Research in Engineering Design 13 (2002) 236–242
DOI 10.1007/s00163-002-0019-8

Rank ordering engineering designs: pairwise comparison charts and Borda counts

Clive L. Dym, William H. Wood, Michael J. Scott

Abstract Designers routinely rank alternatives in a variety of settings using a staple of comparison, the *pairwise comparison*. In recent years questions have been raised about the use of such comparisons as a means of calculating and aggregating meaningful preference or choice data. Results on voting have been used to argue that the positional procedure known as the Borda count is the best pairwise voting procedure, or at least the only one that is not subject to a number of demonstrable problems. We show here that pairwise comparison charts (PCC) provide results that are identical to those obtained by the Borda count, and that the PCC is thus not subject to the arguments used against non-Borda count methods. Arrow's impossibility theorem has also been invoked to cast doubt upon *any* pairwise procedure, including the Borda count. We discuss the relevance of the Arrow property that is lost in the Borda count, the independence of irrelevant alternatives (IIA). While the *theoretical* consequences of IIA are devastating, it is not clear that the same is true of its *practical* consequences. Examples are presented to illustrate some primary objections to pairwise methods.

Keywords Design decisions, Pairwise comparisons, Ranking alternatives, Arrow's theorem, Social choice

Received: 1 February 2002 / Revised: 3 July 2002
Accepted: 3 July 2002 / Published online: 4 September 2002
© Springer-Verlag 2002

C.L. Dym (✉)
Department of Engineering,
Harvey Mudd College,
Claremont, CA 91711-5990, USA
E-mail: clive_dym@hmc.edu
Fax: +1-909-6218967

W.H. Wood
Department of Mechanical Engineering,
University of Maryland, Baltimore County,
Baltimore, MD 21250, USA

M.J. Scott
Department of Mechanical and Industrial Engineering,
University of Illinois at Chicago,
Chicago, IL 60607, USA

The authors are grateful to the National Science Foundation for supporting, and to Donald G. Saari (University of California, Irvine) for organizing the workshop "Decisions and Engineering" in October 2001. This paper derives directly from observations made and discussions initiated at that workshop. Professor Saari also provided helpful comments on an early version of the paper, as did Prof. Patrick Little (Harvey Mudd College).

1 Introduction

Designers routinely rank objectives, design attributes, and designs. In each of these circumstances the designer is charged with creating a set of alternatives – and for efficiency's sake the designer must at some point stop generating and start choosing among the alternatives. So, too, the customer must at some point stop evaluating and start buying among the available attributes and design choices. Simon articulated the principle of *bounded rationality*, which states that we cannot afford to make a decision *and* to generate all of the options because we would then never get beyond generating options (Simon 1996). Thus, any set of options is, by definition, a subset or culling of a global set of options.

In recent years, questions have been raised about the means by which designers establish rankings of alternatives, with a special focus on how pairwise comparisons are performed as a means of assembling information on which basis rankings can be obtained. In *pairwise comparisons*, the elements in a set (i.e. the objectives, design attributes, or designs) are ranked on a pair-by-pair basis, that is, two at a time, until all of the permutations have been exhausted. In each comparison, points are awarded to the winner.[1] Then the points awarded to each element in the set are summed and the rankings are obtained by ordering the elements according to points accumulated.

This methodology has been criticized from two standpoints. In the first, Hazelrigg (1996; G.H. Hazelrigg, personal communication, 2001) argues that aggregating pairwise comparisons violates *Arrow's impossibility theorem*, which can be characterized as a proof that a "perfect" or "fair" voting procedure cannot be developed whenever there are more than two candidates or elements that are to be chosen. Arrow's theorem has been stated in many ways, as we note in the footnotes that accompany our own statement (see Arrow 1951; Scott and Antonsson 1999). A voting procedure can be characterized as fair if five axioms[2] are obeyed:

[1] As both described and practiced, the number of points awarded in pairwise comparisons is often nonuniform and subjectively or arbitrarily weighted. But it is quite important that the points awarded be measured in fixed increments.

[2] The version of Arrow's theorem presented here conforms to its original presentation. Arrow (1951) subsequently showed that axioms 2 and 4 could be replaced by the "Pareto condition," which states that if everyone ranks *A* over *B*, the societal ranking has *A* ranked above *B*. Arrow's presentation also formally states that both individual and group rankings must be *weak orders*, that is, transitive orderings that include all alternatives and allow for indierence among alternatives.

