

## COMPARISONS OF SEVERAL GRAPHICAL METHODS FOR REPRESENTING MULTIVARIATE DATA

C. G. HAMNER, D. W. TURNER and D. M. YOUNG  
Baylor University, Waco, TX 76798, U.S.A.

(Received 4 November 1986)

Communicated by P. L. Odell

**Abstract**—Graphical displays provide a powerful tool for presenting and studying many types of data. This article presents an evaluation of several special graphical methods for representing multivariate data, including three face-type methods and a function-plot method. The evaluation utilizes a split-plot-factorial experimental design. Under the conditions of this experiment, the modified Chernoff-face data-representation method is clearly superior.

### 1. INTRODUCTION

Graphical-data-representation methods provide powerful tools for data display and description, analysis and interpretation, summarization and exploration. Examples of graphical methods which are easy to understand, use, and interpret are well-drawn histograms, stem and leaf plots and bar charts. Other methods, many of which are practical only through computer graphics, are considerably more complicated. Examples of these more complex procedures include certain applications of color graphics (e.g. map shading), stereo graphic projections, trees, human-like faces and function plots.

Although many researchers have developed new graphical-data-representation techniques, only Lambertina and Louv [1], Flury and Riedwyl [2] and Moriarity [3] reported results on the fundamental question of comparative performance of selected graphical methods. Specifically, only Flury and Riedwyl [2] employed a quasi-experimental design in an effort to contrast the information-perserving capabilities of three face-type data-representation techniques. This lack of planned experiments to adequately evaluate the efficacy of graphical-data-representation methods is similar to the lack of experimental-design methods employed in simulation experiments as reported by Hauck and Anderson [4].

To partially remedy this void, we applied a split-plot-factorial design to two separate experiments on two real-data sets to evaluate the performance of five face-type graphical representations for clustering purposes while controlling for the order of presentation. We limited the number of graphical representations for clustering purposes while controlling for the order of presentation. We limited the number of graphical representations in our experiments to six in the first experiment and four in the second experiment to insure manageable experiment sizes. We analyzed four measures of clustering accuracy as performance criteria.

The analyses and contrasts discussed in this paper utilized two data sets previously analyzed by Flury and Riedwyl [2]. For the restricted conditions of these experiments, the modified Chernoff-face method [5] yielded superior clustering accuracy. This result seemingly contradicts the results of Flury and Riedwyl [2]. Later we describe in detail the two data sets, the experimental design, the competing graphical-data-representation methods, and the evaluative criteria.

We emphasize that our goal is to provide a carefully run experiment to study the relative performance of several well-known and complex methods of representing multivariate data. Unlike Cleveland *et al.* [6] or Cleveland *et al.* [7], who have done experimentation addressing very basic details of graphical perception, our experiments do not formally attempt to unravel the individual components that make each of the complex displays "tick". We are studying the actual relative performance of complicated graphical transformations without trying (in this experiment) to isolate individual factors. Our emphasis is on studying the final relative utility of a class of complicated

graphical procedures involving many single geometric substructures interacting in ways yet to be described or predicted by any known theory of graphical perception.

### 2. THE COMPETING GRAPHICAL-DATA-REPRESENTATION METHODS

Flury and Riedwyl [2] contrasted the clustering effectiveness of their asymmetric and symmetric faces, the Chernoff [8] face and a numerical method on data for 20 pairs of 10- to 11-year-old twins: 10 pairs of monozygotic (identical) twins and 10 pairs of dizygotic (fraternal) twins. The data set consisted of 17 anthropometrical variables for each twin; the actual data are available in Hamner [9]. Flury and Riedwyl's symmetric face and the Chernoff face utilized one symmetric face for each twin, and Flury and Riedwyl's asymmetric face utilized one asymmetric face for each pair of twins.

