# An Integrated Approach to the Analysis of Binary Choice Data

Gregory W. Cermak, Gerald J. Lieberman, and Harold P. Benson
General Motors Research Laboratories

An integrated package of statistics is described for measuring the relationship between binary dominance data, such as choices, and predictor variables that measure the objects or stimuli. The degree of fit between predictors and choices is defined, and procedures for determining the maximum possible fit and minimum acceptable fit are described. The statistics are applied, for illustrative purposes, to a practical data set.

In the course of work on a problem in applied psychological measurement, a collection of techniques for analyzing binary choice data was assembled. Individually, the techniques are not new. Collectively, the techniques allow the asking of a set of questions that previously had not been applied to binary choice data in an integrated way. The questions themselves are commonly asked of other sorts of data. To make the discussion concrete, suppose there is a set of items and an ordering of them by preferability (or by utility or goodness, for example) is being sought. Suppose also that various attributes of each item can be measured and that there is some point in accounting for the preferability of items in terms of these measured attributes. Some index of fit between preferability and a measure of an attribute can be computed, for example, a correlation. The following questions can be asked or are implicit in the use of most fit statistics: (1) How well does a given measure account for preferability compared to the maximum possible fit? (2) How well does the measure perform compared to the minimum possible fit? If rated preferability and some measure of the items were being correlated, these questions would be trivial: The maximum possible correlation is 1.0 (or −1.0) and the minimum is 0.0. If accounting for choices between pairs of items by means of measured attributes were being attempted, the same questions would be more difficult. The topic of this paper is the problem of assessing how well a measure, or an independent variable, accounts for observed choices. In solving this problem several questions must be asked of the data, including the two questions above. Methods are presented for answering these questions, and the present approach to the problem is illustrated with a practical data set.

First, why collect disaggregated choice data? More generally, why observe binary dominance data? Dominance is one of the most primitive relationships between two items, and choice is a basic observable indicator of the relative preferability of two items. Choice between two items is a simple task for untrained individuals; it avoids the use of response scales, and it does not require persons to match numbers to items. The experimenter assumes very little about persons,

31

the task, and the data when using binary choice. In contrast, any procedure for producing multiple-state choice data requires assumptions that binary choice avoids. A multiple-state response, such as rating confidence of a choice, requires the assumption of a common response scale when data for different pairs of items are compared. Furthermore, that common response scale would usually be assumed invariant under linear transformations. If binary data from individuals are aggregated to form choice probabilities, then it must be assumed that there is some function to transform probabilities into an underlying unobserved scale. Naturally, information about individual differences is sacrificed. If within-person repeated measures of binary choice are obtained, choice probabilities could be estimated, but then it would again be necessary to assume a function relating probability to an underlying scale.

Given, then, that there is interest in collecting disaggregated binary choice or dominance data, and interest in accounting for those data using independent predictor variables, three major questions arise: (1) Are the data significantly nonrandom? (2) Which orderings of the stimuli are most consistent with the data, if the data are nonrandom? (3) How well are the data accounted for by variables specified by the experimenter? Existing ideas about binary dominance data can be used to create statistics that help answer these questions. Previous work has been primarily directed toward the first two questions, with the third appearing as a minor special case (although see a recent discussion by Bezembinder, 1981). In contrast, the initial concern here was with the third question; but to answer it, answers were required for the first two questions.

The nonrandomness question is the one which prompted the main line of modern work on binary choice begun by Slater (1961) and developed by DeCani (1969), Phillips (1967, 1969), and Thompson and Remage (1964), among others (see Hubert, 1976, for a review). Slater (1961) argued that a randomly generated set of binary data would not generally be compatible with any ordering of the items, so that a hypothesis of randomness could be tested by determining how inconsistent a data set is with respect to the best possible ordering of the items. ("Compatible," "inconsistent," and "best possible ordering" will be more fully defined below.) Kendall and Babington Smith (1939) had proposed a different measure of randomness of binary choice data, but the authors agree with Slater's (1961) argument for a measure based on pairs of items rather than on Kendall and Babington Smith's triads (also see Bezembinder, 1981).