1. *Unrestricted*: All conceivable rankings registered by individual voters are actually possible.
2. *No imposed orders*:[3] There is no pair A, B for which it is impossible for the group to select one over the other.
3. *No dictator*: The system does not allow one voter to impose his/her ranking as the group's aggregate ranking.
4. *Positive response*:[4] If a set of orders ranks A before B, and a second set of orders is identical to the first except that individuals who ranked B over A are allowed to switch, then A is still preferred to B in the second set of orders.
5. *Independence of irrelevant alternatives* (IIA): If the aggregate ranking would choose A over B when C is not considered, then it will not choose B over A when C is considered.

Arrow proved that at least one of these properties must be violated for problems of reasonable size (at least three voters expressing only ordinal preferences among more than two alternatives). Hazelrigg states Arrow's theorem informally by saying that "in general, we cannot write a utility function for a group" (Hazelrigg 1996). It is worth noting that a consistent social choice (voting) procedure can be achieved by violating any one of the five (or four) conditions. Indeed, in this paper we address the questions: Which axioms are violated by designers as they make sensible choices? and, What are the consequences of these "violations"? Dictatorship (axiom 3) and violations of the Pareto condition (axioms 2 and 4) are intuitively offensive. Further, Scott and Antonsson argue that engineering approaches that use quantified performance rankings do not violate axiom 5, since comparison of two alternatives on the basis of measurable criteria is independent of the performance of other alternatives. However, these approaches often violate axiom 1 since many theoretically possible orders are not admissible in practice, as many engineering criteria must be of the less-is-better, more-is-better, or nominal-is-best varieties (Scott and Antonsson 1999).

Saari (1995; D.G. Saari, personal communication, 2001) notes that some voting procedures based on pairwise comparisons are faulty since they can produce ranking results that offend our intuitive sense of a reasonable outcome. He further claims that virtually any final ranking can be arrived at by correct specification of the voting procedure. Saari also suggests that among pairwise comparison procedures, the Borda count most "respects the data" in that it avoids the counter-intuitive results that can arise with other methods. Indeed, Saari notes (D.G. Saari, personal communication, 2001) that the Borda count

> never elects the candidate which loses all pairwise elections . . . always ranks a candidate who wins all pairwise comparisons above the candidate who loses all such comparisons.

In the case of the Borda count, the fifth Arrow axiom, the *independence of irrelevant alternatives* (IIA) is violated. What does this mean for design? In principle, and in the conceptual design phase where these questions are often most relevant, the possible space of design options is infinite. Bounded rationality insists that the designer must find a way to limit that set of design alternatives, to make it a finite, relatively small set of options. As a process, then, design involves generating the design alternatives and selecting one (or more) of them. Options may be eliminated because they do not meet some goals or criteria, or because they are otherwise seen as poor designs. Given these two bases for selection, how important is IIA? Does it matter if we violate IIA? Are we likely to erroneously remove promising designs early in the process? Is our design process flawed because of these removed designs?

The violation of IIA leads to the possibility of *rank reversals*, that is, changes in order among n alternatives that may occur when one alternative is dropped from a once-ranked set before a second ranking of the remaining n–1 alternatives (Sect. 2). The elimination of designs or candidates *can* change the tabulated rankings of those designs or candidates that remain under consideration. The determination of which design is "best" or which candidate is "preferred most" may well be sensitive to the set of designs considered.

It is thought that these rank reversals occur because of a loss of information that occurs when an alternative is dropped or removed from the once-ranked set (D.G. Saari, personal communication, 2001). In addition, rank reversals occur when there are *Condorcet cycles* in the voting patterns: $[A \succ B \succ C,\ B \succ C \succ A,\ C \succ A \succ B]$. When aggregated over all voters and alternatives, these cycles cancel each other out because each option has the same Borda count. When one of the alternatives is removed, this cycle no longer cancels. Thus, removing C from the above cycle unbalances the Borda count between A and B, resulting in a unit gain for A that is propagated to the final ranking results.