Our experiments, although similar in purpose to the Flury and Riedwyl experiments, entailed much more than simply replicating their experiment. In addition to Flury and Riedwyl's symmetric and asymmetric faces, we examined the performance of a modified version of the Chernoff face developed by Turner and Hall [5] and the Turner and Tidmore [10] line-printer face. We also incorporated one non-face graphical display, the polar Fourier method, into our study. It is a variation of Andrews' function plots [11], in which a finite Fourier series is plotted in polar coordinates. We included the polar Fourier method based on its superior performance over competing function plots reported in Mezzich and Worthington [12]. Hence, we included the following graphical methods of data representation in our experimental design:

- Method 1: polar Fourier plots (PF)
- Method 2; modified Chernoff faces (MC)
- Method 3: Flury and Riedwyl symmetric faces (FRS)
- Method 4: Flury and Riedwyl asymmetric faces (FRA)
- Method 5: symmetric line-printer faces (LPS)
- Method 6: asymmetric line-printer faces (LPA).

Methods 2, 3 and 5 used one symmetric face for every twin; Methods 4 and 6 required only one asymmetric face for every twin pair. Method 1 utilized overlapping polar Fourier plots, one plot for each twin pair. Methods 5 and 6 were essentially asymmetric faces and were limited to a maximum of 12 variables. Through factor analysis we reduced the number of variables to five for inclusion in the symmetric line-printer face. The asymmetric line-printer face for the twin pair held two features constant and utilized separate right and left features to create a total of 10 distinct features.

As an historical note, although Flury and Riedwyl [2] is the first widely publicized application of asymmetric faces, Turner and Tidmore [10] first utilized asymmetric faces for graphical-multivariate-data representation.

It is not practically feasible to show the graphical representation of the entire twin-data set for each of the six graphical-representation methods in this article. Thus, we selected a few figures as examples of some of the basic displays analyzed in the current experiment. In particular, Fig. 1 shows selected modified Chernoff faces for the twin data. Similarly, Fig. 2 displays some of the twin data for the Flury and Riedwyl asymmetric-face method.

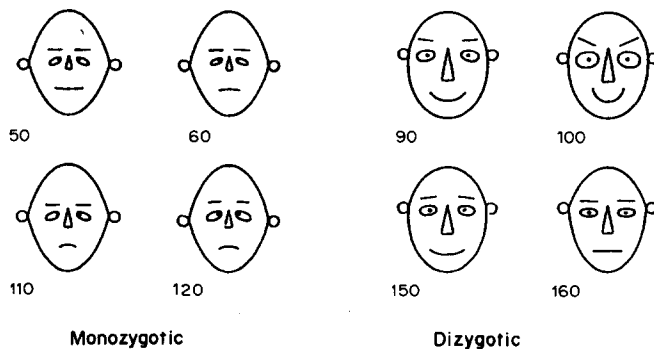


Fig. 1. Twin study: Method 2—modified Chernoff faces of selected monozygotic and dizygotic twin pairs.

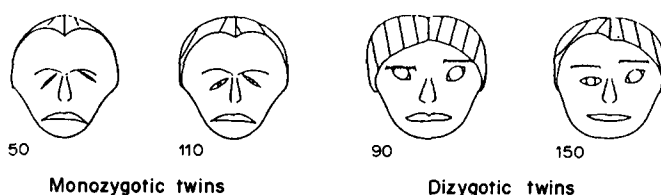


Fig. 2. Twin study: Method 4—selected Flury and Riedwyl asymmetric faces.

We also applied several numerical clustering methods to the twin-data set to provide benchmark information. Four standard hierarchical clustering algorithms were applied to the data. They were single linkage (nearest neighbor), complete linkage (farthest neighbor), average linkage (group-average method) and Ward's minimum variance method. The statistical software used for implementation of the hierarchical methods is referenced in Section 4. In addition, a numerical method called  $S$ , which is fully discussed in Flury and Riedwyl [2], was used. If we let  $d_i$  be the difference between twins on the  $i$ th standardized body measurement, then  $S$  is defined to be the sum of the squares of the  $d_i$ .

### 3. MEASURES OF CLUSTERING ACCURACY

We applied several different measures of clustering accuracy to facilitate the contrast of the six graphical-data-representation methods. The most naive measure of clustering performance was the average proportion of correct classifications,  $Z$ . We calculated this statistic for each of the six methods and contrasted the differences in the mean-correct proportion for several contrasts after testing for equality of the means.

