SESSION VIII

CONTRIBUTED PAPERS: ON-LINE STATISTICAL SYSTEM

N. JOHN CASTELLAN, JR., *Indiana University*, *Presider*

An interactive graphics approach to dimensional analysis*

EDWARD J. SCHNEIDER

Center for Political Studies, Institute for Social Research, University of Michigan, Ann Arbor, Michigan 48104

and

HERBERT F. WEISBERG

Department of Political Science, University of Michigan, Ann Arbor, Michigan 48104

This paper describes an interactive approach to dimensional analysis. This approach permits the researcher and student to inspect alternative spatial representations along with a wide range of rotations. Some understanding of problems inherent in multidimensional scaling is presented to help provide an appreciation of the utility of the method.

Social science researchers are accustomed to a passive mode of data analysis. They "run a computer program" and accept its results as givens. The development of interactive computer technology permits a more active mode. Analysis strategies can be modified on the basis of partial results. Intervention by the researcher can redirect the analysis and permit an unusually exhaustive examination of the data structure.

We have decided to adopt an interactive approach toward dimensional analysis. Dimensional analysis seeks a spatial representation for a set of data. For example, public attitudes toward contenders for the presidency might be analyzed to determine the dimensional underpinnings of those attitudes.¹ If the public views the various Democratic candidates as more similar to each other than they are to the Republican candidates, the Democrats would be placed near each other in the spatial representation. If attitudes toward the candidates depend mainly on partisan considerations, then the main dimension in the space would contrast the Republican candidates on one side with the Democratic ones on the other side. More formally, the dimensional techniques seek spatial representation of measures of proximity or dominance between pairs of stimuli (or possibly between individuals and stimuli).²

Interactive dimensional analysis would permit the researcher to inspect alternative spatial representations along with a wide range of rotations. Chang (1973) has employed interactive graphics for rotation of

dimensional configurations, but interactive graphics can play a more central role in the preceding scaling process. The iterative minimization procedures employed in some dimensional techniques may result in suboptimal solutions. Interactive analysis permits the researcher to test the optimality of a solution by modifying that solution by hand. The researcher might be able to force a solution that is superior (on substantive and/or statistical grounds) to that of a noninteractive analysis. An appreciation of the utility of an interactive approach requires some understanding of problems inherent in multidimensional scaling.

Nonmetric multidimensional scaling seeks a spatial representation in which interpoint distances are monotone with the corresponding data values.³ The greater the similarity (or the less the dissimilarity) between a pair of variables, the closer together those points belong in a spatial solution. The origin of the space is arbitrary (as it does not affect interpoint distance calculations), as is axis placement in Euclidean space. The analyst specifies the number of dimensions in an initial spatial configuration. The monotonic fit of the distances in that space with the data values is assessed. Kruskal's (1964) stress coefficient is the common measure of fit, where a nonzero stress indicates an imperfect solution. The mathematics of steepest descent is used to determine the directions in which points should be moved to reduce stress. The points are shifted, stress is recalculated, steepest descent is used again if the stress remains nonzero, and the process continues iteratively until an acceptably low stress is achieved or until the stress stabilizes and does not fall as iteration continues. This yields the solution for the chosen

*This research was supported by the Center for Political Studies Computer Support Group at the University of Michigan with the gracious encouragement and considerable assistance of Gregory Marks, Sylvia Barge, Neal Oden, and Stewart Robinovitz.

number of dimensions. Solutions employing different dimensionalities are compared to determine the final scaling solution.

NONMETRIC MULTIDIMENSIONAL SCALING PROBLEMS

Experience with nonmetric multidimensional scaling has shown⁴ that these techniques are subject to several problems of varying levels of seriousness. Among these problems are (1) effects of alternative starting configurations, (2) local minima, (3) indeterminacy of point location within regions, and (4) difficulty of interpretation of results due to arbitrariness of rotation. We believe that an interactive graphics approach can aid in overcoming or minimizing the seriousness of each of these problems.

Arbitrary Starting Configuration

Theoretically, multidimensional scaling techniques should converge to the optimal solution regardless of the structure of the initial configuration. However, a poor choice of starting configuration can result in a local minimum. Blind starts (such as use of a random number generator or Kruskal's multidimensional simplex) are more likely to have local minimum problems than more rational starts (such as a metric start) which take advantage of the structure of the data. An interactive mode permits easy experimentation on a variety of starts.