Paralleling the role of the Borda count in voting procedures, the PCC is the most consistent pairwise procedure to apply when making design choices. Both implementations are better than "drop and revote," whether viewed from the standpoint of bounded rationality embedded in Simon's concept of "satisficing" (Simon 1996), or from Saari's analysis of voting procedures (Saari 1995): both say we should *consider all of the information we have*. We may not attain perfect rationality and complete knowledge, but we should proceed with the best available knowledge. Design iterates between generating options and selecting among them, with the richness of information increasing as the process proceeds. At each stage, design selection tools must operate at an appropriate information level – as more information is developed, more complex tools can be applied: decision and information value theory, demand modeling, etc. While these tools can overcome the IIA violations inherent to the Borda count, they do so at a cost. Selection actions could be delayed until design information is rich enough to apply techniques that will not violate IIA, but this would commit the designer to the added expense of further developing poor designs. Rather than "drop and revote," design is more akin to sequential runoff elections in which the (design) candidates continue to "debate" throughout the design selection process.

[3]Also called *citizen's sovereignty*. Pareto is equivalent to citizen's sovereignty plus positive response.

[4]Positive response is an ordinal version of *monotonicity*.

In the end, no selection method can overcome poor design option generation. However, the act of selection helps to clarify the design task. From a practical standpoint, both designers and teachers of design have found that pairwise comparisons appear to work well by focusing their attention, and by bringing order to large numbers of seemingly disparate objectives, attributes, or data points. In addition, these rankings often produce good designs.

We are interested in enabling and contributing to a positive discussion of improving methods of design decision making. In this spirit, we describe here a way to use pairwise comparisons in a structured approach that produces results that are identical to the accepted vote-counting standard, the Borda count. The method is a structured extension of pairwise comparisons to a pairwise comparison chart (PCC) or matrix (Dym and Little 1999). We show that the PCC produces consistent results quickly and efficiently, and that these results are identical with results produced by a Borda count. We illustrate this in two of the examples that have been used to show the inconsistencies produced by pairwise comparisons, and we provide a formal proof of the equivalence of the PCC and the Borda count.

2 Example 1

We begin with an example to highlight some of the problems of (non-Borda count) pairwise comparison procedures, and to suggest the equivalence of the Borda count with the PCC (the proof of equivalence is presented in Sect. 3).

Twelve designers are asked to rank order three designs: *A*, *B*, and *C*. In doing so, the 12 designers have, collectively, produced the following sets of orderings (Saari 2001):

$$\begin{array}{llll} 1 \text{ preferred} & A \succ B \succ C & 4 \text{ preferred} & B \succ C \succ A \\ 4 \text{ preferred} & A \succ C \succ B & 3 \text{ preferred} & C \succ B \succ A \end{array} \tag{1}$$

Saari has shown that pairwise comparisons other than the Borda count can lead to inconsistent results for this case (Saari 2001). For example, in a widely used plurality voting process called *the best of the best*, *A* gets 5 first-place votes, while *B* and *C* each get 3 and 4, respectively. Thus, *A* is a clear winner. On the other hand, in an "antiplurality" procedure characterized as *avoid the worst of the worst*, *C* gets only 1 last-place vote, while *A* and *B* get 8 and 4, respectively. Thus, under these rules, *C* could be regarded as the winner. In an iterative process based on the best of the best, if *C* were eliminated for coming in last, then a comparison of the remaining pair *A* and *B* quickly shows that *B* is the winner:

$$\begin{array}{llll} 1 \text{ preferred} & A \succ B & 4 \text{ preferred} & B \succ A \\ 4 \text{ preferred} & A \succ B & 3 \text{ preferred} & B \succ A \end{array} \tag{2}$$

On the other hand, a Borda count would produce a clear result (Saari 2001). The Borda count procedure assigns numerical ratings separated by a constant to each element in the list. Thus, sets such as (3,2,1), (2,1,0) and (10,5,0) could be used to rank a three-element list. If we use (2,1,0) for the rankings presented in Eq. (1), we find total vote counts of (*A*: 2+8+0+0=10), (*B*: 1+0+8+3=12) and (*C*: 0+4+4+6=14), which clearly shows that *C* is the winner. Furthermore, if *A* is eliminated and *C* is compared only to *B* in a second Borda count,

$$\begin{array}{llll} 1 \text{ preferred} & B \succ C & 4 \text{ preferred} & B \succ C \\ 4 \text{ preferred} & C \succ B & 3 \text{ preferred} & C \succ B \end{array} \tag{3}$$

then *C* remains the winner, as it also would here by a simple vote count. It must be remarked that this consistency cannot be guaranteed, as the Borda count violates the IIA axiom.

