FURTHER CONTRIBUTIONS TO THE THEORY OF PAIRED COMPARISONS[1]

By

M G. Kendall
Visiting Professor
Institute of Statistics
North Carolina State College

Institute of Statistics
Mimeograph Series No. 108
Limited Distribution

1. This research was supported by the United States Air Force, through the office of Scientific Research of the Air Research and Development Command.

FURTHER CONTRIBUTIONS TO THE THEORY OF PAIRED COMPARISONS[1]

by

M. G. Kendall

Visiting Professor, Institute of Statistics, North Carolina State College

1. When a pair of objects is presented for comparison and the two are placed in the relationship preferred: not-preferred, we have what is known as a paired comparison. A set of n objects can be compared, a pair at a time, in some or all of the possible $n(n-1)/2$ ways of choosing a pair, and the set of paired comparisons so derived gives us a picture of the inter-relationships of the objects under preference. A paired-comparison scheme is more general than a ranking; for with the latter A-preferred-to-B and B-preferred-to-C automatically ensures A-preferred-to-C, whereas with paired comparisons it might happen that C was preferred to A. The existence of these departures from the ranking situation may be due to various reasons, such as the fact that 'preference' is a complicated comparison being made with reference to several factors simultaneously; and one reason for using paired comparisons is to give such effects a chance to show themselves.

2. Situations often occur in which a set of m observers express preferences among n objects and we have to select that object, or perhaps that sub-set of objects, which are, in some sense, "most preferred." The simplest case is the one where there are only two objects, A and B, and

---

[1] This research was supported by the United States Air Force, through the Office of Scientific Research of the Air Research and Development Command.

every observer votes for either A or B as president of an institution. If 51 per cent of the votes are cast for A and 49 per cent for B we declare A elected. In doing so we have satisfied 51 per cent of the preferences but have had to proceed contrary to 49 per cent; we may say that 49 per cent of the preferences were violated. More generally, when we have to select a subset of the n objects as "elected" we shall in general, in the absence of complete unanimity, violate a number of preferences. Circumstances force us to do so to some extent. The problem is to do so to the least possible extent.

3. Consider the case in which 8 members of a body have to elect a committee of three from among themselves. We will suppose that no member votes for himself (though this makes no essential difference) and that there are no abstentions (though this too makes no essential difference). If the 8 members are represented by the letters A to G they might vote as follows:

Member	Members Preferred	
A	BDE	
B	DAF	
C	DGA	
D	CBE	
E	ABC	
F	ACD	
G	BAC	(1)

Here, for the moment, we suppose that there is no preference expressed among the triplets of members preferred; that is to say, A prefers B,D,E but does not say whether B is preferred to D or E, or D to E. He might then have written down his nominees in any order.

Under this system each elector expresses 9 preferences. A, for example, says, in effect, that he prefers B to C, F, and G, prefers D to C, F, and G, and prefers E to C, F, and G. There are thus 63 preferences altogether. We will represent this scheme in a two-way array of the following kind:

	A	B	C	D	E	F	G	No. of preferences
A	—	11		111	1111	111	111	15
B	1	—	1	11	1	1111	111	12
C	1	1	—	111	111	111	1111	15
D		11	1	—	11	11	11	9
E	1		1		—	11	11	6
F				1	1	—	1	3
G		1			1	1	—	3
Totals	3	6	3	9	12	15	15	63

(2)

Here, if A is preferred to B (a relationship we shall henceforward write as A pref. B or A → B) we write a unit in the row A, column B. For example C prefers D, G, A to each of B, E, F. We therefore have units in row D, Col. B; row D, Col. E; row D, Col. F; row G, Col. B; row G, Col. E; row G, Col. F; row A, Col. B; row A, Col. E; row A, Col. F. The totality of preferences expressed in (1) is given in the array (2), together with row and column totals.

Notice that: (a) the sum of row and column totals for each letter is 18. This provides a check. The reason is that each of the letters is compared with three others by each of six observers, so that each letter has 18 preferences (one way or the other).

(b) each column or row total is a multiple of three; for if any letter is preferred at all by an observer it is preferred to three others.

4. From the array (2) we see that A and C had 15 preferences each. If all preferences expressed by all observers have equal weight there is nothing to choose between them. B comes next with 12 preferences. All the others have fewer. Thus, if we have to elect three out of the seven to form a committee, we elect A, B and C. In doing so we satisfy as many preferences as possible; and since the total number of preferences is constant we minimize the number of violated preferences.

5. The procedure we have followed exhibits the structure of the preference scheme most clearly, but for the purposes of electing a committee of three we can proceed much more expeditiously. In fact, from array (1) we see that the voting is as follows:

Member	Number of votes
A	5
B	4
C	5
D	3
E	2
F	1
G	1
	21

(3)

A comparison of this with (2) shows that in the latter the row totals are thrice the number of votes. The reason is easy to see, for if any letter gets a vote it is thereby preferred to three others. In short, our procedure of electing those letters which get the greatest number of votes is equivalent to a method of minimizing violated preferences.