A second measure of clustering accuracy was the Rand statistic,  $R$ , developed by Rand [13]. The Rand statistic has values in  $[0, 1]$  with larger values indicating higher similarity between partitions. For the Rand statistic we compared each partition with the correct partition of twins. For each person performing the clustering task, we computed the Rand statistic for all six methods.

Morey and Agresti [14] formulated a third measure of clustering performance by introducing an adjustment to the Rand statistic. This adjusted  $R$  value, denoted by  $R_{adj}$ , can have negative values. A value of 0 indicates partition similarity no better than random clustering. Values  $> 0$  indicate better than random clustering with 1 as an upper bound. We computed  $R_{adj}$  for each subject/method combination.

The fourth measure of clustering accuracy we applied in our experiment was the relatively new Fowlkes and Mallows'  $B_k$ , found in Fowlkes and Mallows [15].  $B_k$  takes on values in  $[0, 1]$ . This statistic is another measure of the degree of similarity between a subject's partition and the correct partition of twins. The value of  $B_k$  is directly proportional to the clustering-partition accuracy of a subject.

### 4. AN ANALYSIS OF THE PERFORMANCE OF SIX GRAPHICAL-DATA-REPRESENTATION METHODS ON THE TWIN DATA

Ninety-six students enrolled in introductory statistics courses separated the 20 pairs of twins into two groups of 10 each, one group of monozygotics and the other of dizygotics. Each student performed this clustering task for all six methods. Prior to the actual experiment, we held a training session to introduce the students to each of the graphical-clustering techniques. At the data-gathering session, we asked students to appropriately label the observations from the monozygotic group and the dizygotic group.

For this experiment, concatenating the asymmetric and symmetric facial-representation methods of both the Flury and Riedwyl and line-printer faces reduced the order of presentation of fundamentally different face-type-representation techniques to four. Thus, there were  $4!$  orders of presentation, and, therefore, we applied a split-plot factorial for the experimental design. The 24 levels of treatment A correspond to the orders of presentation; the levels of treatment B correspond

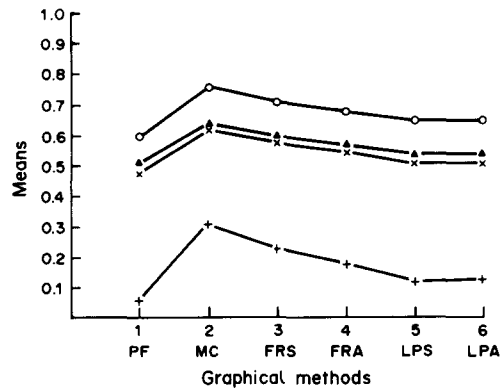


Fig. 3. Twin study: plots of means of evaluative criteria by a graphical method. ○, The proportion correctly classified; △, Rand statistic; +, adjusted Rand statistic; ×, Fowlkes and Mallows  $B_k$ .

to the 6 graphical methods. Subjects received only one level of treatment A but all levels of treatment B.

Regardless of which measure of clustering accuracy we analyzed as the dependent variable, analysis of the associated split-plot-factorial design yielded identical results. We note that the Greenhouse–Geisser conservative  $F$  tests for the effect due to the graphical-representation method did not contradict the conventional test results. Order of presentation was not a significant factor, confirming a similar result from the Flury and Riedwyl experiment. Moreover, the Flury and Riedwyl experiment examined only 6 possible orders of presentation for 3 methods as compared with the 24 orders for the 6 methods in our study. Interaction between the order of presentation and method was also insignificant. However, the test for difference in the means of the graphical-data-representation methods was highly significant ( $P < 0.0001$ ).

For the six graphical-data-representation techniques, we plotted sample means of the four evaluative statistics, as shown in Fig. 3, which revealed that Methods 2, 3, 4 and 1 were ranked first, second, third and sixth, respectively, for all evaluative criteria.