We now make the same comparisons in a PCC matrix (Table 1). As noted above, a point is awarded to the winner in each pairwise comparison, and then the points earned by each alternative are summed. In the PCC of Table 1, points are awarded row by row, proceeding along each row while comparing the row element to each column alternative in an individual pairwise comparison. This PCC result shows that the rank ordering of preferred designs is entirely consistent with the Borda results just obtained:

$$C \succ B \succ A \tag{4}$$

The PCC matrix exhibits a special kind of symmetry, as does the ordering in the "Win" column (largest number of points) and the "Lose" row (smallest number of points): the sum of corresponding off-diagonal elements $X_{ij}+X_{ji}$ is a constant equal to the number of comparison sets.

We have noted that a principal complaint about some pairwise comparisons is that they lead to rank reversals when the field of candidate elements is reduced by removing the lowest-ranked element between orderings. Strictly speaking, rank reversal can occur when any alternative is removed. In fact, and as we note further in Sect. 4, examples can be constructed to achieve a specific rank reversal outcome (Saari 2001). Such examples usually include a dominated option that is not the worst. Also, rank reversals are possible if new options are *added*.

Practical experience suggests that the PCC generally preserves the original rankings if one alternative is dropped. In example 1, if element *A* is removed and a two-element runoff is conducted for *B* and *C*, we find the results given in Table 2. Hence, once again we find

Table 1. Pairwise comparison chart (PCC) for example 1

Win/Lose	*A*	*B*	*C*	Sum/Win
A	–	1+4+0+0	1+4+0+0	10
B	0+0+4+3	–	1+0+4+0	12
C	0+0+4+3	0+4+0+3	–	14
Sum/Lose	14	12	10	–

Table 2. Reduced pairwise comparison chart (PCC) for example 1 wherein the "loser" *A* in the first ranking is removed from consideration

Win/Lose	*B*	*C*	Sum/Win
B	–	1+0+4+0	5
C	0+4+0+3	–	7
Sum/Lose	7	5	–

$$C \succ B \tag{5}$$

The results in Eq. (5) clearly preserve the ordering of Eq. (4), that is, no rank reversal is obtained as a result of applying the PCC approach. In those instances where some rank reversal does occur (see example 2, Sect. 4), it is often among lower-ranked elements where the information is strongly influenced by the removed element (which we will see and explain in example 2).

3
Proof of the PCC–Borda count equivalence

The PCC is an implementation of the Borda count. In both procedures, the number of times that an element outranks another in pairwise comparisons is tallied to determine a final overall ranking. More formally, we prove here that these methods are identical, always producing the same rank order for a given collection of individual orderings.[5]

3.1
Preliminaries

Let us suppose that a set of n alternatives

$$\{A_1, A_2, \ldots A_n\} \tag{6}$$

is ranked individually m times. Each rank order R_i takes the form

$$A_{i_1} \succ A_{i_2} \succ \cdots \succ A_{i_n} \tag{7}$$

where $A \succ B$ indicates that A outranks or is ranked ahead of B. Each rank order R_i can then be expressed as a permutation σ_i of (1,2,...,n):

$$\sigma_i = (i_1, i_2, \ldots, i_n) \tag{8}$$

Let σ_{ij} be the jth entry of σ_i, so $\sigma_{ij}=i_j$. Let $\sigma_i(k)$ be the index of the entry with value k in σ_i (for k=1,2,...,n). Then:

$$\sigma_i(\sigma_{ij}) = j \tag{9}$$

Then $\sigma_i(k)$ is equal to the ordinal position that alternative A_k holds in the rank order σ_i. To take an example with n=3, if R_i expresses the ranking

$$A_3 \succ A_1 \succ A_2 \tag{10}$$

that is, if σ_i=(3,1,2), then $\sigma_i(1)=2$, $\sigma_i(2)=3$, and $\sigma_i(3)=1$.