6. Now let us suppose that the rules of election are altered slightly and that each elector writes down the three members he prefers <u>in order of preference</u>. Such an order might be that of array (1) where, for example, A gives B his first preference, D his second and E his third. Each elector now expresses 12 preferences, three among the set he names and 9 by implication between those three and the three he omits. If we now form an array of preferences we get, instead of (2)

	A	B	C	D	E	F	G	Totals
A	--	3	3	4	4	4	3	21
B	2	--	3	3	3	4	3	18
C	2	2	--	4	4	4	4	20
D	1	2	1	--	3	2	3	12
E	1	0	1	0	--	2	2	6
F	0	0	0	1	1	--	1	3
G	1	1	0	0	1	1	--	4
Totals	7	8	8	12	16	17	16	84

(4)

The antisymmetry of the table has now been lost and row or column totals are no longer divisible by three. But we could still pick out the three members with the greatest number of preferences (A, C, B as before) without constructing a full table. In fact from (1) we score for A the following preferences allotted by the electors B to G:

$$4 + 3 + 0 + 5 + 5 + 4 = 21$$

and so for the other letters. The scores are the preference totals in the final column of (4).

7. The same method can obviously be applied to any number of voters and any size of committee. Under the condition that there are no abstentions and that nobody votes for himself, the total number of preferences expressed by m voters for a committee of n (no preferences between committee nominees) is $mn(m - n - 1)$; or if preferences are expressed by ranking nominees, is $mn(m - n/2 - 3/2)$. We may now, if we wish, relax some of the conditions without affecting essentials.

(a) If every man is allowed to vote for himself nothing new is introduced so long as we adhere to the principle of giving each preference the same weight;

(b) The same principles apply when a number of electors express preferences concerning a group of individuals who are not members of themselves. If m judges express preferences for k out of n objects (without ordering them) the number of preferences is $mk(n - k)$.

(c) If there are any abstentions we can continue as before to count those preferences which are expressed. Suppose, for example, that instead of (1) we had the following preferences expressed (second column):

Member	Preferences	Corrected Preferences	
A	B D E	B D E	
B	C A	C A	
C	D G A B	D G A	
D	C B E	C B E	
E	A B	A B	
F	A C D	A C D	
G	B B B	B	(5)

We suppose that these are in order. Member C has overstepped the mark. Unless we reject his ballot as spoiled we delete B from his ordering. Member B prefers C to A and both to the other four, but cannot express a preference between those other four and hence submits only two names. Member G tries to "plump" but we disallow this and count his expression as a preference for B only. We now have the preferences in the third column of (5) giving the following:

Preferences for

A		4	+3		+5	+5		= 17	
B	5			+4	+4		+5	= 18	
C		5		+5		+4		= 14	
D	4		+5			+3		= 12	
E	3			+3				= 6	
F								= 0	
G			4					= 4	(6)
								71	

A, B and C are still elected but B now gains more preferences than A.

We notice that election on this principle maximizes the number of satisfied preferences and minimizes the violated preferences, as before.

(d) If any voter "ties" certain nominees, this is equivalent to expressing no preference between them and everything proceeds as before. For example, if in (5) member D tied C, B, E there would be two fewer preferences for C and one fewer preference for B in (6).

(e) In particular this method covers the case when each of a set of judges ranks all the objects, and not merely a preferred sub-set of them. The whole method, in fact, is very flexible in this respect. So long as any preferences are expressed we can pursue the same technique. The only thing to take particular care about is that one judge has the same opportunities as another for expressing the same number of preferences, even though he may not avail himself of them. We clearly introduce bias if we give one judge a chance to express two preferences and another only one. The system proposed is in accordance with the best democratic principles in that each judge has the same number of votes, and all votes have the same weight.

(f) It is possible to order the members, according to the number of preferences allotted to them, in a ranking (which may itself contain tied members). Thus we constrain a paired-comparison system into a ranking at the expense of violating a number of preferences. The fewer the violations the nearer the scheme to an actual ranking. In tables of the type of (2) or (4) a perfect and unanimous ranking would correspond to a situation in which all the non-zero cells were above the main diagonal.

(g) In those cases where we choose to regard any object as compared with itself, as for example if we wish to complete the diagonals in (2), we may allot 1/2 to the cell in the same row and column. This will clearly not affect the order of the objects according to numbers of preferences received, for each object then receives an extra 1/2 for each observer.

(h) Likewise, if an observer cannot express a preference between a given pair A, B we may allot 1/2 to each of the cells in row A, column B and row B, column A in arrays of type (2).

(i) We can, if we wish, give effect to differences in reliability between judges. For example, if in array (2) we regard D as twice as important in his preferences as the others, we enter 2 for each preference instead of unity in the table.