The question of which ordering is the "best possible" for a given data set has received attention in several fields (Hubert, 1976). The question is of interest as a mathematical problem, as a computational problem, and as a practical data-analysis problem. Sometimes nonoptimal orders have been used for convenience in dealing with binary choice data (e.g., Carroll, 1972; David, 1969; Slater, 1960). The applications considered here require a best or optimal order.

The question of how well a binary choice data set is accounted for by variables specified by the experimenter has received very little attention. Hubert (1976) shows how his general formulation of the seriation problem subsumes hypothesis testing in this sense as a special case. In contrast, the approach here begins with the question of hypothesis testing or assessing the fit of unidimensional measures to data. Bezembinder (1981) proposes four measures of the fit of choice data to independent variables, but each of these measures requires assumptions that can be avoided by the simple method proposed here. Also, Kendall's (1970) coefficient of agreement between two choice matrices might be viewed as a method for assessing agreement between choice data and a unidimensional measure, although it is not presented for that purpose. However, the null hypothesis of random data considered by Kendall when matching matrices is not appropriate for data and unidimensional measures that are known to be nonrandom (see

Bezembinder, 1981). Similarly, the use of rank correlation statistics is not appropriate for the case considered here in which the binary data are not necessarily consistent with any ranking.

## Other Approaches to Choice Analysis

Apart from the approaches above, which deal with disaggregate binary choice (at least as a special case), four other classes of choice analyses and models might be compared to the optimal linear order approach.

1. A Thurstonian analysis could produce a group preference scale for comparison with independent measures of the stimuli via some form of correlation or multiple regression. However, that approach is not designed to produce a preference scale for an individual given a single choice for each ordered pair of stimuli. Also, it assumes a particular functional relationship between choice probability and preference scale.

2. Unfolding seeks to model preference data (e.g., a preference order) as a function of an underlying scale and a person's ideal point on that scale. An optimal order for a person's choices would serve nicely as input to such an analysis (e.g., Cermak & Cornillon, 1976).

3. Multidimensional analyses of choice data are useful, and in fact were applied to the illustrative data described later (Cermak, 1979; Cermak & Cornillon, 1976). These analyses begin with an ordering of the stimuli for each person and find multidimensionality in between-person differences.

4. Tversky's (1969) lexicographic semiorder model can be used to model choice data. That model, as well as other nonadditive multiattribute models, is incompatible with an ordering approach. However, the lexicographic semiorder model does not happen to describe the illustrative data presented here as well as do optimal orders (Cermak, 1979).

## Three Key Statistics and Their Computation

Consider the general case in which (1) a simple binary dominance relationship is observed for every ordered pair of $n$ items and (2) some unidimensional measure $M$ with at least ordinal properties is available for each item. The goal is to decide how well the measure $M$ accounts for the dominance data even if these data are not consistent with a ranking of the $n$ items. Note that since a ranking of the items does not necessarily exist, rank-order correlation is not appropriate. Three key statistics can be used to determine the goodness of fit of $M$.

## The Statistics

The first statistic is a distribution-free statistic for the fit of $M$, or any order of the integers 1 through $n$, to the dominance data. This fit is given by the number of pairs whose dominance relationship is consistent with $M$ (Cermak & Cornillon, 1976; Hubert, 1976). It is calculated by simply counting the number of choices consistent with the order implied by $M$. By "order" is meant a linear order, namely, a complete, transitive, asymmetric relation. By "consistent with" it is meant that if the measure for item $i$ exceeds the measure for item $j$, and if $i$ is chosen when $i$ and $j$ are presented, then one choice is counted as being consistent with $M$.