Local Minima

The steepest descent algorithm is subject to local minima. Rather than finding the global minimum stress value, the iteration might accidentally "get stuck" in a local minimum. The current solution might yield lower stress than any solution which could be obtained by small movements of the points; yet, a much lower stress might be found by drastically moving the points. Unfortunately, the analyst cannot determine whether the solution represents a global or local minimum by inspection alone. The points must be moved and the stress recalculated to determine if further iteration would return to the original solution or lead to a superior one. An interactive mode facilitates the movement of the points and provides quick feedback on the resultant effect on the stress value.

Indeterminacy Within Region

Since the points are situated in accord with a nonmetric rule, the locations of variables are not precisely pinpointed. A specific point location is given, but moving the variable in a narrow region around that point might not violate the ordinal constraints. Thus, a spatial solution might not seem to be in accord with the researcher's prior hypotheses, but interactively moving a few points might confirm the hypothesized structure without increasing the stress.

Arbitrary Rotation

Nonmetric multidimensional scaling seeks to construct a space in which interpoint distances are monotone with similarity measures. Yet, distances are invariant under rotation in Euclidean space, so orthogonal rotation of dimensions is permissible. An interactive graphics mode facilitates interpretation of the space by permitting the researcher to see the effect of various rotations. Principal axis and varimax rotation might be used as in factor analysis or the researcher might experiment with rotations of different numbers of degrees to yield maximally interpretable axes.

INTERACTIVE ANALYSIS

We have adapted Kruskal's MDSCAL-V program for nonmetric multidimensional scaling into an Interactive Multidimensional Scaling Program (IMDS). Component analysis and orthogonal rotation procedures are incorporated into the package. IMDS is written in FORTRAN and is currently operative on an IBM Model 360/67 computer. The Michigan Terminal System (MTS) provides the interactive capabilities. The program can be run at a standard Teletype, but it is most effective at a graphics terminal (such as a Computek or a Tektronix) with a cathode ray tube (CRT) display.

IMDS is run by issuing a set of commands specifying particular operations to be performed. The user types a one-word command. If the command is straightforward, it will be executed immediately and the results will be displayed on the terminal. For example, typing the command VARIMAX results in a display of the varimax rotation. If more information is required for the command, the program prompts for that information. A typical command involving prompting is the READ command, which directs that the data matrix be read by the program. When the user types READ, IMDS prompts for parameters such as the location of the data file, the format of the data, and missing data conventions. Almost all parameters have appropriate defaults. Once the parameters have been specified or defaulted, the command is executed. After a command is executed and the results are displayed, IMDS invites the user to type in another one-word command.

The command structure permits the user to perform (and reperform) the several steps in the scaling process in a very flexible order. The data matrix must be read and an initial configuration generated (using a random start, Kruskal's arbitrary start, a principal component analysis, a user's hypothesized configuration, or a configuration saved from a previous analysis), though the user may later change data matrices or initial configurations to determine the effect on the scaling. The stress can be calculated and/or iteration can be performed whenever desired.

The user can shift points in the space to determine the effect on the stress. A cursor can be used at a graphics terminal to move the points manually. Several pieces of

information can be displayed to suggest which points might usefully be shifted. The stress for each variable can be examined. The residual (error in monotonic fit) for each original data entry is available. Also, the user can see the directions in which each point would be moved at the next iteration by the steepest descent minimization (gradient). Thus, the user can be told how the program would attempt to decrease stress, and the user can seek to improve it directly. Similarly, the user can check whether deletion of a set of variables would markedly improve the stress.

The scaling solution is normally shown at the end of the iteration sequence. The first two dimensions of the space are plotted on the CRT unless otherwise specified. Rotation commands are available to assist the user in attempting to understand the structure of the space. Principal axis and varimax options are provided, plus a capability of turning any pair of dimensions a specified number of degrees. Dimensions can be translated (to add a constant to every value) and reflected (to reverse the sign of every loading). The space can be centered (so the mean value on each dimension is zero) and normalized (so the sum of squared projections on the dimensions equals the number of variables).

When iteration is requested, the program gives the stress as each iteration is performed. When the iteration is completed, the user at a graphics terminal can obtain a "trace" of the iteration. The trace locates the points at the initial configuration, draws a line from each showing the movement during the first iteration, and so on. When the trace is completed, the movements of each point through the iterative process are visible. A "Shepard diagram" can also be plotted to show the degree of monotonicity between original data and actual distance values.