8. Finally, let us note that the number of preferences can be used to calculate a coefficient of agreement among judges. This is another aspect of the coefficient of agreement in paired comparisons proposed by Babington Smith and myself some years ago. (See my Advanced Theory of Statistics, vol. 1, chapter 16). In fact if the total possible number of agreements is N and the actual number of agreements is M, the coefficient of agreement would be simply $2M/N - 1$ which varies from $-1/m$ or $-1/(m - 1)$ to 1. In table (2) for example the cells (A, B) and (B, A) have respectively 2 and 1 members. The pair A, B are compared three times and of these comparisons two are in agreement; there is thus one agreement out of a possible 3; likewise for AG, there are three agreements, each in the all AG, out of a possible 3. For the whole table it will be found that there are 47 agreements out of a possible 74 and the coefficient of agreement is 0.270.

9. We may also use the table to calculate a coefficient of departure from the ranking situation. Suppose we arrange the table so that rows and columns follow the order of the number of preferences expressed; in the case of table (2) this merely amounts to interchanging the rows and columns

corresponding to B and C. The number of units below the diagonal is then 13 and that above the diagonal is 50. No other arrangement of rows and columns can divide the 63 preferences so unequally. If all were above the diagonal the preferences would be consistent with a ranking. We might then take as our measurement of departure from the ranking situation the coefficient (13/63) x 2 = 0.413. We have multiplied the factor 13/63 by two because the furthest situation from ranking occurs when one half of the total preferences are allotted to the cells below the diagonal.

10. So much for the elements of the subject. I now proceed to consider sundry developments which are necessary to enable a more penetrating study of a paired-comparison situation to be made. The first arises from the nature of paired comparisonsin themselves and may best be introduced by an example.

Let us suppose that six players A to F are engaged in a chess tournament in which each plays the other once. The set of scores (1 for a win, 1/2 for a draw and 0 for a loss) then represents a set of paired comparisons made in all possible ways between them. We assume that all games reach a decision so that there are no missing values. A possible set of results is as follows:

	A	B	C	D	E	F	Total score
A	1/2	1	1	0	1	1	4 1/2
B	0	1/2	0	1	1	0	2 1/2
C	0	1	1/2	1	1	1	4 1/2
D	1	0	0	1/2	0	0	1 1/2
E	0	0	0	1	1/2	1	2 1/2
F	0	1	0	1	0	1/2	2 1/2

(7)

The simple way of arranging the competitors in order of success is to add up their scores, as is done in the final column. If we had three prizes we should divide the first and second between A and C and divide the third among B, E and F. Only D does not qualify for a share of the prize money. This situation, as we have seen, minimizes the violated "preferences." Such a procedure would be adopted in most tournaments of the kind.

11. But we now notice one rather anomalous effect. D, the only player to receive nothing, has in fact beaten one of the winners, A. We are not allowed to dismiss this as a mere fluke, because all preferences are equally valid. Furthermore A has beaten C but is nevertheless ranked with him. Vague but genuine feelings for general equity lead us to inquire whether something should not and cannot be done to restore the balance. Such a method was suggested by Dr. T. H. Wei (1952) in an unpublished thesis successfully submitted to the University of Cambridge for the Ph.D. degree. In effect Wei's procedure amounts to this:

We recalculate a score for each player by giving him the score of every player he has beaten and half the score of every player with whom he has drawn. This leads to the following new scores:

$$
\begin{aligned}
A &= \tfrac{1}{2}(4\tfrac{1}{2}) + 2\tfrac{1}{2} + 4\tfrac{1}{2} + 0 + 2\tfrac{1}{2} + 2\tfrac{1}{2} = 14\tfrac{1}{4}\\
B &= 0 + \tfrac{1}{2}(2\tfrac{1}{2}) + 0 + 1\tfrac{1}{2} + 2\tfrac{1}{2} + 0 = 5\tfrac{1}{4}\\
C &= 0 + 2\tfrac{1}{2} + \tfrac{1}{2}(4\tfrac{1}{2}) + 1\tfrac{1}{2} + 2\tfrac{1}{2} + 2\tfrac{1}{2} = 11\tfrac{1}{4}\\
D &= 4\tfrac{1}{2} + 0 + 0 + \tfrac{1}{2}(1\tfrac{1}{2}) + 0 + 0 = 5\tfrac{1}{4}\\
E &= 0 + 0 + 0 + 1\tfrac{1}{2} + \tfrac{1}{2}(2\tfrac{1}{2}) + 2\tfrac{1}{2} = 5\tfrac{1}{4}\\
F &= 0 + 2\tfrac{1}{2} + 0 + 1\tfrac{1}{2} + 0 + \tfrac{1}{2}(2\tfrac{1}{2}) = 5\tfrac{1}{4}
\end{aligned}
\qquad (8)
$$

We now arrange the players in order of new scores; and we now notice that A and C have separated, A being first and C second, while D has moved up to equality with B, E, and F.