The second statistic is the minimum acceptable fit of $M$ to the dominance data. $M$ provides a minimum acceptable fit if it fits the data better than 95% (or 99%) of the random permutations of the integers 1 through $n$. To calculate this statistic, monte carlo methods can be used. Random permutations of the integers 1 through $n$ are generated, and for each permutation the fit to the given choice data set is found. The minimum acceptable fit is read from the distribution of fits for the random permutations. Note that the null hypothesis implied is quite different than one commonly seen in which two random binary choice matrices are compared. The order implied by any $M$ or any permutation

of integers leads to a binary matrix that is complete, transitive, and asymmetric, i.e., far from random. The choice data are not assumed random. Thus, the null hypothesis is that the choice data may be structured and that any measure is structured, but measure $M$ does not fit the data better than do random permutations.

The third statistic is the maximum possible fit of $M$ to the dominance data. Its definition is somewhat more involved than those of the first two statistics. First, assume that any candidate measure must order the items on which the choices are based. For every ordering of the items, there is a unique fit, which is just the number of binary data that are consistent with the order in the sense described above. Since there is a finite number of orders of $n$ items ($n!$), at least one order can always be found whose fit to the data is greater than (or equal to) the fit of all other orders. An order with the maximum fit to a given set of binary dominance data can be called an "optimal order." This maximum fit is not generally equal to the total number of binary data because of the fact, pointed out by Slater (1961), that real binary data are not generally completely consistent with any order. (That is, the data may not always be transitive and asymmetric, although an order is.) Since a measure $M$ is also an order, clearly the fit of any $M$ cannot exceed the fit of an optimal order. Thus, in contrast to regression or rank-order correlation, the maximum fit of a variable (or measure $M$) to a data set is not fixed at the total number of pairs but is limited by the fit of an optimal order for that particular data set.

## Computing the Statistics

The problem of finding an optimal order for a given set of binary dominance (or choice) data has been thoroughly reviewed by Baker and Hubert (1977), Hubert (1976), and Hubert and Schultz (1975). Methods for solving the problem are well known in the field of operations research. Two methods are presented here that enable the data analyst to answer useful questions about psychological data. One approach is heuristic in nature; the other uses a technique known as branch-and-bound.

*The heuristic algorithm.* This approach was constructed to successively improve an order so that it more and more closely fits a binary choice data set in the sense described above. Heuristics, in general, do not necessarily guarantee an optimal order, but this heuristic often does produce one. When it produces a suboptimal order, that order is usually very close to optimal. The advantage of the heuristic is its running speed on a computer and its relative insensitivity to the size and lack of structure of the choice data set. In the illustration to follow, the average CPU time for the heuristic on an IBM 370-168 computer was .075 seconds per preference matrix. An optimal order was achieved for 49 of 51 preference matrices on which it was run. The heuristic missed four choices that optimal orders correctly accounted for, yielding a 99.94% success rate for the heuristic compared to best orders.

The heuristic begins by constructing a good, easy-to-compute order, the so-called "voting order." The voting order ranks each stimulus item (or choice alternative) according to the number of stimuli it was preferred to in binary comparisons minus the number of stimuli preferred to it. This type of order, incidentally, is used in Carroll's (1972) MDPREF preference analysis program, and is extensively used in David's (1969) work on paired comparisons. The voting order is not optimal, in general (see Appendix). In the illustration to follow, the voting order was optimal for 4 of the 51 choice matrices for which it was computed.

To improve a voting order, the heuristic searches for a pair of stimuli $i$ and $j$ for which the voting order contradicts the given choice data. Suppose the given order has ..., $i$, ..., $j$, ..., but in the choice data $j$ is preferred to $i$. The heuristic seeks a new order that contradicts the data less often by moving $i$ back in the order and/or $j$ forward, leaving the order of the other stimuli unchanged. The result is a new order ..., $j$, $i$, ... with $j$ in the $k$th position. Such an exchange is called an $i$, $j$ exchange in position $k$. It will, of course, eliminate the contradiction for

pair $i$, $j$, but it may create new contradictions for other pairs involving $i$ and $j$. The heuristic iteratively searches for $i$, $j$ exchanges in all positions $k$ that will decrease the contradictions between the order and the data. The process stops when no further improvements are possible using any $i, j$ exchange (even though improvements may be possible with more complicated exchanges). The resulting order is at least as good as the voting order, is usually better, (e.g., in 45 of 51 cases examined here), and often is optimal.