At any stage in the analysis, the user can retrieve much detailed information. For example, the interpoint distances can be shown, parameter values at each iteration can be recovered, and the sum of squared projections on each dimension can be computed to indicate their relative importance. The user is not subjected to large amounts of display information but can selectively retrieve the information most desired.

An essential feature of an interactive graphics approach is an ability to obtain hard copy of the results. A full transcript of the session can be printed remotely at a batch station. Included in the hard copy is more detailed information about the scaling than the user may have cared to have displayed during the interactive session. CALCOMP plots can be obtained of any of the diagrams produced in the scaling.

ILLUSTRATIVE SESSION

The potential of the program is best illustrated by giving an actual research example. Space limitations preclude reproducing the entire analysis, but we can quickly portray the scaling strategy employed. We have

Fig. 1. Kruskal's arbitrary starting configuration in two dimensions (stress = 1.000).

chosen to illustrate how the interaction can aid in the detection of a local minimum solution. The example will suggest strategies for avoiding local minima, both in terms of choice of starting configuration and verification of whether a solution is a global minimum.

The data involve public reactions to political parties in a multiparty country. Respondents in a national sample were asked to rate on a 100-deg thermometer scale how warm or cold they felt toward each party. The correlations between the ratings of a pair of parties indicate the extent to which they evoke similar evaluations. The multidimensional scaling logic specifies that the more the ratings of a pair of parties covary, the closer those parties belong in a dimensional space.⁵

We began the IMDS analysis by issuing a READ command to read in the data matrix. An INITIAL command was used to establish Kruskal's simplex as the starting configuration in two dimensions. Figure 1 presents the L-shaped arbitrary configuration that is displayed at this point on the user's terminal.⁶ The Communist and Conservative parties are identified by name in the figure; other parties are identified only as to whether they are usually considered to be parties of the left, center, or right.

The multidimensional scaling iteration was invoked by means of an ITERATE command with the TRACE option employed. The results are shown in Fig. 2. Notice that the interpoint distances in the starting configuration were not at all monotone with the interparty correlations, as shown by an initial stress value of 1.00 (indicating total lack of fit). The iteration algorithm

MINIMUM ACHIEVED	ITER	STRESS	NEXTSTEP
LAST ITER = 14	0	1.000000	.000038
	1	1.000000	.000103
	2	1.000000	.000351
	3	1.000000	.001238
	4	1.000000	.004312
	5	1.000000	.014705
	6	1.000000	.049237
	7	1.000000	.161487
	8	1.000000	.523525
	9	1.000000	1.680412
	10	1.000000	5.431786
	11	1.000000	17.41340
	12	.999999	55.69710
	13	1.000000	78220528
	14	.999999	70175744

Fig. 2. Iteration trace using Fig. 1 as the starting configuration.

Fig. 3. Solution using Fig. 1 as the starting configuration (stress = .999999).

moved the points from the starting configuration, indicated by the stars in Fig. 2, along the lines to the

Fig. 4. Shift of Communist and Conservative parties in Fig. 3 (stress = .946).

Fig. 5. Solution using Fig. 4 as the starting configuration (stress = .032).

final configuration, indicated by the variable numbers. The final solution is Fig. 3, which still has a complete lack of fit (stress = .999999).

These results could be read as meaning that more than two dimensions are essential for a spatial representation of these data. However, our substantive knowledge

about party systems in general and about the present party system in particular makes us expect that a low dimensionality is appropriate. An alternative interpretation is that these results constitute an unusually severe local minimum. To test this possibility, a normal batch MDSCAL-V analysis would require a resubmission with a new starting configuration, but IMDS permits us to shift the point locations in an effort to improve the fit.

The ALTER command was used to establish a more rational starting configuration, with the CRT cursor moved to where we wished parties shifted. Placing the Communists on the left and the Conservatives on the right, as shown in Fig. 4, was sufficient to reduce stress to .946, a value below that for the original two-dimensional solution. Further use of the ITERATE command shifted the points to minimize the stress function. The final solution (Fig. 5) has an excellent stress of only .032. To aid in interpreting the solution, we used a PRAX command which gave the principal axis rotation displayed in Fig. 6. This rotation emphasizes the importance of a left-right dimension underlying popular evaluations of the several parties. Obviously the data can be represented very well in two dimensions; the original poor fit was indeed a local minimum.