This is as far as one would wish to go on practical grounds, perhaps, but now a further point raises itself. We have re-allocated the scores once. Why not do so again? If we re-allocate the scores of (8) by the same method we find

$$
\begin{array}{llllllllllllll}
A & \frac{1}{2}(14\frac{1}{4}) & + & 5\frac{1}{4} & + & 11\frac{1}{4} & + & 0 & + & 5\frac{1}{4} & + & 5\frac{1}{4} & = & 34\frac{1}{8} \\
B & 0 & + & \frac{1}{2}(5\frac{1}{4}) & + & 0 & + & 5\frac{1}{4} & + & 5\frac{1}{4} & + & 0 & = & 13\frac{1}{8} \\
C & & & & & & & & & & & & = & 26\frac{5}{8} \\
D & & & & & & & & & & & & = & 16\frac{7}{8} \\
E & & & & & & & & & & & & = & 13\frac{1}{8} \\
F & & & & & & & & & & & & = & 13\frac{1}{8}
\end{array}
\qquad (9)
$$

A and C are still first and second but D takes third place and B, E, F share the fourth position.

If we reallocate the scores once more we find scores

A	824.375
B	365.625
C	695.625
D	425.625
E	365.625
F	365.625

(10)

The order is now the same as we derived from (9); and if we ascertain new scores on the same principle we shall find that no new ordering has appeared. Later I shall prove that after a time the situation always "settles down" in this way.

13. There are two interesting features of this procedure. Let us revert to the preference scheme of (7) and regard the scores as a matrix. If we square this matrix we obtain

$$\begin{pmatrix} \frac{1}{4} & 3 & 1 & 4 & 2 & 3 \\ 1 & \frac{1}{4} & 0 & 2 & 1 & 1 \\ 1 & 2 & \frac{1}{4} & 4 & 2 & 2 \\ 1 & 1 & 1 & \frac{1}{4} & 1 & 1 \\ 1 & 1 & 0 & 2 & \frac{1}{4} & 1 \\ 1 & 1 & 0 & 2 & 1 & \frac{1}{4} \end{pmatrix} \qquad \begin{matrix} \underline{\text{Row totals}} \\ 14\frac{1}{4} \\ 5\frac{1}{4} \\ 11\frac{1}{4} \\ 5\frac{1}{4} \\ 5\frac{1}{4} \\ 5\frac{1}{4} \end{matrix} \qquad (11)$$

and the row totals are those previously obtained in (8) by the first re-allocation of scores. The reason for this will be obvious to anyone familiar with the rules of matrix multiplication and the result is generally true for all preference matrices. Furthermore, if we multiply (11) again by the matrix (7) and add row totals we shall get the scores of (9); and so on. The continual reallocation of scores is equivalent to taking successive powers of the matrix.

14. Let us now consider what interpretation can be given to the process in terms of comparisons. The following diagram shows the scheme of (7) in geometrical form

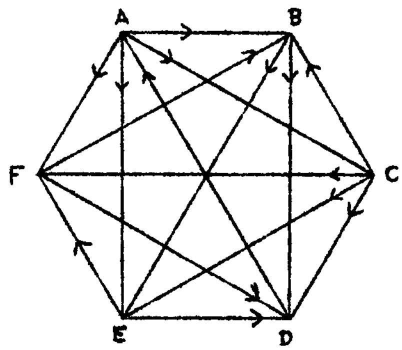

Figure 1 (12)

The six players are represented by the six vertices of a regular hexagon, which are joined by straight lines in all possible ways. If A pref. B we draw an arrow from A towards B. If no preference was expressed (or the game was drawn) we do not draw an arrow.

15. It will be seen that the score of any player in (7) is the number of arrows leaving his vertex, together with 1/2 (as the conventional score in the diagonal, when he is compared with himself) and 1/2 for any line passing through his vertex on which no arrow is drawn. When we proceed to the next stage we count the number of paths leaving the vertex and taking two steps. For example, for A we have the following paths leaving A and also leaving the vertex next visited:

ABD, ABE; ACB, ACD, ACE, ACF; AED, AEF; AFB, AFD.

There are ten of these "transitive" preferences. We also count the preference of B with itself, C with itself, etc., as 1/2 each, making a further score of 2; and finally we score 1/2 of 1/2 for the double preference of A with itself. The total score is $14\frac{1}{4}$, which is the score for A in (11). It may be verified that the same procedure gives the other scores in that array.

Similarly the scores obtained by the next reallocation, as given in (9), are the numbers of paths of three lines leaving the respective vertices, all arrows going the same way, with similar conventions about vertices taken with themselves; and so on. Our reallocation is equivalent to powering the matrix or to counting paths of transitive preferences of increasing extent.

16. From the geometrical viewpoint it is seen that in proceeding by reallocation we are extending our concept of comparison. We began by considering comparisons of pairs by themselves. When we proceed to the next stage we compare pairs which form part of triads; but we do not compare the triads by considering them as three pairs (which would bring us back to the first situation). Thus it is possible to "compare" A and C by the route $A \rightarrow B \rightarrow C$ or A and B by $A \rightarrow C \rightarrow B$. Both of these "comparisons" do not count in our score because they cannot both happen together; but either counts when it occurs.