*Branch-and-bound.* This approach is an enumerative technique for solving an optimization problem without explicitly enumerating all of the possible solutions. In this case the problem is to find an ordering of the stimuli with a minimum number of contradictions to the actual choice data. The number of possible orders, $n!$, can be extremely large. To find an optimal order, branch-and-bound explicitly enumerates only a small fraction of these orders. The remaining orders are considered implicitly.

To accomplish this, a branch-and-bound technique repeatedly partitions the set of all orders into mutually exclusive and jointly exhaustive categories (the branches). For each category, a lower bound is determined for the number of contradictions that orders in the category could possibly achieve. Periodically, orders are evaluated in this process. Any category of orders whose lower bound exceeds the number of contradictions of one of these known orders is eliminated from further consideration. This process of partitioning into categories, establishing lower bounds, and eliminating categories continues until the lower bounds on all remaining categories equal or exceed the number of contradictions of the best candidate order yet evaluated. When this occurs, any order having that number of contradictions must be optimal. For details and a more general description of branch-and-bound, see the review articles of Agin (1966), Lawler and Wood (1966), and Mitten (1970).

The branch-and-bound algorithm for finding an optimal order used in the current application was quite similar to an algorithm proposed by Phillips (1969). To decrease computational effort, the algorithm employed here differed in certain procedural details from Phillips' algorithm. Most major steps in the two algorithms were identical, including the steps for the formation of categories of orders and for the establishment of lower bounds.

The branch-and-bound algorithm, unlike the heuristic algorithm, is guaranteed to produce an optimal order. However, this guarantee is usually achieved at the expense of longer computer running times than those required by the heuristic, especially when the number of stimuli is large and the data set is only weakly structured. Average CPU time for the data to follow was .20 seconds per matrix on an IBM 370-168. For randomly generated binary matrices, average CPU time was 6.3 seconds. In the following illustration the branch-and-bound algorithm was used on all data sets, while the heuristic was applied only to data from the first three experiments.

## Application to a Practical Data Set

### Method

The data set introducing the optimal order problem has characteristics that are fairly general in practical applications. Binary choice on pairs of items was the appropriate response because a simple indication of preference was required, yet the stimuli were complex and multidimensional, and the subjects had no special training. In particular, the stimuli were 13 traffic sounds recorded at a wide variety of locations—central business district, freeway, urban arterial, suburban collector street, and so forth (see Cermak, 1979; Cermak & Cornillon, 1976, for complete descriptions of the experiments). The stimuli might as easily have been a set of products or representations of environments or descriptions of other people or many other perceptual stimuli. The 65 subjects partitioned among the experiments were mainly suburban housewives with high school educations. By using binary choice, no assumptions were required concerning subjects' knowledge of number scales. Further, by defining the task as simply

choosing the sound in each pair that the person would rather hear, verbal characterizations of the sounds' attributes, and of the subjects' reactions, were unnecessary (see Molino, 1974, for arguments against verbal descriptors in psychophysical tasks).

