures. If the point ranks are measured on an ordinal scale (see Sect. 4.2), the weighted-sum method gives meaningless scores, because it uses an inapplicable arithmetic operation (addition of performance measures) to calculate the scores of the global performance.

4.1 Design for an uncertain future

The selection methods aim to support the process of selecting a single or a few promising design concepts from among various alternative concepts. In order to make a decision, the performance of the designed artifact has to be evaluated. The design concepts under evaluation are only abstract embodiments of designs, which are given physical shape in detailed design and subsequent construction. The decision maker thus needs to predict the final physical form of the artifact under design. Furthermore, to evaluate the performance of this still to-be created artifact, the decision maker needs to envision the mode of the artifact's application and the effects this application will bring about. In other words, the decision process requires the prediction of future states. These predictions are made under uncertainty, which is the result of various factors: lack of knowledge (known unknowns), ignorance (unknown unknowns), system complexity, and ambiguity.

The degree of uncertainty depends on the type of design. Both Vincenti (1992) and Van Gorp (2005) make a distinction between normal and radical design. In normal design, both the operational principle and the configuration are kept similar to already implemented designs, while in radical design, the operational principle and/or configuration deviates from the convention or is unknown. In normal design, the degree of uncertainty in predicting future states is smaller compared to radical design, because in the former case, there is an effective basis of experience to base the predictions on. The degree of the uncertainty that is faced in the decision process is also dependent on the lifetime of the artifact and/or the duration of its (potential) effects. Forecasting becomes more difficult when it spans larger amounts of time. The degree of uncertainty also depends on the design phase (e.g., conceptual, preliminary, or detailed). As the design project proceeds, more information comes available and more features become defined. Moreover, the time span to the actual construction and application becomes discernibly smaller, so decreasing uncertainty. For the clarity of the debate, it is thus important that authors make explicit what degree of uncertainty the selection method is expected to incur in relation to the design type, design phase, and kind of artifact.

In the impossibility theorems for preference aggregation, as well as in the one for performance aggregation, it is assumed that adequate predictions can be made so that preferences and performance measures can be obtained without the burden of uncertainty. It is thus clear that the uncertainties, especially in conceptual phase of radical design, will complicate the selection procedure beyond the difficulties presented by the impossibility theorem. Although normal design work in the detailing phase has far fewer uncertainties about the possible future states that

need to be assessed in selecting design concepts, it does not avoid the clutches of the impossibility theorem. Even in this case, the information basis is too limited for proper aggregation due to problems of measurability and comparability.

4.2 Measurability

In his work, Sen (1977) has pointed out that the impossibility result of Arrows theorem can be interpreted from an informational perspective. Arrow (1950) did not include the uncertainties surrounding the interpretation of future states, which are necessary for the formation of preferences. He presupposed an ideal situation in which the options of selection are known. This makes his impossibility so strong, because in practice even more difficulties will arise. Furthermore, Arrow (1950) assumes that certain information is not available as he presumes the incomparability of preferences and supposes that the measurability of preferences is restricted to ordinal scales. It is possible to relax this restriction by assuming measurements on interval or ratio scales, which would make intensity and ratio information available for the decision. The most important properties of the four fundamental measurement scales are presented in Table 3 (Stevens 1946).

However, allowing for more information about preference/performance measurements on interval scales does not avoid the clutches of an Arrowian type of theorem. Kalai and Schmeidler (1977) presented a proof of an impossibility theorem that uses interval scale information; however, this proof needs an additional condition. Hylland (1980) extended the proof of Kalai and Schmeidler in such a way that the added condition is not required.

One step further is to consider ratio scale information for preference/performance measurements. However, this also seems no way out of the impossibility result, because Tsui and Weymark (1997) have shown that the theorem also

holds when one allows for ratio scale measurements. Furthermore, it seems unfeasible to obtain such scales for the performance/preference measures, because design concepts are abstract embodiments that still need to be given physical form. The performance/preference measures thus refer to mental notions about what accounts for a "good" design. The measure represents a (value) judgment, and arguably, this does not allow for ratio scale measurements. Moreover, almost no examples of ratio scale measurement exist in the behavioral sciences. Summarizing, the Arrowian type of theorem for multi-criteria concept decisions in engineering design can be extended to interval and ratio scale measurements of preferences and performances, though the impossibility result is preserved.