17. Or we may put it another way by saying that we compare two members AB not directly, but through their comparisons with other members, e.g. by ACB, ADB, AEB and AFB. We choose the leading members in the final order so as to maximize the agreement with transitive preferences; or conversely, so as to minimize the violation of transitive preferences. Whether this is the right thing to do depends to some extent on practical circumstances. The process of continual reallocation has the advantage that it results in an objective final ordering; but whether this is what we want depends on whether we are considering a situation in which direct comparison is the basic generator of the data, or whether we wish to give scope for more reflective judgment in roundabout comparisons involving other members.

18. Let us now consider the case when several judges make paired comparisons, or several tournaments are played between the same set of players. For each observer we shall have a preference matrix of the type of (7). To obtain a composite picture, on the supposition that the judges are equally reliable, we superpose the matrices. Thus if (7) represents the preferences of a judge for 6 varieties of ice cream when offered to him in pairs, two additional judges might have the following preference matrices:

	A	B	C	D	E	F	Totals
A	$\frac{1}{2}$	1	0	1	1	0	$3\frac{1}{2}$
B	0	$\frac{1}{2}$	$\frac{1}{2}$	1	0	1	3
C	1	$\frac{1}{2}$	$\frac{1}{2}$	1	0	1	4
D	0	0	0	$\frac{1}{2}$	1	$\frac{1}{2}$	2
E	0	1	1	0	$\frac{1}{2}$	1	$3\frac{1}{2}$
F	1	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	2

(13)

	A	B	C	D	E	F	Totals
A	$\frac{1}{2}$	0	1	1	$\frac{1}{2}$	1	4
B	1	$\frac{1}{2}$	1	0	0	1	$3\frac{1}{2}$
C	0	0	$\frac{1}{2}$	1	1	1	$3\frac{1}{2}$
D	0	1	0	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$2\frac{1}{2}$
E	$\frac{1}{2}$	1	0	$\frac{1}{2}$	$\frac{1}{2}$	1	$3\frac{1}{2}$
F	0	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	1

(14)

Adding these and (7) together we get

	A	B	C	D	E	F	Totals
A	$1\frac{1}{2}$	2	2	2	$2\frac{1}{2}$	2	12
B	1	$1\frac{1}{2}$	$1\frac{1}{2}$	2	1	2	9
C	1	$1\frac{1}{2}$	$1\frac{1}{2}$	3	2	3	12
D	1	1	0	$1\frac{1}{2}$	$1\frac{1}{2}$	1	6
E	$\frac{1}{2}$	2	1	$1\frac{1}{2}$	$1\frac{1}{2}$	3	$9\frac{1}{2}$
F	1	1	0	2	0	$1\frac{1}{2}$	$5\frac{1}{2}$
Totals	6	9	6	12	$8\frac{1}{2}$	$12\frac{1}{2}$	54

(15)

On the basis of simple paired comparisons we should place A and C as bracketed equal, E as third, B as fourth, D as fifth and F as last.

19. The question now arises whether we should reallocate the scores by powering the matrix (15); or whether it would be preferable to power each matrix and then amalgamate the rankings at the end so as to minimize violated preferences. The two processes will not always lead to identical results, although in practice they should not differ very much. Arithmetically it is simpler to power just the one matrix (15), and in cases where there are many judges this would be almost decisive. This is the procedure I would recommend myself, but if there were any serious doubts I would perform the analysis both ways and compare the results. A wide disparity would, in my view, suggest that neither was very reliable. It would arise

mostly in cases where there were substantial disagreements among judges.

20. I now prove that the process of repeated powering does in fact converge to a limiting ranking. Dr. Wei offered a proof of the result for one observer and a complete set of preferences in his thesis.

First of all we define a matrix A of non-negative elements to be <u>indivisible</u> if it cannot be expressed in the form (by rearrangement of rows and columns)

$$A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix} \qquad (16)$$

If a preference matrix of type (15) is divisible in this sense the members of one block of objects are always preferred to every member of another. In such a case we divide the data into the two blocks and operate on each, finally ranking the members of the first group and then the members of the second. Similarly, if one of these blocks is itself divisible we divide it up; and so on. We clearly lose no generality by doing this, and divisibility is not a handicap in our preference situations.

21. I now require a theorem of Frobenius (cf. Wielandt, 1950[2]) which says that for indivisible matrices A with non-negative elements and positive elements in the diagonal there exists a unique simple positive root of the equation $|A - \lambda I| = 0$ which is greater than all other roots in absolute

---

[2] I am indebted to Sir Alexander Aitken and Dr. F. G. Foster for some references on this subject. The preference matrices are similar to, but not identical with, the matrices of transition probabilities studied in the theory of stationary stochastic processes.

value; and that the corresponding characteristic vector has all its elements of the same sign (which we may take to be positive).