The object of the experiments was to determine preference differences among the sounds, and then to account for these differences using various acoustical measures of the sounds. Hopefully, one measure would have a clear advantage over others in accounting for subjects' choices and thus emerge as the most psychologically justifiable way to measure traffic sound. Four experiments were conducted, arranged as a replicated pair of experiments. Experiments I and II used stimuli collected in one environment; Experiments III and IV used stimuli collected in another. Within each pair, the first experiment used unaltered recorded traffic sounds. The second used the same sounds, which had been rerecorded such that whatever acoustical measure happened to best account for choices in the first experiment was held constant for all sounds in the second experiment. (As it turned out, this objective was technically realizable in both pairs of experiments, and in fact the same measure was equalized in both pairs.) The questions in Experiments II and IV were whether the equalization of the main acoustic cue in the sounds would remove all preference information so that choices would be random, and if not, what acoustical measures would best account for the choices.

## Analysis

Individual subjects' data were analyzed separately. This point is to be emphasized, since separate analyses are not possible for some alternative methods. Questions asked of the data were (1) Are choices nonrandom for unaltered sounds? for altered sounds? (2) How well do acoustical measures account for choices? (3) What is the minimum acceptable level that an acoustical measure should achieve in accounting

for choices? and (4) What is the maximum possible fit that an ideal acoustical measure could achieve? Again, these same questions might be asked if other sorts of stimuli were employed in a binary choice experiment. The answers to these questions involve properties of orders. Although each piece of the computational machinery required to answer these questions exists somewhere in the literature, this may be the first instance in which the machinery has been brought together to form a unified analysis package and has actually been demonstrated.

The hypothesis tested in determining whether choices were reliably nonrandom was, not that each choice was made randomly with probability .5. Instead, a stronger hypothesis was tested—that the choices were both nonrandom and consistent with a linear ordering of the stimuli. The test of the hypothesis consists of (1) finding an optimal order for a subject's data; (2) noting the number of choices not accounted for by this order, essentially Slater's (1961) $i$ statistic; (3) randomly generating 200 artificial data sets; (4) computing the number of choices not accounted for by an optimal order for each data set; and (5) noting whether the number of choices not accounted for in a subject's data was smaller than 99% of the distribution for the randomly generated data.

All choice data sets passing this test were examined for their degree of agreement with acoustical measures of the stimuli. The assumption was that if an acoustical measure for sound $i$ exceeded that measure for sound $j$, then sound $j$ would be chosen. That prediction was compared to the actual choice. The number of choices compatible with, or predicted by, a given measure was counted for each measure for each subject's choices. The actual measures used consisted of indicators of average sound level, of change in level with time, and of spectral shape (Cermak & Cornillon, 1976).

