

A new technique for ranking a large number of attributes

M. Jacobs

Management services Division, Unilever, Durban

M. Hedley

Aluman, Durban

In certain market research surveys, problems may be encountered when respondents are required to rank a large number of attributes. In this article, we propose a method of breaking down the attribute list into manageable groups, and present an algorithm for constructing an overall ranking. A step-by-step illustration of the algorithm is given. The performance of the algorithm in a practical situation is examined.

S. Afr. J. Bus. Mgmt. 1983, 14: 105 – 108

In sekere marknavorsings duik probleme op wanneer respondente 'n groot getal kenmerke in rangorde moet plaas. In hierdie artikel, word 'n nuwe metode voorgestel wat die kenmerke in beheerbare groepe lys en opbreek, en 'n algoritme vir die daarstelling van 'n algehele rangorde word voorgestel. 'n Stap-vir-stap illustrasie van die algoritme word gegee en die prestasie daarvan word in 'n praktiese voorbeeld ondersoek.

S.-Afr. Tydskr. Bedryfsl. 1983, 14: 105 – 108

In many market research applications, it is necessary for a respondent to rank a number of product attributes. Miller holds a view that the human mind is only capable of making correct judgments on up to six sets of stimuli at any one time.¹

'If the human observer is a reasonable kind of communication system, then when we increase the amount of input information, the transmitted information will increase at first, and will eventually level off at some asymptotic value.'

Hence, if the number of attributes to be ranked exceeds six, the complete set of ranking orders is likely to become inaccurate.

In this article, we propose a method of breaking down the list of attributes into manageable groups, and an algorithm to construct an overall ranking. The algorithm is easily programmed into a straightforward computer routine.

Firstly, the list of attributes should be broken down into n groups of m attributes each, where $n < m$ and $m < 7$. Where the list cannot be broken down conveniently, one or two attributes may be repeated to facilitate a convenient breakdown. Attributes should be assigned randomly to the groups.

For example, if there are 15 attributes to be ranked, these may be broken into three groups of five. If there are 19 attributes, one should be repeated to achieve a breakdown into four groups of five. In principle, it is possible to have groups of different sizes, but this may be confusing to the respondent.

Secondly, another breakdown of the attributes is performed. Judgment must be used to ensure that the second breakdown is materially different from the first. The following guidelines should be followed:

- (a) Each group in the second breakdown must contain at least one attribute from every group in the first breakdown.
- (b) If attributes were repeated to create a convenient grouping, then different attributes should be repeated in the second breakdown.
- (c) Attributes in groups of the first breakdown should be spread as evenly as possible over groups of the second breakdown.

The algorithm for attaining the overall ranking will be presented by means of an example. For the sake of simplicity, the example will be confined to a small number of attributes.

M. Jacobs*

Management Services Division, Unilever S.A., P.O. Box 3612, Durban 4000, Republic of South Africa

M. Hedley

Aluman. (Pty) Ltd., P.O. Box 1064, Durban 4000, Republic of South Africa

*To whom correspondence should be addressed

Suppose that eight attributes are to be ranked. They are broken into two groups of four. The arrangement of attributes and a sample response are given below:

	Attribute Ranking		Attribute Ranking		
(1)	A	3	(3)	C	1
	B	4		E	3
	C	1		A	4
	D	2		G	2
(2)	E	1	(4)	B	4
	F	4		H	2
	G	2		F	3
	H	3		D	1

An 8 × 8 cross table of attributes is drawn up. (Figure 1)

	A	B	C	D	E	F	G	H
A	X							
B		X						
C			X					
D				X				
E					X			
F						X		
G							X	
H								X

Figure 1 A blank cross table of attributes

From the first group of rankings, (1) above it is seen that attribute C is preferred to all others in the group, namely A, B and D. We can say, therefore, that 'C exceeds A' and write $C > A$. Also $C > B$ and $C > D$. The exact meaning of 'exceeds' will depend on the type of attribute being ranked. The term will be used for generality.

The relationship $C > A$ is expressed in the cross table by writing a 1 in row C column A. The other relationships determined by group (1) are similarly entered in the cross table. (Figure 2).

Similarly, all relationships determined by rankings in groups (2), (3) and (4) are entered into Figure 3.

The next stage consists of checking for logical inconsistencies in the rankings. For example, in group (2), attribute E exceeds G, while in group (3) G exceeds E. From this, we can only conclude that the respondent is not really sure which attribute he prefers. Hence, there is no preference between E and G, and both ones in row E column G and in row G column E are eliminated.

In general, an inconsistency is expressed by ones in both (i, j) and (j, i) positions in the cross table. It is removed by setting both positions to zero. (Figure 4).

These inconsistencies may prove valuable in themselves. For example, in pretesting a questionnaire, if certain pairs of attributes are regularly inconsistent, this might indicate

	A	B	C	D	E	F	G	H
A	X	1						
B		X						
C	1	1	X	1				
D	1	1		X				
E					X			
F						X		
G							X	
H								X

Figure 2 Relationships determined by group (1)

	A	B	C	D	E	F	G	H
A	X	1						
B		X						
C	1	1	X	1	1		1	
D	1	1		X		1		1
E	1				X	1	1	1
F		1				X		
G	1				1	1	X	1
H		1				1		X

Figure 3 All primary relationships in groups (1) (2), (3) and (4)

	A	B	C	D	E	F	G	H
A	X	1						
B		X						
C	1	1	X	1	1		1	
D	1	1		X		1		1
E	1				X	1		1
F		1				X		
G	1					1	X	1
H		1				1		X

Figure 4 Removal of inconsistencies

that the questionnaire is ambiguous with regard to these attributes, or that the attributes are not clearly differentiated in the minds of the respondents.