4.3 Comparability

Let us turn to the comparability part of the informational restriction. Arrow's theorem in social choice theory is set up "to exclude interpersonal comparison of social utility either by some form of direct measurement or by comparison with other alternative social states" (Arrow 1950, p 342). For the Arrowian impossibility theorem for multi-criteria concept decisions in engineering design, this would mean that it rules out the possibility to make direct comparisons between the performances of the conceptual designs on various design criteria. In the field of social choice theory, the possibility of interpersonal comparisons has been investigated. Hammond (1976) has shown that there are some possible aggregation methods when the information basis is limited to ordinal scale comparability, even when using stronger forms of the Pareto principle and non-dictatorship conditions plus an additional separability condition. D'Aspremont and Gevers (1977) have proved the feasibility of aggregation with interval scales and unit comparability. Deschamps and Gevers (1978) have developed this further for cases of full comparability. Roberts (1980) showed that there are even more possible aggregation methods when allowing for ratio scale comparability, under the conditions of Arrow's impossibility theorem.

The mutual effects of comparability and measurements scales on the feasibility of aggregation are presented in Table 4. It thus becomes clear that if Arrow's theorem is relaxed in such a way that it allows for more preference/performance information in the form of comparability, there are admissible aggregation methods. However, the comparability of preference as well as performance measures is not straightforward.[9] Comparability means that

Table 3 Four fundamental measurements of scale and their properties

Scale	Mathematical group structure	Admissible transformation	Example
Nominal[a]	Permutation	One-to-one substitution	Classification of data
Ordinal	Isotonic	Monotonic increasing	Mohs scale of mineral hardness
Interval[b]	General linear	Positive affine	Celsius scale of temperature
Ratio	Similarity	Scalar	Kilogram scale of weight

[a] Also refers to as a categorical scale

[b] In economics, the term "cardinal value scale" is used

---

[9] Even if one allows for interpersonal comparability, it is not obvious that the analogon in inter-criteria comparability also holds. In interpersonal comparisons, preferences could be compared with respect to the same value (e.g., human welfare), whereas design criteria may relate to different values.

Table 4 Aggregation of preference/performance structures into a global preference/performance structure in accordance with the Arrowian conditions depending on measurability and comparability

Information basis	Scales of measure		
	Ordinal scale	Interval scale	Ratio scale
Incomparable	Impossible (Arrow 1950)	Impossible (Hylland 1980)	Impossible[a] (Tsui and Weymark 1997)
Comparable[b]	Feasible (Hammond 1976)	Feasible[c] (Deschamps and Gevers 1978)	Feasible (Roberts 1980)

[a] A feasibility result follows if only nonnegative or positive preference/performance measures are used

[b] The notion of comparability varies with the choice of measurement scales

[c] See d'Aspremont and Gevers (1977) for proof of feasibility result under unit comparability instead of full comparability

there is a relationship between the various measurement scales for preferences or performances. For example, assume that the strength of the product, the speed of production, and needed investment cost are used as performance measures that account for a "good" design. The strength of the used product cannot be lowered under a certain level regardless of the possible gains in production speed or reduction in investment cost, because the design will certainly fail and be of no value. This results in a weak form of comparability, because trade-offs can only be made within a limited range.[10] So, to escape the impossibility result, all the windows of comparability should match, which would be the exception rather than the rule in engineering design.

There is still a more pressing argument against the comparability of preference/performance measures in engineering design practice. The reality of modern engineering design is that engineers are increasingly required to factor in value-laden criteria (e.g., safety, sustainability, and reliability) into their decision-making process as early as the conceptual design phase. By their very nature, some moral values are considered incomparable. These moral values, which are precluded from trade-offs, are called protected or sacred values (Baron and Spranca 1997; Tetlock 2003), and the inclusion of such value-laden measures will make comparison between them next to impossible. All things considered, it is in principle possible to avoid the kind of difficulties associated with Arrow's theorem by

---

[10] To make performance measures more "comparable," mathematical normalization procedures are sometimes applied. However, such normalization only shifts the comparability issue away from the aggregation method toward the normalization procedure and hence does not resolve the comparability issue.

using comparable preference/performance measures; however, it is our opinion that it would be very exceptional to find that all considered measures are comparable in engineering design practice.