3.2
Borda count sums

In a Borda count, each alternative A_k is assigned a number of points for each individual rank order R_i depending on its place in that order, and then the numbers are summed. Although there is an infinite number of equivalent numbering schemes, the canonical scheme, which may be used without loss of generality, assigns $(n-\sigma_i(k))$ points to alternative A_k from rank order R_i. For example, the rank ordering in Eq. (10) assigns two points to alternative A_3, one point to A_1, and no points to A_2. The Borda sum for the alternative A_k is obtained by summing over all individual orders R_i

$$A_k^{\mathrm{B}} = \sum_{i=1}^{m} (n - \sigma_i(k)) \tag{11}$$

3.3
Pairwise comparison chart (PCC) sums

To generate the kth row of the PCC, for each $j \neq k$ we count the number of permutations σ_i for which $\sigma_i(k)<\sigma_i(j)$, assigning one point for each such σ_i. Notice that $\sigma_i(k)<\sigma_i(j)$ if and only if A_k outranks A_j in R_i. For any σ_i, if $\sigma_i(k)<n$, then one point will be added to the A_k row in each of the columns $A_{\sigma_i(k)+1}, \ldots, A_n$. If $\sigma_i(k)=n$, no points are added to that row. Thus, the total points added to the A_k row as a result of R_i is $(n-\sigma_i(k))$. The grand total for A_k in the "sum/win" column is simply

$$A_k^{\mathrm{PCC:W}} = \sum_{i=1}^{m} (n - \sigma_i(k)) \tag{12}$$

which is exactly equal to the Borda sum given in Eq. (11). Therefore, the two methods are equivalent: the PCC is thus either an alternate representation of or a simple method for obtaining a Borda count (or vice versa).

Note that the sum for the "sum/lose" row in the PCC is just

$$A_k^{\mathrm{PCC:L}} = mn - \sum_{i=1}^{m} (n - \sigma_i(k)) \tag{13}$$

Therefore, the information contained in the "sum/lose" row is immediately available if the Borda count is known.

4
Example 2

Rank reversals sometimes occur when alternatives are dropped and the PCC procedure is repeated. We now show how such an example can be constructed.

Thirty designers (or consumers) are asked to rank order five designs A, B, C, D, and E as a result of which they produce the following sets of orderings:

$$\begin{aligned} &10 \text{ preferred } A \succ B \succ C \succ D \succ E \\ &10 \text{ preferred } B \succ C \succ D \succ E \succ A \\ &10 \text{ preferred } C \succ D \succ E \succ A \succ B \end{aligned} \tag{14}$$

Here too, the procedure chosen to rank order these five designs can decidedly influence or alter the results. For example, all of the designers ranked C and D ahead of E in the above tally. Nonetheless, if the following sequence of pairwise comparisons is undertaken, an inconsistent result obtains:

$$C \text{ vs } D \Rightarrow C; \quad C \text{ vs } B \Rightarrow B; \quad B \text{ vs } A \Rightarrow A; \quad A \text{ vs } E \Rightarrow E \tag{15}$$

If we construct a PCC matrix for this five-design example, we find the results shown in Table 3, and they clearly indicate the order of preferred designs to be

[5]An anonymous reviewer has suggested that a similar proof can be found in a different context in Zangemeister (1970) with the procedures identified under altogether different names.

Table 3. Pairwise comparison chart (PCC) for example 2

Win/Lose	*A*	*B*	*C*	*D*	*E*	Sum/Win
A	–	10+0+10	10+0+0	10+0+0	10+0+0	50
B	0+10+0	–	10+10+0	10+10+0	10+10+0	70
C	0+10+10	0+0+10	–	10+10+10	10+10+10	90
D	0+10+10	0+0+10	0+0+0	–	10+10+10	60
E	0+10+10	0+0+10	0+0+0	0+0+	–	30
Sum/Lose	70	50	30	60	90	–

Table 4. Borda count for example 2 using the weight set (4,3,2,1,0)

Element	Points
A	40+0+10=50
B	30+40+0=70
C	20+30+40=90
D	10+20+30=60
E	0+10+20=30

$$C \succ B \succ D \succ A \succ E \tag{16}$$

If the same data are subjected to a Borda count using the weights (4,3,2,1,0) for the place rankings, we then find the results displayed in Table 4. When we compare these results to the PCC results shown in Table 3, we see that the PCC has achieved the same Borda count results, albeit in a slightly different fashion.