The minimum acceptable fit of acoustical measures to the choice data was defined with respect to a null hypothesis that they predict no better than random. A reasonably inexpensive
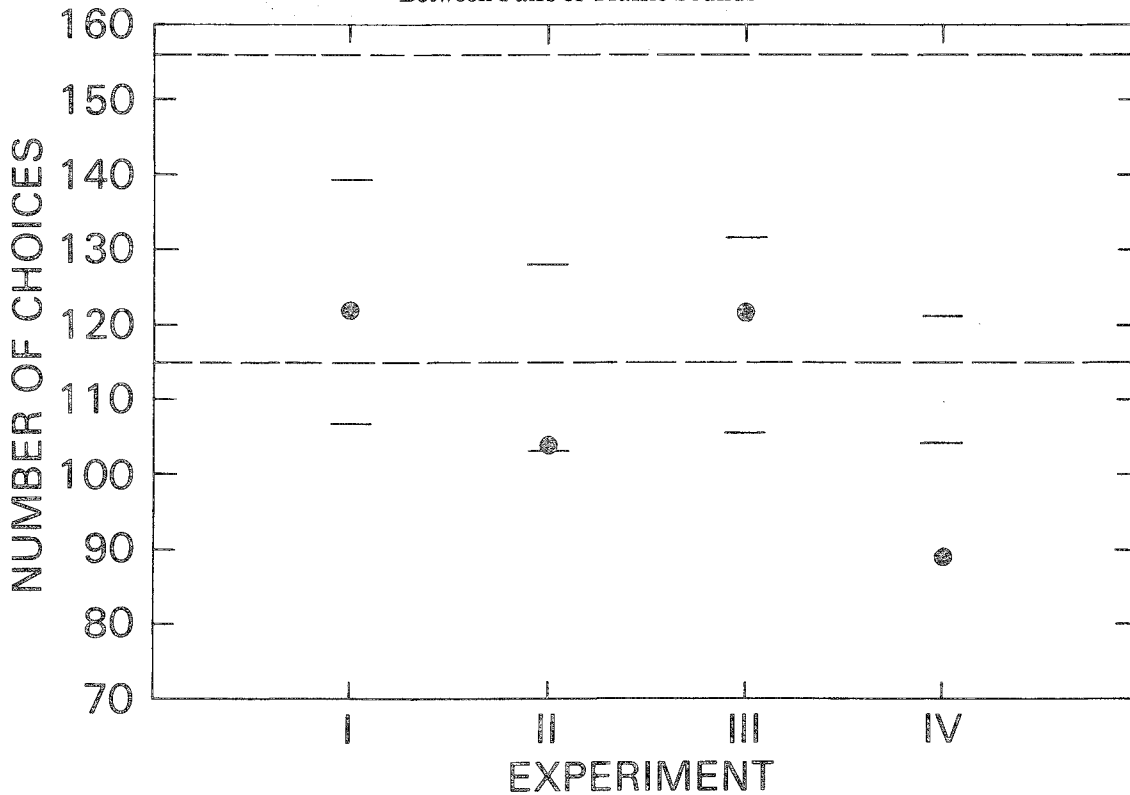
brute force approach to finding the minimum acceptable fit is a monte carlo study of the fit of random permutations of integers to a given subject's choice data. In the present case, 100 random permutations of the integers 1 through 13 were fit to each subject's choice data, and the fits were ordered from worst to best. The 96th fit in the list was taken as an estimate of the fit beyond which a measure could be said to predict choices reliably better than chance. Note that this minimum acceptable fit depends on the individual subject's data, and substantial differences in minimum acceptable fit occur for differences in the amount of structure in subjects' data.

The maximum possible fit of any measure to the choice data is given directly by the same statistic that was computed in the monte carlo evaluation of whether the data were nonrandom, namely, the fit of an optimal order. The fit of an optimal order, it should be noted, is specific to a particular data set, as is the case with the minimum acceptable fit.

The main outlines of the results for the four experiments are summarized in Figure 1. Again, these particular results are presented to illustrate the analysis. The top dashed line represents the total number of choices in each experiment. The dashed line toward the middle of the figure represents the minimum number of choices that must be consistent with an optimal order if it is to be concluded that a choice data set is reliably nonrandom. The mean number of choices actually consistent with an optimal or-

Figure 1
Results of Four Experiments in Which Subjects Chose
Between Pairs of Traffic Sounds

der, for a given experiment, is given by the upper bar for that experiment. These upper bars all lie between the two dashed lines, indicating that the choice data for all experiments were less structured than data generated by linear orders of the stimuli (as would be expected), but more structured than random data, on the average. These results held also for all but two subjects' individual data (in Experiment IV).

The upper bar for each experiment also refers to the mean over choice data sets of the maximum possible fit for any measure of the stimuli that orders the stimuli. The lower bar for each experiment refers to the mean of the minimum fits. The dots represent the mean number of choices accounted for the the best measure of the stimuli (of those examined) for each experiment. For three of the experiments the dots are bounded by the bars, that is, the best acoustical measure did worse than optimum (as would be expected) but better than chance in accounting for choices. For the fourth experiment, the best measure failed to perform significantly better than chance; in the second experiment the best measure only slightly outperformed the upper 5% of random permutations.