The current discussion would be improved if authors are explicit about the information basis that they assume to be available. This can be illustrated by the seemingly contradictory opinions of Hazelrigg (1999), Franssen (2005) and Keeney (2009). Keeney claims that the "Arrow's impossibility theorem has been misinterpreted by many," (Keeney 2009, p 14) and he explicitly mentions Hazelrigg en Franssen as examples. However, this is not a case of misinterpretation, but of different assumptions about the information basis. Keeney provides a framework, very similar to that of Arrow's theorem, in which the preferences are expressed as von Neumann–Morgenstern expected utilities. In this framework, interpersonal comparability of these preferences is assumed, because "the group expected utility U_j of an alternative is calculated using the individual's expected utilities U_{jk}" (Keeney 2009, p 14). In contrast, Hazelrigg and Franssen do not explicitly deviate from the original comparability assumption of Arrow and thus exclude interpersonal comparability.

5 Conclusions

In order to achieve the reflective debate Reich (2010) envisioned in his editorial, it is necessary that authors are explicit about their personal goals as well as methodological aims in terms of coherence and/or correspondence (Katsikopoulos 2009, 2012). We have presented here several additional issues that cloud the current debate. First of all, difficulties in the debate result from the fact that Arrow's impossibility theorem can affect the decision-making in engineering design in two distinct ways: (a) it impedes methods to aggregate preferences and (b) obstructs ways to combine various performance criteria. Secondly, misconceptions are caused sometimes by an unclear translation of Arrow's original impossibility theorem to the aggregation problem of design performances. In order to resolve this ambiguity, we have presented the Arrowian kind of impossibility theorem for performance aggregation. Thirdly, clarity is also required about the uncertainties associated with the predictability of future states in engineering design decisions, which are dependent on the considered (a) kind of designed artifact, (b) type of design, and (c) design phase. Finally, the debate would be clarified, if the assumed information basis for the decision is made explicit, especially with regard to comparability and measurability. We think the explicit consideration of all these issues will go a long way to come to a truly

reflective debate on decision-making methods for engineering design.

References

Akao Y (2004) Quality function deployment: integrating customer requirements into product design. Productivity Press, New York

Arrow KJ (1950) A difficulty in the concept of social welfare. J Polit Econ 58:328–346. doi:10.1086/256963

Baron J, Spranca M (1997) Protected values. Organ Behav Hum Decis Process 70:1–16. doi:10.1006/obhd.1997.2690

Blackorby C, Donaldson D, Weymark JA (1984) Social choice with interpersonal utility comparisons: a diagrammatic introduction. Int Econ Rev 25:327–356. http://www.jstor.org/pss/2526200

Bouyssou D, Marchant T, Perny P (2009) Social choice theory and multicriteria decision aiding. In: Bouyssou D, Dubois D, Pirlot M, Prade H (eds) Decision-making process: concept and method. ISTE Ltd., London, pp 741–770

D'Aspremont C, Gevers L (1977) Equity and the informational basis of collective choice. Rev Econ Stud 44:199–209. http://www.jstor.org/stable/2297061

Deschamps R, Gevers L (1978) Leximin and utilitarian rules: a joint characterization. J Econ Theory 17:143–163. doi:10.1016/0022-0531(78)90068-6

Eggert RJ (2005) Engineering design. Pearson Education, New Jersey

Franssen M (2005) Arrow's theorem, multi-criteria decision problems and multi-attribute preferences in engineering design. Res Eng Des 16:42–56. doi:10.1007/s00163-004-0057-5

Frey DD, Herder PM, Wijnia Y, Subramanian E, Katsikopoulos K, Clausing DP (2009) The Pugh controlled convergence method: model-based evaluation and implications for design theory. Res Eng Des 20:41–50. doi:10.1007/s00163-008-0056-z

Frey DD, Herder PM, Wijnia Y, Subramanian E, Katsikopoulos K, de Neufville R, Oye K, Clausing DP (2010) Research in engineering design: the role of mathematical theory and empirical evidence. Res Eng Des 21:145–151. doi:10.1007/s00163-010-0085-2

Hammond PJ (1976) Equity, Arrow's conditions, and Rawls' difference principle. Econometrica 44:793–804. http://www.jstor.org/stable/1913445