What happens if we drop the lowest-ranked design and redo our assessment of alternatives? Here design *E* is least preferred, and we find the results shown in Table 5 if it is dropped. These results show a rank ordering of

$$C \succ B \succ A \succ D \tag{17}$$

Rank order is preserved here for the two top designs, *C* and *B*, while the last two change places. Why does this happen? Quite simply, because of the relative narrowness of the gap between *A* and *D* when compared to the gap between *A* and *E*, the two lowest ranked in the first application of the PCC in this example.

It is also useful to "reverse engineer" this example. Evidently it was constructed by taking a Condorcet cycle [$A \succ B \succ C$, $B \succ C \succ A$, $C \succ A \succ B$] and replacing *C* with an ordered set ($C \succ D \succ E$) that introduces two dominated (by *C*) options that are irrelevant by inspection. Removing only *E* produced a minor rank reversal of the last two alternatives, *A* and *D*. Removing only *D*, the third-best option, produces the same result among *A*, *B*, and *C* as removing *E*, although without creating a rank reversal. Removing both *D* and *E* produces a tie among *A*, *B*, and *C*.

In a design context, assuming that designs *D* and *E* are always inferior to design *C*, they would seem to be dominated members of the same basic design family. Thus, in order to avoid these (minor) rank reversals, it is important to group designs into similar families, pick the best, and then use PCC to rank the best across families. In other words, we need to be sure that we are not evaluating inferior alternatives from one class of design along with the best options from that class and from other classes. This suggests that PCC should be applied hierarchically to avoid artificial spacing in the Borda count among design alternatives. In early design, it is too costly to acquire quantitative measures of performance that can indicate how much better one alternative is than another. By grouping alternatives into families, we can lessen the chance that alternatives that are actually quite close to each other in performance will appear far apart due to the sheer number of alternatives deemed to fall in the middle.

It is also worth noting here that rank reversals of any two alternatives can be "cooked" by adding enough irrelevant alternatives to a Borda count. This follows directly from the fact that the Borda count depends upon the number of alternatives between two alternatives, as does its PCC equivalent. Consider the following example: there are n+1 alternatives and m+1 voters. Alternative *A* is ranked first (n points) and alternative *B* last (0 points) by m voters, while the remaining voter casts *B* as second-to-last (1 point) and *A* as last (0 points). Thus, *A* has $m \times n$ points, and *B* has 1. It is clear that it does not really matter what the absolute rankings are: *A* has gotten n more points than *B* from m voters and B has gotten 1 *more* than *A* on the last criterion – as far apart as the two alternatives can be without having *A* dominate *B*. Suppose new alternatives are added. Any new alternative that is either better than both *A* and *B* or worse than both will not affect the ranking of *A* and *B*. However, if a new alternative falls between *A* and *B*, the relative ranking will change. Therefore, if we find $m \times n$ new alternatives that are more or less preferred than both *A* and *B* by the original m voters that favor *A*, but that fall between *B* and *A* for the last voter, we can change the aggregated scores to $m \times n$ for *A* and $(m \times n)+1$ for *B*. Thus, again we have changed the aggregate scores by (artificially) introducing a large number ($m \times n$) of irrelevant "ringers."

Perhaps one of the main points of all of the above discussion is that the tool that should be used to do

Table 5. Reduced pairwise comparison chart (PCC) for example 2 wherein the "loser" *E* in the first ranking is removed from consideration

Win/Lose	*A*	*B*	*C*	*D*	Sum/Win
A	–	10+0+10	10+0+0	10+0+0	40
B	0+10+0	–	10+10+0	10+10+0	50
C	0+10+10	0+0+10	–	10+10+10	60
D	0+10+10	0+0+10	0+0+0	–	30
Sum/Lose	50	40	30	60	–

ranking or to calculate aggregate demand depends on how much data is available, with what granularity, and on how much the data-gatherers are prepared to spend. Pairwise comparisons are cheap and require little detailed knowledge, and are thus valuable in conceptual design. Focusing on the best candidates or exemplars in a set introduces a certain granularity in the data that can help avoid IIA-induced rank reversals. Alternatives that fit an existing group do not earn a separate, distinguishable space in the PCC, and the spacing between different alternatives is less likely to be "padded" by alternatives that are actually quite close in performance.

5 Example 3

We now present an example that shows how pairwise ranking that does not consider other alternatives can lead to a result exactly opposite to a Borda count, which does consider other alternatives. It also indicates that attempting to select a single best alternative may be the wrong approach.