The top three graphical methods in decreasing order of their ability to distinguish between twin pairs were modified Chernoff, Flury and Riedwyl symmetric and Flury and Riedwyl symmetric faces. Polar plots consistently ranked last. Flury and Riedwyl [2] recorded means of the proportion of correctly classified observations for three graphical methods: Chernoff faces and the Flury and Riedwyl symmetric and asymmetric faces. The average proportions correctly classified from the Flury and Riedwyl study and the corresponding means of this study, respectively, were as follows: 0.66 and 0.76 for the modified Chernoff-faces method; 0.73 and 0.71 for the Flury and Riedwyl symmetric-faces method; and 0.71 and 0.68 for the Flury and Riedwyl asymmetric-faces method.

We computed 10 nonorthogonal, *a priori* multiple comparisons using a Bonferroni multiple-comparisons method. We summarized the main results of the multiple comparisons below. The conclusions were identical regardless of which of the four evaluative statistics we chose as the dependent variable in the experiment. We found the following results:

- (1) no significant difference in the means of the two line-printer-face methods ( $P > 0.05$ );
- (2) slight differences in the two Flury and Riedwyl method means ( $0.01 < P < 0.05$ ), with the Flury and Riedwyl symmetric-face method being superior;
- (3) highly significant differences in the five face-type data-representation methods in that the mean of the modified Chernoff-face method was greater than the average of the two Flury and Riedwyl methods and the average of the two Flury and Riedwyl methods was greater than the average of the two line-printer methods;
- (4) highly significant differences between the means of the five facial-graphical-data-representation methods and the mean of the polar Fourier plot in that the means of all five face-type-representation methods were greater than the mean of the polar Fourier plots.

Table 1. Twin study: correlation matrix among the proportions of subjects choosing a twin pair as dizygotic for every graphical method and for the log  $S$  method

Method	PF	MC	FRS	FRA	LPS	LPA
MC	0.32					
FRS	0.54	(0.66) 0.72				
FRA	0.15	(0.79) 0.78	(0.89) 0.75			
LPS	0.16	0.64	0.51	0.44		
LPA	0.21	0.54 (0.67)	0.42 (0.90)	0.32 (0.91)	0.72	
Log $S$	0.45	0.88	0.86	0.79	0.71	0.44

The numbers in parentheses are from Flury and Riedwyl [2].

For every twin pair, we calculated the proportion of subjects assigning the observation to the dizygotic group. Correlations for all six graphical methods among these proportions and log ( $S$ ), which are presented in Table 1, showed that the highest correlations in descending order occurred among log ( $S$ ) and the modified Chernoff-face method (0.88), log ( $S$ ) and the Flury and Riedwyl symmetric-face method (0.86), log ( $S$ ) and the Flury and Riedwyl asymmetric-face method (0.79) and the modified Chernoff-face method and the Flury and Riedwyl asymmetric-face method (0.78). Recall that Flury and Riedwyl [2] defined the clustering-evaluation criterion  $S$ .

The largest discrepancies in the correlations between the 1981 Flury and Riedwyl study and this study occurred in the correlation of the two Flury and Riedwyl methods [0.89 (FR), 0.75 (present study)] and the correlation of log ( $S$ ) and the modified Chernoff method [0.67 (FR), 0.88 (present study)]. This last discrepancy contradicted the 1981 Flury and Riedwyl conclusion that the numerical-clustering technique  $S$  showed much better agreement with the Flury and Riedwyl facial-data-representation versions than with the Chernoff-facial method. The results of our experiment supported a higher correlation of log ( $S$ ) with the modified Chernoff-facial-data-representation method.