The possibility remains that even better solutions could be obtained by use of an even more rational starting configuration. To establish the rational start we used the PRIN command, which performs a principal component factor analysis (Fig. 7).⁷ The component solution has a stress of .150, but multidimensional scaling invoked by the ITERATE command shifted the

Fig. 6. Principal axis rotation of Fig. 5.

Fig. 7. Principal components analysis of correlation matrix (stress = .150).

Fig. 8. Solution using Fig. 7 as the starting configuration (stress = .026).

points about until the final solution (Fig. 8) was achieved with an excellent stress of only .026. Again to aid interpretation we applied two transformations: The PRAX command provides principal axis rotation and the REFLECT command reverses the direction of the major

Fig. 9. Best two-dimensional solution (a rotation and reflection of Fig. 8).

Fig. 10. Shepard diagram for solution of Fig. 9 (stress = .026).

INDIVIDUAL POINT CONTRIBUTIONS TO TOTAL STRESS OF 0.0262 IN 2 DIMENSIONS

STRESS POINT

0.02112669 403I
0.01659922 401I
0.01629348 397I
0.01501861 400I
0.00976007 398I
0.00786942 402I
0.00761387 404I
0.00752102 405I
0.00295889 399I

Fig. 11. Individual point stress for solution of Fig. 9.

axis so the left-wing parties are on the left side of the plot. The resulting configuration is shown in Fig. 9. This solution has marginally lower stress than the previous solution (Fig. 6), indicating that was also a local minimum.⁸ Points could be shifted around further or additional new starting configurations could be employed to seek an even better solution, but the basic elements of such analysis have by now been demonstrated sufficiently.

The quality of the fit of Fig. 9 to the original data can be further gauged by some auxiliary IMDS diagrams. The SHEPARD command plots (Fig. 10) the interpoint distances of our solution on the horizontal axis against the original correlation values on the vertical axis. A zero stress solution would have all the *s in this diagram fall along a monotone line from the upper-left corner of the diagram down and to the right, to the lower-right corner. There is some dispersion around the monotone regression line of Fig. 10, but the fit of the regression line to the *s is impressive. The .026 stress value summarizes the fit of the regression line in this "Shepard diagram."

The INDIVIDUAL command displays (Fig. 11) the stress associated with each individual point in Fig. 9. There was noticeably more difficulty in fitting some parties than others, but the stress is not just associated with one or two parties that are clearly out of position relative to the others.

In some party systems the extremist parties of the left and right are perceived as being more similar to each other than they are to the centrist parties. To examine this possibility, we used the ALTER command to move the Conservative party to the cluster of left-wing parties (Fig. 12). This dramatically increased the stress to .907. By using ITERATE with the TRACE option (Fig. 13), we can see that the scaling iteration pulled the Conservatives back to the cluster of right-wing parties. Note the movement in Fig. 13. The Conservatives were too far from the right-wing parties, so the program pulls the Conservatives over toward the right. But also the left-wing parties were too close to the Conservatives, so the stress minimization pushes the left-wing parties down away from the Conservatives. Thus, the iteration trace shows the Conservatives returning to the right, the left-wing parties moving down, and, correspondingly, the right-wing parties moving up. The final solution is not as

good as that achieved in Fig. 9, but what is more important is that the hypothesis of adjacent extremist parties has been disproved. Additionally, Fig. 13 illustrates how the iteration process works.

During the same IMDS run, we obtained solutions for these data in three and one dimensions, though to simplify this illustration we have not included those solutions. At the end of the session, we employed the ELBOW command to plot the stress against the number of dimensions (Fig. 14). The drop from three to two dimensions has little effect on the stress, while the drop from two to one drastically increases the stress. On the basis of the "elbow" in Fig. 14 at two dimensions, we conclude that a two-dimensional solution is appropriate. Thus, Fig. 9 is our best spatial representation of the data.

Above we described four multidimensional scaling problems which interactive analysis would help solve. The rotation of the solution is necessarily arbitrary, but the principal axis of Fig. 9 does correspond to the conventional left-right ordering of the parties. Interestingly enough, it does not suggest any single ordering of the parties within the left or right blocs; the cluster separation seems more fundamental here. We have not attempted to shift the points around to test indeterminacy of the solution within small regions, but we do know from Fig. 13 that the fit is destroyed by large shifts of points.