Let λ_1 be this largest root and Y_1 the corresponding vector. Then if $\lambda_2, .. \lambda_p$ are the other roots and $Y_2 ... Y_p$ the corresponding vectors, and if P be the preference matrix, we have

$$PY = \Lambda Y \tag{17}$$

where Λ is the diagonal matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{pmatrix} \tag{18}$$

It is now easy to show that for any positive integer k

$$P^k Y = \Lambda^k Y \tag{19}$$

As the powering proceeds the major root λ_1 becomes dominant and (19) tends to the equation

$$P^k Y_1 = \lambda_1{}^k Y_1 \tag{20}$$

Thus from some k onwards the final ordering will be determined by the vector Y_1, which has non-negative elements.

22. We notice that the proof remains applicable to preference matrices in which some preferences may be missing, or when ties are present, provided that the matrix is not divisible. If any cell in a combined preference matrix contains no entries we insert a zero.

23. It is also of some interest to note that we may prove that the preference matrix is never singular. In fact, we can always express it (apart from positive numerical factors) in the form

$$(Q + U) \qquad (21)$$

where Q is an anti-symmetric matrix and U is the matrix all of whose elements are unity. For example (15), after division of rows by 1 1/2, can be expressed as U plus the matrix

$$Q = \begin{pmatrix} 0 & 1/3 & 1/3 & 1/3 & 2/3 & 1/3 \\ -1/3 & 0 & 0 & 1/3 & -1/3 & 1/3 \\ -1/3 & 0 & 0 & 1 & 1/3 & 1 \\ -1/3 & -1/3 & -1 & 0 & 0 & -1/3 \\ -2/3 & 1/3 & -1/3 & 0 & 0 & 1 \\ -1/3 & -1/3 & -1 & 1/3 & -1 & 0 \end{pmatrix} \qquad (22)$$

We reduce Q + U systematically by subtracting the first column from the second column, then the first row from the second row; then the first column from the third column, then the first row from the third row; and so on. The effect on Q is to reduce it to another antisymmetric matrix, say Q', and the effect on U is to reduce it to a unit in the top left-hand corner and zero elsewhere. Thus the determinant of Q + U is the determinant of Q' plus the determinant of the principal minor obtained by omitting the first row and column, which is also antisymmetric.

Now the determinant of p x p antisymmetric matrix is zero if p is odd and positive if p is even. Hence the determinant of Q + U is the sum of two components, one zero and the other positive; and hence it does not vanish.

22. In practice the number of paired comparisons arising from n objects may be inconveniently large and the question arises whether it is possible to economize in the number of comparisons made. In the example of the chess tournament which has been mentioned above (paragraph 10) if each player is to play every other, 15 games must be played. But only three can be conducted at once, so at best 5 sessions are necessary. If this is too long, and, say, three sessions are all that can be allowed, only nine games can be played and six have to be sacrificed. The question is, which six? Or again, if an individual is comparing items by taste testing, his patience or his palate may not endure the presentation of all the possible pairs, and a problem arises as to how best to cut down the number of pairs and which pairs to present.

23. Problems like this arise in many fields of experimentation and are usually dealt with by incomplete balanced blocks. Some new points, however, arise in paired-comparison work. Durbin (1951) has considered the use of Youden designs in ranking experiments. More recently Benard and van Elteren (1953) have discussed tests of significance where incomplete rankings are concerned. Without trying to exhaust the subject I proceed to consider the use of incomplete balanced blocks in preference schemes.

24. Consider first of all the case of a single observer. Of the $n(n-1)/2$ preferences which he could make we require to pick out a sub-set. Certain elementary principles of choice at once suggest themselves:

(a) every object should appear equally often. In this sense the design should be balanced;

(b) the preferences should not be divisible in the sense that we can split the objects into two sets and no comparison is made between any object in one and any object in the other.

In terms of preference matrices (a) means that there should be the same number of non-empty items in each row and column; (b) means that the matrix does not divide into two blocks and become of the form $\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}$ when the zeros represent empty cells. In terms of the preference diagram (a) means that there are the same number of paths direct between points leaving or entering each vertex and (b) means that the figure does not separate into two distinct polygons.

25. When possible I add a further condition of symmetry to the situation, that is to say

(c) In the preference diagram the number of paths of length ℓ proceeding from any point to any other point shall be the same for all pairs of points.

The length ℓ here means the number of lines traversed in the path, e.g. the path (in Figure 1, section 14) ABC from A to C is of length 2 and AEBDC from A to C is of length 4. Where no pair of objects is compared in these "partial" situations we omit the line between them. If they are joined by a line without an arrow this means that they have been compared but that no preference has been expressed.

In terms of preference matrices this condition implies a kind of symmetry of interlocking. A path ABC implies entries in row A, column B and column C (and the reflections column A, row B and row C); and analogous entries must occur in other rows in such a way that all the objects are symmetrically involved.

26. Under these conditions we can meet a requirement suggested to me in conversation by Dr. R. C. Bose: if all the preferences are exerted at random (e.g. if we toss up for it which of a pair shall be preferred) all possible final orderings of the objects produced by powering the matrix should be equally probable. This follows from the symmetry of the situation, for we can interchange two objects in the designs without altering the preference matrix, so far as concerns the underlying probabilities, and all final orders are therefore equally probable.