One hundred customers were "surveyed on their preferences" with respect to five mutually exclusive design alternatives *A*, *B*, *C*, *D*, and *E* (G.H. Hazelrigg, personal communication, 2001). The survey reported that "45 customers prefer *A*, 25 prefer *B*, 17 prefer *C*, 13 prefer *D*, and no one prefers E." These data seem to indicate that *A* is the preferred choice, and that *E* is entirely "off the table."

However, as reported, these results assume either that the customers are asked to list only one choice or, if asked to rank order all five designs, that only their first choices are abstracted from their rank orderings. Suppose that the 100 customers were asked for rankings and that those rankings are (G.H. Hazelrigg, personal communication, 2001)

$$\begin{aligned} &45 \text{ preferred } A \succ E \succ D \succ C \succ B \\ &25 \text{ preferred } B \succ E \succ D \succ C \succ A \\ &17 \text{ preferred } C \succ E \succ D \succ B \succ A \\ &13 \text{ preferred } D \succ E \succ C \succ B \succ A \end{aligned} \tag{18}$$

Again, the procedure used to choose among the rank orderings of these five designs can decidedly influence or alter the results. For example, if *A* and *B* are compared as a (single) pair, *B* beats *A* by a margin of 55 to 45. And, continuing a sequence of pairwise comparisons, we can find that

$$A \text{ vs } B \Rightarrow B; \quad B \text{ vs } C \Rightarrow C; \quad C \text{ vs } D \Rightarrow D; \quad D \text{ vs } E \Rightarrow E \tag{19}$$

Proposition (19) provides an entirely different outcome, one that is not at all apparent from the vote count originally reported. How do we sort out this apparent conflict?

We resolve this dilemma by constructing a PCC matrix for this five-product example (Table 6), whose results clearly indicate the order of preferred designs to be

$$E \succ D \succ A \succ C \succ B \tag{20}$$

A Borda count of the same data of Eq. (18), using the weights (4,3,2,1,0) for the place rankings, confirms the PCC results, with the Borda count numbers being identical to those in the "win" column of the PCC in Table 6, that is,

$$E(300) \succ D(226) \succ A(180) \succ C(164) \succ B(130) \tag{21}$$

In this case, removing *B* and revoting generates a relatively unimportant rank reversal between *A* and *C*, thus demonstrating the meaning of IIA and showing that dropping information can have consequences.

This example is one where the "best option" as revealed by the PCC/Borda count is not the most preferred by anyone. Is the PCC lying to us? In a real market situation, where all five options are available, none of the surveyed customers would buy *E*. Two explanations for this survey come to mind. First, this data could have been collected across too broad a spectrum of customers in a segmented market in which design *E* is something of a "common denominator"; the other four designs respond better to four disparate market "niches." Under this explanation, there is really no "best design," although *E* seems to be a good starting point from which to search. Unfortunately, there is also no identifiable "worst design," although one could also argue that *E* is the "worst."

A second explanation is that these designs are all extremely close to each other in performance, so that small variations in performance have translated into large differences in the PCC. If this is the case, a designer might try to generate new design options by better merging the apparent desires of consumers. Methods such as the *House of Quality* require that designs be ranked along several significant (and possibly linguistic or nonquantifiable) performance criteria (Pugh 1990; Ulrich and Eppinger 2000). The goal in such a process shifts from selecting the "best" design to identifying the characteristics of a composite winning design. Of course, there is no guarantee that such a winning composite design exists, but PCCs can help the ranking process that might lead to its generation.

Both of the above explanations point to the need to integrate the PCC into a hierarchy of design decision methods. Deciding just when the PCC should give way to more information-rich methods is perhaps the main problem in this task. The PCC calculated for example 3

Table 6. Pairwise comparison chart (PCC) for example 3

Win/Lose	*A*	*B*	*C*	*D*	*E*	Sum/Win
A	–	45+0+0+0	45+0+0+0	45+0+0+0	45+0+0+0	180
B	0+25+17+13	–	0+25+0+0	0+25+0+0	0+25+0+0	130
C	0+25+17+13	45+0+17+13	–	0+0+17+0	0+0+17+0	164
D	0+25+17+13	45+0+17+13	45+25+0+13	–	0+0+0+13	226
E	0+25+17+13	45+0+17+13	45+25+0+13	45+25+17+0	–	300
Sum/Lose	220	270	236	174	100	–

shows strong support for option *E*, yet we have argued that more information should be developed before a design is selected. Inconclusive results generated by the PCC are generally easy to detect and can be corrected by moving to a more detailed selection method. While such graceful degradation of performance is typical of the PCC in practice, the above example, unfortunately, is of a case in which the PCC yields clear selection results at a point where more detailed selection procedures might be more appropriate.