The acoustical measure that fit the choice data best in Experiment I was the energy-equivalent sound level, Leq. In Experiment II the same sounds were rerecorded to be equal to their Leq levels. Although the choice data were nonrandom, the best acoustical measure (a new measure of spectrum shape) fit them relatively poorly. This indicates that the residual information in the stimuli, on which choice was based, was not being picked up well by noise measures. The same pattern of results was obtained in Experiments III and IV: Leq was again the best predictor in Experiment III, using a new sample of sounds. When these sounds were made equal in Leq (Experiment IV), choices were nonrandom but were not well accounted for by any noise measures examined.

Thus, the set of experiments that illustrate the use of the fit statistics showed (1) that one particular traffic noise measure, Leq, was the best predictor of subjects' choices between traffic sounds, (2) that some room for improvement in prediction is possible, (3) that if Leq were removed as a variable (by recording sounds at equal Leq levels) then no other acoustical measure predicted subjects' responses well, and (4) that usable information remains in equal-Leq sounds. These results can be used in models of traffic noise impact (e.g., Cermak, von Buseck, Blanchard, & Cervenak, 1981).

## Summary

Existing ideas about binary dominance data can be used to create statistics for evaluating the fit of unidimensional independent variables to binary choices. The application of these statistics was illustrated with a practical data set. The statistics described were (1) a measure of fit between binary dominance (choice) data and a unidimensional descriptor of the stimuli, (2) a maximum possible fit between a given choice data set and any unidimensional descriptor, and (3) a minimum acceptable fit, reliably greater than chance, between a given choice data set and any unidimensional descriptor. The maximum fit measure is essentially Slater's (1961) *i* statistic, which was originally proposed as a measure of internal consistency in a set of choice data. The maximum fit measure was also used in this role as an indicator of nonrandomness in choice data.

Because these statistical tools were available, it was feasible to design experiments on perception of traffic noise around binary choice and still analyze disaggregate data (as opposed to using Thurstonian scaling on aggregate data). The use of binary choice as the response solved potential experimental design problems concerning subjects' analytical and verbal abilities and provided a simple operational definition of "relative aversiveness of traffic noise." The statistical tools, of course, apply more broadly to any binary dominance data.

## References

Agin, N. Optimum seeking with branch and bound. *Management Science*, 1966, *13*, B-176–B-185.

Baker, F. B., & Hubert, L. J. Applications of combinatorial programming to data analysis: Seriation using asymmetric proximity measures. *British Journal of Mathematical and Statistical Psychology*, 1977, *30*, 154-164.

Bezembinder, T. G. G. Circularity and consistency in paired comparisons. *British Journal of Mathematical and Statistical Psychology*, 1981, *34*, 16-37.

Carroll, J. D. Individual differences and multidimensional scaling. In R. N. Shepard, A. K. Romney, & S. B. Nerlove (Eds.), *Multidimensional scaling* (Vol. 1). New York: Seminar Press, 1972.

Cermak, G. W. Exploratory laboratory studies of the relative aversiveness of traffic sounds. *Journal of the Acoustical Society of America*, 1979, *65*, 112-123.

Cermak, G. W., & Cornillon, P. C. Multidimensional analyses of judgments about traffic noise. *Journal of the Acoustical Society of America*, 1976, *59*, 1412-1420.

Cermak, G. W., von Buseck, C. R., Blanchard, R. D., & Cervenak, J. G. Analysis of traffic noise abatement strategies. *Journal of the Acoustical Society of America*, 1981, *69*, 158-170.

David, H. A. *The method of paired comparisons*. London: Griffin, 1969.

DeCani, J. S. Maximum likelihood paired-comparison ranking by linear programming. *Biometrika*, 1969, *59*, 537-545.

Hubert, L. J. Seriation using asymmetric proximity measures. *British Journal of Mathematical and Statistical Psychology*, 1976, *29*, 32-52.

Hubert, L., & Schultz, J. Maximum likelihood paired-comparison ranking and quadratic assignment. *Biometrika*, 1975, *62*, 655-659.

Kendall, M. G. *Rank correlation methods*. London: Griffin, 1970.