In addition to the evaluation of graphical-clustering techniques, we applied five numerical techniques to the twin data, four of which are standard hierarchical-clustering algorithms. We performed single linkage, complete linkage and average linkage with BMDP subprogram PIM, using Euclidean distance [16]. We utilized the program CLUSTAR [17] to implement Ward's clustering method. Information consistent with that known by the experimental subjects applying the graphical methods was a factor in deciding the dendrogram cuts. We also employed a fifth numerical technique using Flury and Riedwyl's  $S$  statistic. We then compared the evaluative statistics for applying  $S$  to form the two clusters with the evaluative statistics obtained through the other numerical and graphical methods. The test statistics showed that all of the numerical-clustering techniques were better than the graphical techniques in recovering the *known* dichotomous groups for all evaluative criteria. The average-linkage algorithm ranked first with only one misclassification per twin-pair group. Single linkage, complete linkage, Ward's method and the  $S$  method tied with only two misclassifications per twin-pair group. The values of the average proportion correctly classified, Rand, adjusted Rand and  $B_k$  for average linkage were 0.90, 0.81, 0.64 and 0.80, respectively. These values should be compared with the results in Fig. 3. Since the numerical techniques performed so well, one might question the utility of the graphical displays to perform clustering. However, in addition to the obvious visual advantages of graphical displays, one can apply the facial-representation techniques to help make an optimal choice among competing numerical clustering algorithms. We refer the reader to Tidmore and Turner [18] for further discussion of this application.

##### 5. AN ANALYSIS OF THE PERFORMANCE OF FOUR GRAPHICAL-DATA-REPRESENTATION METHODS ON THE SWISS-BANKNOTE DATA

In a separate experiment, we asked 96 students (different from the twin-data-experiment subjects) enrolled in introductory statistics courses to separate 40 plots, representing 25 real and 15 forged

banknotes, into groups using each of four graphical-representation methods, as described below. Six measurements were made on each banknote. These 40 banknotes were a subset of the original 200 analyzed in Flury and Riedwyl [2] and were documented in Hamner [9]. For the Swiss-banknote data, the graphical-multivariate-data-representation methods applied to the clustering task were as follows:

- Method 1: polar Fourier plots (PF)
- Method 2: modified Chernoff faces (MC)
- Method 3: Flury and Riedwyl asymmetric faces (FRA)
- Method 4: Line-printer asymmetric faces (LPA).

In a training session we introduced the students to each of the four methods of graphical-data representation. This experiment differed from the previously discussed twin-data experiment in that we did not inform the students for this experiment about the nature of the data, and we allowed them to choose the number of groups (four maximum) and the size of each group.

We assigned variables to face parameters (method parameters) to maximize the visual distinction between groups, as in the 1981 Flury and Riedwyl experiment. Figure 4 shows the graphical representations of two typical real banknotes and two typical forged banknotes for each of the four graphical techniques applied to the Swiss-banknote data.

Two split-plot-factorial analyses, using the Rand statistic and the adjusted Rand statistic revealed surprisingly similar results to the previously discussed twin study. The Greenhouse–Geisser conservative  $F$  test was the second step in testing the hypothesis of no difference in the means of the clustering-accuracy criteria associated with the four graphical-data-representation methods described in Section 5. The Greenhouse–Geisser tail probabilities for the  $F$  tests did not contradict the results of the conventional tests. Order of presentation and interaction of order and method were insignificant (all  $P$  values  $> 0.13$ ). The difference in means of the four graphical methods was highly significant ( $P < 0.0001$ ).