We have shown how shifting points about can allow improvement from a local minimum. Indeed, a conventional multidimensional scaling analysis might not even have caught the fact that Fig. 3 was a local minimum. Simultaneously we have shown that

Fig. 12. Shift of Conservative party in Fig. 9 (stress = .907).

MAXIMUM NUMBER OF ITERATIONS USED
LAST ITER = 15

ITER	STRESS	NEXTSTEP
0	.915362	.465094
1	.807262	.662667
2	.446689	.743176
3	.318990	.577043
4	.238999	.125456
5	.125935	.289287
6	.114891	.107743
7	.100845	.145670
8	.095264	.039262
9	.086756	.073868
10	.084645	.020566
11	.080298	.045433
12	.075338	.037132
13	.072168	.015011
14	.064879	.014667
15	.062465	.016179

Fig. 13. Iteration trace using Fig. 12 as the starting configuration.

Fig. 14. Stress by dimensionality.

experimentation with different starts is useful in detecting local minima. In the present case the Kruskal start led to a very poor solution, while the rational metric start led to a very good result. We will not contend that the principal component start is always

optimal, but we find the IMDS capability to experiment with different starts invaluable in finding a global minimum solution.

PROGRAM DESIGN ISSUES

Several problems confronted us in designing a program for interactive graphics analysis. The design implications of interactive graphics were not fully apparent to us when we began work on this project, but we have since been made quite aware of them. Even the most experienced batch user/programmer has difficulty suggesting design specifications which will be appropriate for an interactive conversational environment. Input/output, command structure, and documentation are critical problem areas.

The input/output requirements of interactive graphics are diverse. Different input/output routines are required to handle graphics and nongraphics terminals.⁹ Special provisions must be made to route output to a batch printer in order to preserve a record of the session. Perhaps the most important problem is that of guessing what the user would like to see displayed on the screen after each command. We have provided automatic displays, though we realize that feedback from users may suggest improvements on these display alternatives.

The command structure of a conversational program is complicated to design. Program functions can be divided into computation and display categories, but that distinction may not be apparent to the user. Consequently, the user should be able to request computation and receive an automatic display or be able to request display without realizing that this is different from computation. This suggests a command structure in which all commands are equal rather than a hierarchical command structure with a distinct display mode. A further question involves the independence of commands. Our initial design for IMDS emphasized the independence of commands so the researcher could follow any path of immediate interest, but we found the inexperienced user did not know what to do after receiving the results of a command. A command default structure is used to step the novice through the important parts of the normal batch run, while the experienced user can readily depart from the conventional command order. Our default settings may require improvement, but the program became useful to students only when some default structure was implemented.

Documentation needs remain in a conversational environment. The command structure had seemed self-explanatory to the authors, but users have not found it such. The full set of commands is listed at the beginning of a program session, and that set can be displayed again whenever the user issues an appropriate instruction. Additionally, the default structure makes suggestions about command sequencing. However, the meaning of commands and their sequencing for

alternative analytic purposes is still best explained by conventional written documentation.

Some of the above issues may simply reflect the authors' own lack of prior experience in an interactive graphics environment. Yet we suspect that the design problems we encountered are quite general. A self-explanatory graphics routine is probably an illusion. The command structure must take pains to permit novices to benefit from a sophisticated program design. Only exploratory use of the program will show what display techniques are really required. Each of these issues increases the design and programming costs, but that may be an inevitable consequence of moving to a novel computing environment.

CONCLUSIONS

What are the concrete gains in developing interactive graphics analysis programs? An evaluation of the project should carefully distinguish three user audiences: the student, the methodologist, and the researcher.

Students must learn a series of key concepts about multidimensional scaling. Interactive graphics can demonstrate these concepts in an unusually concrete form. The iteration process in multidimensional scaling is best explained by means of an iterative trace such as Figs. 2 and 13. The Shepard diagram of Fig. 10 gives the feeling for the monotonic relationship between interpoint distances and data values that underlies the calculation of the stress coefficient. The local minimum problem cannot be better explained than by the sequence of Figs. 1-5. The effects of alternative starts is seen by contrasting Figs. 3, 5, and 8. The rotation issue is brought to life by Figs. 6 and 9. The ability to move quickly from one solution to another is vital in demonstrating several of these points, so the interactive graphics approach should be particularly vital for student use.