Kendall, M. G., & Babington Smith, B. On the method of paired comparisons. *Biometrika*, 1939, *31*, 324-345.

Lawler, E. L., & Wood, D. E. Branch-and-bound methods: A survey. *Operations Research*, 1966, *14*, 699-719.

Mitten, L. G. Branch-and-bound methods: General formulation and properties. *Operations Research*, 1970, *18*, 24-34.

Molino, J. A. Measuring human aversion to sound without verbal descriptors. *Perception & Psychophysics*, 1974, *16*, 303-308.

Phillips, J. P. N. A procedure for determining Slater's *i* and all nearest adjoining orders. *British Journal of Mathematical and Statistical Psychology*, 1967, *20*, 217-225.

Phillips, J. P. N. A further procedure for determining Slater's *i* and all nearest adjoining orders. *British Journal of Mathematical and Statistical Psychology*, 1969, *22*, 91-101.

Slater, P. The analysis of personal preferences. *The British Journal of Statistical Psychology*, 1960, *13*, 119-135.

Slater, P. Inconsistencies in a schedule of paired comparisons. *Biometrika*, 1961, *48*, 303-312.

Thompson, W. A., Jr., & Remage, R., Jr. Rankings from paired comparisons. *Annals of Mathematical Statistics*, 1964, *35*, 739-747.

Tversky, A. Intransitivity of preferences. *Psychological Review*, 1969, *76*, 31-48.

## Appendix

The following is a case in which the voting order is not optimal for a set of (constructed) binary data. A "1" in the matrix indicates that the row item was preferred to the column item; a "0" indicates that the column item was preferred. Diagonal elements, marked "X," are not considered. Note that both orders of each pair of items are represented. Recall that voting ranks are determined by summing the number of items to which an item is preferred (row sum) and subtracting the number of items chosen in preference to it (column sum). The original "choices" are given in Table A.

### Table A
### A Constructed Example
### of Binary Choice Data

| Stimulus | Stimulus | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | x | 1 | 0 | 1 | 1 |
| 2 | 0 | x | 0 | 1 | 1 |
| 3 | 0 | 0 | x | 0 | 1 |
| 4 | 0 | 0 | 1 | x | 1 |
| 5 | 0 | 0 | 0 | 1 | x |

The voting ranks, $v(i)$, for these items are $v(1) = 3-0 = 3$, $v(2) = 2-1 = 1$, $v(3) = 1-1 = 0$, $v(4) = 2-3 = -1$, and $v(5) = 1-4 = -3$. Hence, the items are listed in the correct voting order. However, the preference matrix implied by the voting order is given in Table B.

Table B
Preference Matrix Implied
by the Voting Order

| Stimulus | Stimulus | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 1 | x | 1 | 1 | 1 | 1 |
| 2 | 0 | x | 1 | 1 | 1 |
| 3 | 0 | 0 | x | 1 | 1 |
| 4 | 0 | 0 | 0 | x | 1 |
| 5 | 0 | 0 | 0 | 0 | x |

The matrices in Tables A and B differ in five places, so the voting order has an error rate of 5.

Table C
Preference Matrix Implied
by an Optimal Order

| Stimulus | Stimulus | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 4 | 3 | 5 |
| 1 | x | 1 | 1 | 0 | 1 |
| 2 | 0 | x | 1 | 0 | 1 |
| 4 | 0 | 0 | x | 1 | 1 |
| 3 | 0 | 0 | 0 | x | 1 |
| 5 | 0 | 0 | 1 | 0 | x |

However, the order 1, 2, 4, 3, 5 has an error rate of only 3, and is, in fact, the optimal order (Table C). Thus, the voting order need not be an optimal order.

### Author's Address

Send requests for reprints or further information to G. W. Cermak, Societal Analysis Department, General Motors Research Laboratories, Warren MI 48090. Harold P. Benson is currently affiliated with the Department of Management and Administrative Sciences, University of Florida.