6 Conclusions

This paper demonstrates that effective decision making is possible in the practice of engineering design, notwithstanding concerns raised about pairwise comparisons and Arrow's impossibility theorem. The identification of the structured pairwise comparison chart (PCC) as an implementation of the well-known Borda count and its application to oft-cited "pathological" examples suggests several ideas. First, it is not the individual pairwise comparisons that lead to erroneous results. Rather, rank reversals and other infelicities result from serial aggregation of pairwise comparisons when "losing" alternatives are dropped from further consideration. Pairwise comparisons that are properly aggregated in a PCC produce results that are identical to the Borda count, a "unique positional procedure which should be trusted" (D.G. Saari, personal communication, 2001). Indeed, our proof that the PCC is identical to the Borda count confirms that it compensates for and removes the same inherent cancellations.

It is important to recall that, in practice, the PCC and similar methods are used very early in the design process, where rough ordinal rankings are used to bound the scope of further development work. The PCC is more of a discussion tool than a device intended to aggregate individual orderings of design team members into a "group" decision. Indeed, design students are routinely cautioned against over-interpreting or relying too heavily on small numerical differences (Dym and Little 1999). In voting, we usually end up with only one winner, and any winner *must* be one of the entrants in the contest. In early design, it is perfectly fine to keep two or more winners around, and the ultimate winner often does not appear on the initial ballot. Indeed, it is often suggested that designers look at *all* of the design alternatives and try to incorporate the good points of each to create an improved, composite design (Dym and Little 1999; Pahl and Beitz 1996; Pugh 1990; Ulrich and Eppinger 2000). In this framework, the PCC is a useful aid for understanding the strengths and weaknesses of individual design alternatives, holistically or along more detailed performance criteria.

PCC can be used not only to rank designs but also to order design criteria by importance. This information helps to structure other design selection methods (e.g. Pugh concept selection, Pugh 1990), showing the design team where comparative differences among candidate designs are most important. This emphasis on *team* is significant. PCCs that implement the Borda count by having individuals vote in the pairwise comparisons are useful in the design process. However, they are most useful for encouraging student design teams to work on designs *as a team*. True collaboration takes place when team members must reach consensus on each comparison. The discussion necessary to reach this consensus helps to foster the shared understanding that is so important for good design. This collaborative approach might not be relevant to a social choice framework. In design and design education, however, where we are encouraged (and able) to improve design alternatives midstream, fostering constructive discussion is a significant reason for using any structured design approach. Thus, the matrix format of the PCC is perhaps a more useful tool in design education and design practice for conveying the same results obtained with the Borda count implemented as a piece of formal mathematics.

References

Arrow KJ (1951) Social choice and individual values, 1st edn. Wiley, New York

Dym CL, Little P (1999) Engineering design: a project-based introduction. Wiley, New York

Hazelrigg GH (1996) Systems engineering: an approach to information-based design. Prentice Hall, Upper Saddle River, New Jersey

Pahl G, Beitz W (1996) Engineering design: a systematic approach. Springer, London

Pugh S (1990) Total design: integrated methods for successful product engineering. Addison-Wesley, Wokingham, UK

Saari DG (1995) Basic geometry of voting. Springer, Berlin Heidelberg New York

Saari DG (2001) Decisions and elections: explaining the unexpected. Cambridge University Press, Cambridge

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

Simon HA (1996) The sciences of the artificial, 3rd edn. MIT Press, Cambridge, Mass

Ulrich KT, Eppinger SD (2000) Product design and development, 2nd edn. Irwin McGraw-Hill, Boston

Zangemeister C (1970) Nutzwertanalyse in der Systemtechnik; eine Methodik zurmultidimensionalen Bewertung und Auswahl vonProjektalternativen. Verlagskommission Wittemannsche Buchhandlung, Munich