The methodologist is concerned with the effects of alternative strategies. Again interactive graphics permits unusually flexible experimentation with the program. The effects of alternative starts can quickly be assessed. Different solutions to the local minimum problem can be tried. The utility of analytic rotations as opposed to hand rotations can be assessed. The methodologist can try an approach, see where it leads, and revise the scaling whenever that is needed. Particularly useful is the ability to obtain a solution, store that result, try to improve it, and, if the initial improvement attempt fails, the original can be retrieved for further experimentation. A single terminal session can involve as much experimentation as could be done in a few dozen batch runs. We interpret the illustrative session in this paper as demonstrating the utility of rational starts and suggesting that large movements of any points are a good cure and test for local minimum problems, though we expect much more methodological work will be required to specify carefully the conditions under which these solutions are appropriate.

The researcher well acquainted with the technique can benefit considerably from interactive graphics. An amazing amount of exploration of data structure can be performed for less than \$15. If one tried to reproduce an IMDS session through several batch runs, it would prove exceedingly tedious and much more expensive. Perhaps the single most important research gain is that it can make a test for a local minimum a normal part of an analysis rather than a burdensome extra chore which is usually skipped. The researcher studying the party system examined in the illustrative section of this paper would very likely have obtained a solution involving a

local minimum had not the data been analyzed by IMDS.

The interactive approach may provide solutions superior to those that would have been obtained from batch analysis, which means that our tinkering with input/output may have resulted in a pragmatically superior scaling program. We suspect this consequence is more general than it would appear. Interactive graphics so opens up the communication with analysis programs that more efficient analysis strategies can emerge. The analyst prepared to interact dynamically with her/his data may find unexpected rewards.

APPENDIX

The conversational structure of IMDS is illustrated in the following interchange between IMDS and a user. The IMDS prompts are indicated in uppercase, the user's replies are in lowercase. Note that the lines with only a "?" (IMDS outputs the "?" to indicate it is waiting for the user) indicate that the user accepted the default by sending a carriage return.

```
#run icpr:imds
#EXECUTION BEGINS
ARE YOU ON A TYPEWRITER OR CRT?
crt
SUPPLY TITLE FOR OUTPUT < 80 CHARACTERS
? parties data
PRINT OUTPUT ON *PRINT*? YES/NO
? yes
ROUTE? DEFAULTS TO NORTH CAMPUS
?
WRITE DOWN THE FOLLOWING RECEIPT NUMBER
#*PRINT* ASSIGNED RECEIPT NUMBER 613363
READ/INIT/STRE/ITER/ALTE/DELE/SAVE/PRIN/REST/EXPL/
CALC/DEFI/STOP/INPU/WEIG/STAR/HIST/SORT/ACTI/DIST/
RESI/INDI/SHEP/PLAN/DICT/LAST/SSQ/SCAL/ELBO/GRAD/
VARI/PRAX/CENT/NORM/STAN/TURN/ADD/REFL/MODE/ERAS/
COMMAND? DEFAULT IS "READ"
?
MATRIX=STANDARD,NOWEIGHTS,FILE,CUTOFF=(-1,0,1.0),SPLIT=BYDECKS?
?
INPUT MATRIX FILE NAME?
? sw68:parties
COMMAND? DEFAULT IS "INIT"
?
CONF/PRIN/KRUS/RAND=INTEGER
USE RANDOM=0?
? kruskal
NUMBER OF DIMENSIONS? DEFAULT=2
?
COMMAND? DEFAULT IS "ITER"
? center
COMMAND? DEFAULT IS "NORM"
?
COMMAND? DEFAULT IS "PRAX"
? mode
NOECHO/MATRICES/SOMEGRAPHICS/FULLGRAPHICS
? full
ACTIVE, PLANE=(1,2),NUMBERS,MAXIMUM=2.0,DICTIONARY?
? names
```

At this point the screen would contain exactly what is in Fig. 1.