27. In a sense, it seems to me, condition (c) is necessary as well as sufficient for a proper design. If it is not obeyed certain objects become subject to different schemes of preference from others and their final positions are not determined on an unbiased basis. In terms of powered preference matrices, the sums of rows are not based on the same number of transitive comparisons of length ℓ.

28. The conditions laid down above impose certain restrictions on the scope of a paired-comparison experiment. For instance, if there are six objects and the numbers of entries in the rows of the preference matrix are equal, the number of comparisons necessary to obtain a balanced experiment must be a multiple of three. Anything else destroys the balance. The connectivity condition (b) further limits the freedom of choice; for example, with six objects at least six comparisons are required.

29. The setting up of incomplete designs is most easily thought of in terms of tours round the preference polygon. Consider the case n = 7. (Prime numbers are easier to deal with in most experimental designs.) There are 21 comparisons altogether. To obtain a balanced design we must have either 7 or 14 comparisons (or, of course, the full 21). The first 7 may,

without loss of generality, be taken as the tour ABC ... G round the preference heptagon. (No generality is lost because each member must be

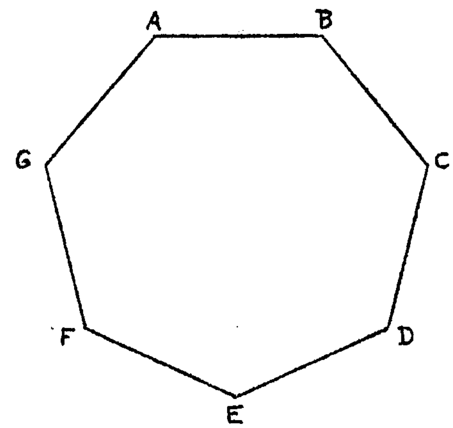

Figure 2

connected to two others and hence they be on a chain which may be taken to be the order A to G.) For the next 7 we have two possibilities: (a) start from A, miss a vertex and go to C, miss a vertex and go to E and so on; (b) start from A, miss two vertices and go to E, then two vertices and go to G and so on. We do not obtain new designs by tours missing three or more vertices because they are equivalent to (a) or (b). The two schemes are shown in Figure 3.

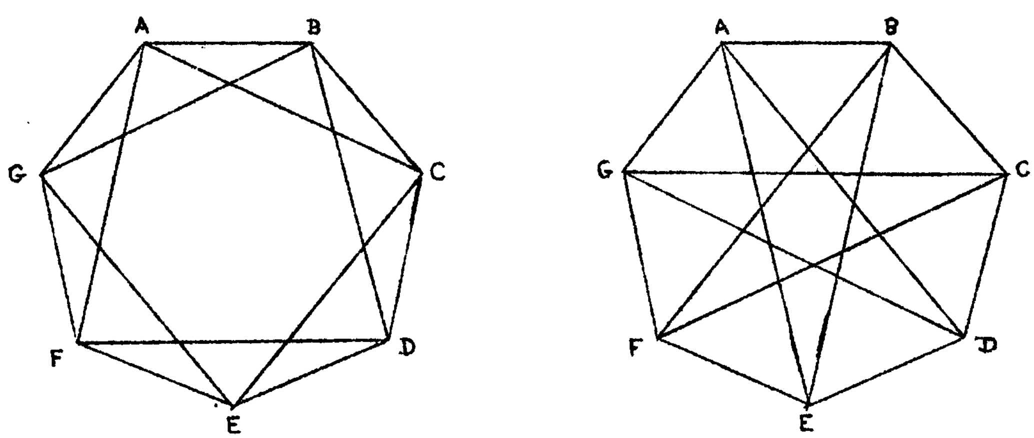

Figure 3

These schemes are not identical. In the former there are two triangular tours connecting any pair, e.g. ACB and AGB, whereas in the second there is only one, e.g. AEB. In terms of time taken in performance there is nothing to choose between them. For example if they represented a chess tournament, each round requires three games, one player having a bye, and for 14 games 5 rounds are required. Such might be

Scheme 1			Bye	Scheme 2			Bye	
AB,	CD,	EF	G	AB,	CD,	EF,	G	
AC,	BD,	EG	F	AD,	BC,	FG,	E	
BC,	DE,	FG	A	BE,	CF,	DG,	A	
AF,	CE,	BG	D	AE,	BF,	CG,	D	
AG,	DF		B, C, E	AG,	DE		B, C, F.	(23)

30. It remains to be considered whether one scheme is preferable to the other by some other criterion. There is nothing to choose between them in relation to balance or the application of the powered-matrix method. We note, however, that the patterns of transitive preferences are different. In the first any pair is connected by two triangles, three quadrilaterals, etc., in the second by one triangle, four quadrilaterals, etc. On the whole, I should be inclined to select the second design from a feeling that it has higher connectivity, but an exact criterion awaits further investigation.