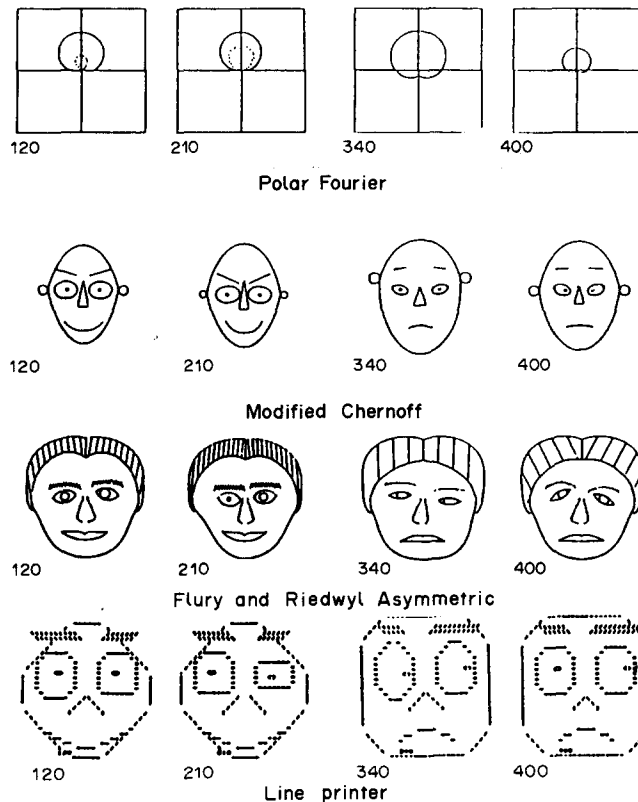


Fig. 4. Swiss-banknote study: Method 1—polar Fourier, Method 2—modified Chernoff faces; Method 3—Flury and Riedwyl asymmetric faces; Method 4—Line-printer faces. Plots 120 and 210 are real banknotes, while plots 340 and 400 are forged banknotes.

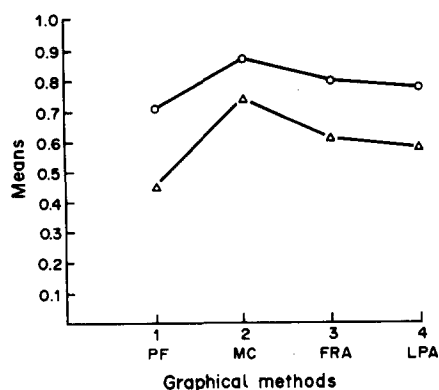


Fig. 5. Swiss-banknote study: plots of means of evaluative criteria by a graphical method. ○, Rand statistic; △, adjusted Rand statistic.

The means of the graphical methods as shown in Fig. 5 yielded the same rank for each of the evaluative criteria: modified Chernoff first, Flury and Riedwyl asymmetric second, line-printer asymmetric third and polar Fourier plots fourth. Combining the two versions of the Flury and Riedwyl faces and the two versions of the Turner and Tidmore line-printer faces yielded the identical rank order of the ability to recover natural clusters as found in the authors' twin-data experiment discussed in Section 4. Multiple comparisons of pairwise differences in the means and the average of the means of the face-type methods minus the polar Fourier method mean indicated significant differences for all comparisons except the means of the evaluative statistics for the Flury and Riedwyl asymmetric-face method and the line-printer asymmetric-face method. For the banknote data as well as the twin data, the three face-type methods outperformed the one polar plots method in distinguishing between groups.

## 6. DISCUSSION

The major purpose of the two experiments discussed in this article is to contrast and evaluate several well-known graphical methods for representing multivariate data. The results of the experiments conducted on two major data sets (the twin data and the Swiss-banknote data) clearly identify modified Chernoff faces as the best of the four graphical methods in recovering the known dichotomous groups. Then for both data sets, the Flury and Riedwyl faces, the line-printer faces and the polar Fourier plots consistently ranked second, third and last, respectively, for both evaluative criteria applied to measure clustering accuracy.

We conjecture that at least two major reasons exist to explain why the modified Chernoff face is clearly the best of the three face-type graphical methods for recovering the known dichotomous groups: feature detail and dimension size. First, the modified Chernoff version has all of the benefits of the familiarity of faces in a cartoon-like face. The line-printer face is a much cruder caricature and suffers from the lack of rounded features. Thus, it is the least face-like of the three versions in terms of feature details. Both types of Flury and Riedwyl faces more closely resemble human faces in their appearance. However, the added intricacy of detail of the facial features does not appear to increase the benefit of the familiarity of a face-type method, and in some cases the detail may hinder the subject's performance of the clustering task.

The second reason for the superiority of the modified Chernoff face is that the 20 dimensions of the modified Chernoff face may be closer to the ideal amount of visual information the human brain can assimilate. The 36-dimensional Flury and Riedwyl asymmetric face may be too complicated. The 18-dimensional symmetric Flury and Riedwyl version, in combination with the detailed features, may not be optimal in conveying information but may be slightly better than its asymmetric counterpart. The 12 dimensions of the line-printer face, in combination with fewer face-like features, may be oversimplified.