It is recommended that the number of inconsistencies occurring in each individual questionnaire be recorded. If the number exceeds a predetermined limit, this would indicate that the respondent is either confused or unqualified to contribute meaningfully to the survey. The particular questionnaire should thus be discarded.

Continuing with the example, we now assume that the relationships between attributes are transitive, that is, if $a_1 > a_2$ and $a_2 > a_3$ then $a_1 > a_3$.

From the cross table, we see that A exceeds B. Therefore, by the assumed transitivity property, A will also exceed any other attribute exceeded by B. In this case, no attributes are exceeded by B, so no new information is gained. However, note that attribute C exceeds D, and also $D > F$ and $D > H$. Hence, we can conclude that $C > F$ and $C > H$. These relationships are now entered into the cross table. (Figure 5).

	A	B	C	D	E	F	G	H
A	X							
B		X						
C			X					
D				X				
E					X			
F						X		
G							X	
H								X

Figure 5 Addition of relationships $C > F$ and $C > H$

Now proceed systematically down the table, entering any new relationships that arise due to transitivity. The cross table will end up as in Figure 6.

It is possible that transitive relationships entered late in the procedure may have an effect on attributes A, B, C, etc. Hence, another pass must be made through the table, entering transitive relationships. Continue in this way until a pass is made through the table without encountering a new transitive relationship — this will indicate that all relationships have been entered. In our example, no new transitive relationships are encountered in the second pass. Thus the final table is as it appears in Figure 6.

It is possible that a 'transitive inconsistency' of the form $a_1 > a_2 > a_3 > \dots > a_N > a_1$ may be present. This will emerge only when the transitive relationships are entered into the table. Thus, at each stage of updating the table, we must check for inconsistencies by the procedure outlined above. The presence of a transitive inconsistency causes a serious problem. Presumably if a_1 exceeds a_2 , which in turn exceeds a_3 etc., then there should be a distinct preference for a_1 over a_N . The fact that the respondent has ranked a_N above a_1 indicates a large degree of confusion. Two courses of action are open: either the entire transitive inconsisten-

	A	B	C	D	E	F	G	H
A	X							
B		X						
C			X					
D				X				
E					X			
F						X		
G							X	
H								X

Figure 6 Inclusion of all relationships that arise due to transitivity

cy and all its transitive implications can be deleted from the table, or the particular questionnaire can be discarded. We recommend the latter option, for simplicity.

Returning to the example, we can gain an idea of the overall ranking of attributes by adding up the rows of the cross table. The scores achieved are as follows:

Attribute	Score
A	1
B	0
C	7
D	4
E	4
F	1
G	4
H	2

It is clear that attribute C is preferred to all others in the survey, and hence is ranked 1. However, attributes D, E and G tie for second place. This is due to a combination of two factors:

- (i) in the design of the groups, attribute D was not compared directly to either E or G, and hence no clear preference could be established, and
- (ii) the logical inconsistency of the respondent in choosing between E and G.

The fact that there is no preference between attributes D, E and G is confirmed by reference to the final cross table. Thus, we are unable to assign unique ranks 2, 3 and 4 to these three attributes. The best we can do is assign a tied rank 3 to all.

Hence the final ranking of attributes in the example is:

Attribute	Rank
C	1
D	3
E	3
G	3
H	5
A	6,5
F	6,5
B	8

In general, if a tied score situation arises, the tie may be broken if a definite preference between the attributes involved has been established in the cross table. Otherwise, a tied rank must be assigned, or the tie can be broken randomly.

If the tie must be broken, (for example if the ranks are to be used as input to a computer package that will not accept ties), then we are stuck with the necessity of reverting to a random choice. If the tie-break is applied truly randomly over a large sample, then the final results should not be seriously affected.

In cases where the rankings assigned by a large sample of respondents will be aggregated, there is little point in breaking ties. Overall, the preferred attributes will come out on top, and the least preferred attributes will score badly.

More direct comparisons between attributes can be achieved by setting up a third grouping of attributes. This will lessen the possibility of ties due to lack of direct comparisons, such as (i) above. On the other hand, the questionnaire will be lengthier, and the likelihood of more logical inconsistencies is increased.

Illustrative example

Hedley recently performed a survey aimed at ranking attributes of hybrid seed maize.² The algorithm described in this note was used to rank 17 attributes studied. The 17 attributes were broken into three groups of six, there being one repetition. The sample size was 341.

The respondents were also asked to rate each of the attributes on a scale of 1 to 10. Hence, an alternative ranking can be deduced for each attribute on the basis of its average score. A summary of scores and ranks appears in Table 1.

The Spearman rank correlation coefficient for the above comparison is 0.87, showing that the two methods result in similar results. Hence, we have shown that the algorithm

Table 1 Summary of scores and ranks

Attribute	Average score	Rank by score	Rank by algorithm
A	9,80	1	1
B	9,27	2	2
C	8,89	3	3
D	7,35	8	4
E	7,83	5	5
F	7,38	7	6
G	7,23	9	7
H	8,02	4	8
I	7,55	6	9
J	6,42	13	10
K	6,41	14	11
L	6,99	10	12
M	6,16	16	13
N	6,44	12	14
O	6,24	15	15
P	6,53	11	16
Q	5,33	17	17

yields reasonable results at least: it is unfortunately impossible to demonstrate that the results are or are not 'correct'. Note particularly the similarity of the results amongst the most highly ranked attributes.

Reference

1. Miller G.A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol. Rev.* 63, p.81.
2. Hedley S.M. (1982). Product attribute analysis in the hybrid maize industry with specific reference to Pioneer Seed Co. (Pty) Ltd. Unpublished technical report, Graduate School of Business, University of Cape Town.