REFERENCES

- Chang, J. J. On-line multidimensional scaling programs with interactive graphical display. *Behavior Research Methods & Instrumentation*, 1973, 5, 99-103.
- Coombs, C. A *theory of data*. New York: Wiley, 1964.
- Daalder, H., & Rusk, J. G. Perceptions of party in the Dutch parliament. In S. C. Patterson and John C. Wahlke (Eds.), *Comparative legislative behavior*. New York: Wiley, 1972. Pp. 143-198.
- Guttman, L. A general nonmetric technique for finding smallest coordinate space for a configuration of points. *Psychometrika*, 1968, 33, 469-506.
- Kruskal, J. B. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 1964, 29, 1-27.
- Rusk, J. G., & Weisberg, H. F. Perceptions of presidential candidates: Implications for electoral change. *Midwest Journal of Political Science*, 1972, 16, 388-410.
- Spence, I. A Monte Carlo evaluation of three nonmetric multidimensional scaling algorithms. *Psychometrika*, 1972, 37, 461-486.
- Weisberg, H. F., & Rusk, J. G. Dimensions of candidate evaluation. *American Political Science Review*, 1970, 64, 1167-1185.

Behavior Research Methods & Instrumentation
1974, Vol. 6, No. 2, 194-200

NOTES

1. Such analysis is reported in Weisberg and Rusk (1970) and Rusk and Weisberg (1972).
2. The formal scaling problem is elegantly stated by Coombs (1964).
3. Nonmetric multidimensional scaling is presented in Kruskal (1964) and Guttman (1968).
4. For example, see Spence (1972).
5. Daalder and Rusk (1972) provide an example of multidimensional scaling of party spaces in the 12-party Dutch setting.
6. The CENTER and NORMALIZE commands were issued after the INITIAL command to facilitate display of the configuration, but these have no effect on the scaling.
7. When all correlations are positive, we have found that dropping the first general component yields a superior starting configuration. IMDS has this capability but we have not used it in this illustrative session, since evaluations of the left- and right-wing parties are negatively correlated.
8. The terminal user cannot see the solutions of Figs. 6 and 9 simultaneously and, therefore, could not determine where the differences between them occur. Addition to the program of an algorithm to compare spatial configurations would solve this problem.
9. Normally the prompting employed with interactive programs must be terse. We have found, however, that this applies only to nongraphic terminals. The speed of graphics terminals eliminates the terseness requirement.

The Minnesota interactive statistical system

RONALD E. ANDERSON, JAMES CLEARY, JONATHON GROSS,
DON McTAVISH, RICHARD MELSON, and JOHN VINCENT
Department of Sociology, University of Minnesota, Minneapolis, Minnesota 55455

The Minnesota Interactive Statistical System (MISS) is reviewed as a successful application of on-line minicomputing in an instructional and research context. The MISS system is described with regard to (1) the instructional and research needs which prompted its development, (2) its programming design, (3) its operating system, (4) its computer hardware requirements, and (5) the continued innovative minicomputing development that it has inspired.

INTRODUCTION

The Minnesota Interactive Statistical System (MISS) is an integrated system of statistical programs for the analysis of social science data through minicomputing. The system was developed by the Sociology Department of the University of Minnesota as an educational and research facility. It was designed to maximize both ease of use and flexibility of analysis, so that both the novice user and the experienced researcher would find interaction with MISS a rewarding experience. The MISS system has some general resemblance to a number of other interactive instructional systems, such as IMPRESS, ISIS, and SIPS (Anderson, 1971; Anderson & Coover, 1972; Cline & Meyers, 1970). It is particularly unique in being operable on a small minicomputer.

As an integrated and interactive system, MISS permits the analysis of a data set to be very flexible. The user is able to develop his statistical analysis in a stepwise manner, viewing the results of one statistical analysis subroutine before deciding which other analysis to

request next; and the output of some analysis subroutines may be optionally used as input for others (for example, the correlation matrix generated from raw data input by subroutine CORR may be stored in memory of MISS for use in higher level correlational analyses). Meanwhile, the rapid feedback of the interactive system permits the user to maintain a close awareness of his data, while it speeds the learning process and makes the task more interesting.

The statistical analysis programs and the possible paths of sequential analysis are represented in the schematic overview of MISS in Fig. 1. As indicated by the diagram, data may be entered into MISS in any of three basic forms: as raw data, as a frequency distribution, or as a matrix (either a correlation matrix, a bivariate table matrix, or a factor matrix). The arrows leading from a data type to the blocks representing the statistical analysis subroutines indicate which programs are compatible with each of the data types. Thus, it is possible to analyze raw data (on more than one variable) with any program in the system, provided that the raw