Hauser J, Clausing D (1988) The house of quality. Harv Bus Rev 66:63–74. doi:10.1225/88307

Hazelrigg GA (1996) The implications of Arrow's impossibility theorem on approaches to optimal engineering design. J Mech Des 118:161–164. doi:10.1115/1.2826864

Hazelrigg GA (1999) An axiomatic framework for engineering design. J Mech Des 121:342–347. doi:10.1115/1.2829466

Hazelrigg GA (2010) The Pugh controlled convergence method: model-based evaluation and implications for design theory. Res Eng Des 21:143–144. doi:10.1007/s00163-010-0087-0

Hylland A (1980) Aggregation procedure for cardinal preferences: a comment. Econometrica 48:539–542. http://www.jstor.org/stable/1911117

Kalai E, Schmeidler D (1977) Aggregation procedure for cardinal preferences: a formulation and proof of Samuelson's impossibility conjecture. Econometrica 45:1431–1438. http://www.jstor.org/stable/1912309

Katsikopoulos KV (2009) Coherence and correspondence in engineering design: informing the conversation and connecting with judgment and decision-making research. Judgm Decis Mak 4:147–153. http://journal.sjdm.org/cck/cck.pdf

Katsikopoulos KV (2012) Decision methods for design: insights from psychology. J Mech Des 134:084504.1–084504.4. doi:10.1115/1.4007001

Keeney RL (2009) The foundations of collaborative group decisions. Int J Collab Eng 1:4–18. doi:10.1504/IJCE.2009.027438

Lowe AJ, Ridgway K (2000) Optimization impossible? The importance of customer segmentation in quality function deployment. Qual Prog 33:59–64. http://www.asq.org/data/subscriptions/qp/2000/0700/qp0700lowe.pdf

May KO (1952) A set of independent necessary and sufficient conditions for simple majority decision. Econometrica 20:680–684. doi:10.2307/1907651

Pahl G, Beitz W (2005) Engineering design: a systematic approach. Springer, London

Pugh S (1991) Total design: integrated methods for successful product engineering. Addison-Wesley, Harlow

Pugh S (1996) Creating innovative products using total design. Addison-Wesley, Reading

Reich Y (2010) Editorial: is my method better? Res Eng Des 21:137–142. doi:10.1007/s00163-010-0092-3

Roberts KWS (1980) Interpersonal comparability and social choice theory. Rev Econ Stud 47:421–446. http://www.jstor.org/stable/2297002

Saaty TL (1980) The analytic hierarchy process: planning, priority setting, resource allocation. McGraw-Hill, New York

Scott MJ, Antonsson EK (1999) Arrow's theorem and engineering design decision making. Res Eng Des 11:218–228. doi:10.1007/s001630050016

Sen A (1977) On weights and measures: informational constrains in social welfare analysis. Econometrica 45:1539–1572. http://www.jstor.org/stable/1913949

Sen A (1986) Chapter 22: social choice theory. In: Arrow KJ, Intriligator MD (eds) Handbook of mathematical economics, vol III. Elsevier, Amsterdam, pp 1073–1181. doi:10.1016/S1573-4382(86)03004-7

Sen A (1995) Rationality and social choice. Am Econ Rev 85:1–24. http://www.jstor.org/stable/2117993

Stevens SS (1946) On the theory of scales of measurement. Science 103:677–680. doi:10.1126/science.103.2684.677

Tetlock PE (2003) Thinking the unthinkable: sacred values and taboo cognitions. Trends Cogn Sci 7:320–324. doi:10.1016/S1364-6613(03)00135-9

Tsui K, Weymark JA (1997) Social welfare orderings for ratio-scale measurable utilities. Econ Theory 10:241–256. doi:10.1007/s001990050156

Ulrich KT, Eppinger SD (2008) Product design and development. McGraw-Hill, New York

Van de Poel I (2007) Methodological problems in QFD and directions for future development. Res Eng Des 18:21–36. doi:10.1007/s00163-007-0029-7

Van Gorp AC (2005) Ethical issues in engineering design: safety and sustainability. Simon Stevin Series in the Philosophy of Technology, Delft

Vincenti WG (1992) Engineering knowledge, type of design, and level of hierarchy: further thoughts about what engineers know. In: Kroes P, Bakker M (eds) Technological development and science in the industrial age. Kluwer, Dordrecht, pp 17–34