31. When we have several judges, an obvious extension of symmetry requirements necessitates that each participates to an equivalent extent: in some sense the design should be balanced by judges as well as by comparisons. Something depends on whether we require to compare judges in

addition to objects. If so, each pair of judges must have certain comparisons in common. With two judges and seven objects, for example, one simple way would be to allot to each 14 comparisons, one judging according to each of the designs of Figure 3. They would then have 7 comparisons in common and all possible comparisons could be made.

32. I do not propose on this occasion to attempt a systematic exposition of the design problems involved in paired comparisons. Designs of an optimum kind which balance by numbers of comparisons, objects compared, numbers of observers on given comparisons and so forth are probably rather rare; and if something has to be sacrificed it depends on what is the point of major interest whether we sacrifice symmetry in comparisons or in judges. A final example will make clear a few of the principles involved.

Consider again the case of seven objects, A B C D E F G. There are three distinct tours round the preference polygon,

A B C D E F G

A C E G B D F

A D G C F B E (24)

Each tour involves seven comparisons and each object is compared with two others in a tour.

For a complete set of comparisons each observer would have to make 21. If this is felt to be too much we may allocate 14, consisting of two tours each. And if the tours are represented by a, b, c, we may allocate to the observers 1, 2, 3

1:	a, b
2:	b, c
3:	c, a

(25)

With these schemes every comparison is made equally often (twice); every tour is made equally often (twice); every observer makes the same number of comparisons (14); every observer has a tour in common with every other observer; and thus every observer can be compared with every other observer in respect of two comparisons involving any specified object.

If we have more than three observers, we take a number equal to a multiple of three and replicate the design.

Now suppose we had eleven objects, A to K. The full set of comparisons numbers 55. There are five distinct tours round the preference polygon

a :	A	B	C	D	E	F	G	H	I	J	K
b :	A	C	E	G	I	K	B	D	F	H	J
c :	A	D	G	J	B	E	H	K	C	F	I
d :	A	E	I	B	F	J	C	G	K	D	H
e :	A	F	K	E	J	D	I	C	H	B	G

(26)

Now if we try to allot two tours to each of five observers we lose symmetry; for there are 10 pairs of tours choosable from these five. We have, to preserve complete balance, to allot four tours to each observer 1, 2, 3, 4, 5

1 : b, c, d, e,

2 : c, d, e, a

3 : d, e, a, b

4 : e, a, b, c

5 : a, b, c, d (27)

Again the tours are balanced, but we have not achieved very much. Each observer now makes 44 comparisons, against the full set of 55.

We can sacrifice symmetry in several ways. We may, for instance, allot two tours to each observer, e.g.

1 : a, b

2 : b, c

3 : c, d

4 : d, e

5 : e, a (28)

Here every observer can be compared with two other observers but not every pair can be compared. Or if we have, say, 10 observers we may allot all the 10 possible pairs of tours one to each. Each observer then makes 22 comparisons and can be compared with four other observers. If 22 comparisons are still felt to be too many for one observer we may allocate the 55 preferences according to a linked design, e.g. (numbering the preferences 1 to 55) with 11 observers, 10 preferences each

1 :	1,	2,	3,	4,	5,	6,	7,	8,	9,	10	
2 :	1,	11,	12,	13,	14,	15,	16,	17,	18,	19	
3 :	2,	11,	20,	21,	22,	23,	24,	25,	26,	27	
4 :	3,	12,	20,	28,	29,	30,	31,	32,	33,	34	
5 :	4,	13,	21,	28,	35,	36,	37,	38,	39,	40	
6 :	5,	14,	22,	29,	35,	41,	42,	43,	44,	45	
7 :	6,	15,	23,	30,	36,	41,	46,	47,	48,	49	
8 :	7,	16,	24,	31,	37,	42,	46,	50,	51,	52	
9 :	8,	17,	25,	32,	38,	43,	47,	50,	53,	54	
10 :	9,	18,	26,	33,	39,	44,	48,	51,	53,	55	
11 :	10,	19,	27,	34,	40,	45,	49,	52,	54,	55	(29)

Here we have cut down the comparisons for each observer to 10 and each comparison is made twice. But we have lost a good deal of the comparison between judges; every judge can be compared with every other judge but only on one comparison of objects.

Benard, A. and Van Elteren, Ph. (1953), A generalization of the method of m rankings. Kon. Neder. Ak. van Weterschappen. A, 56, 358.

Durbin, J. (1951), Incomplete blocks in ranking experiments. Brit. Jour. Psych. 4, 85.

Frobenius, G. (1912), Über Matrizen aus nicht negativen Elementen. Sitz. Preuss. Akad. Wiss., 456

Wei, T. H. (1952), The algebraic foundations of ranking theory. Unpublished thesis, Cambridge University, England.

Wielandt, H. (1950), Unzerlegbare, nicht negative Matrizen. Math. Zeit., 52, 642.