We should note that the modification of the original Chernoff face applied in this experiment study is not identical to the Chernoff face utilized by Flury and Riedwyl. Our version has 20

dimensions; Flury and Riedwyl's version has 18 dimensions. However, the differences in the dimensionality and the feature complexities of the two Chernoff faces are relatively small contrasted to the differences between the Chernoff faces and the competing methods listed in Section 2. Thus, certain discrepancies between the Flury and Riedwyl experiment and our experiment may occur, in part, because of the differences in the dimensionality, features, and complexity of the features in the two Chernoff-face versions. However, it seems highly unlikely that these differences could account for the gross difference in the results of the Flury and Riedwyl experiment and our experiment.

Overlapping polar plots, apparently contrasted to competing graphical-data-representation methods for the first time in our experiment, are appropriate in the twin-data experiment because of the unique nature of the data. One group contains those pairs of plots which appear to look most alike (identical twins). The other group contains pairs of plots which appear to look the least alike (fraternal twins). Overlapping the polar plots is superior to two adjacent plots because memory of the functional arrangements is particularly difficult, and size and position are important in judging similarities between pairs of plots. Overlapping polar plots would provide an interesting application for color graphics (distinct colors for each twin in an overlapping pair), and we conjecture that this addition would significantly improve their performance. Also, we believe that more experience with unfamiliar polar plots would improve their performance relative to the familiar face-like displays.

The effectiveness of polar plots is extremely sensitive to the order of variable assignment. It is difficult to obtain informative plots for the clustering task because arbitrary variable assignments create "noisy" plots. For this reason, researchers often apply a data-reduction technique to order transformed variables for assignment to the coefficients of the polar plots [12]. For the Swiss-banknote experiment, we chose factor analysis as the data-reduction method because it not only reduces the dimension but also provides an ordering of the factors for variable assignment.

One of the difficulties involved in studying graphical methods of the type presented here is that designed experiments for evaluation of graphical-multivariate-data-representation methods require human subjects. Training sessions and experimental data-collecting sessions require high concentration and a significant time commitment from each subject. These demands put substantial limitations on the number of factors one can practically analyze. Nonetheless, the two experiments described in this article are large when contrasted to other experiments concerning this topic, and the results provide comparative information obtained under controlled conditions.

## REFERENCES

1. W. J. F.-T. Lambertina and W. C. Louv, Comparisons of some graphical methods for exploring multivariate data analysis. *Am. Statist.* **38**, 184-188 (1984).
2. B. Flury and H. Riedwyl, Graphical representation of multivariate data by means of asymmetrical faces. *J. Am. statist. Ass.* **76**, 757-765 (1981).
3. S. Moriarity, Communicating financial information through multidimensional graphics. *J. Account. Res.* **1**, (17) 205-224 (1979).
4. W. W. Hauck and S. Anderson, A survey regarding the reporting of simulation studies. *A. Statist.* **38**, 214-215 (1984).
5. D. W. Turner and K. A. Hall, GRAPHPAK: interactive graphics for analysis of multivariate data. In *Proc. 15th Symp. on the Interface of Computer Science and Statistics*, pp. 328-332. North-Holland, Amsterdam (1983).
6. W. S. Cleveland, P. Diaconis and R. McGill, Variables on scatter plots look more highly correlated when the scales are increased. *Science* 1138-1141 (1982).
7. W. S. Cleveland, C. S. Harris and R. McGill, Judgment of circle sizes on statistical maps. *J. Am. statist. Ass.* **77**(379), 541-547 (1982).
8. H. Chernoff, The use of faces to represent points in  $k$ -dimensional space graphically. *J. Am. statist. Ass.* **68**, 362-368 (1973).
9. C. G. Hamner, A statistical evaluation of several graphical methods for representing multivariate data. Unpublished Ph.D. Dissertation, Baylor Univ., Waco, TX, U.S.A. (1983).
10. D. W. Turner and F. E. Tidmore, Clustering with Chernoff-type faces. In *Proc. Statistical Computing Section, American Statistical Association*, pp. 372-377 (1977).
11. D. F. Andrews, Plots of high dimensional data. *Biometrics* **28**, 125-136 (1972).
12. J. E. Mezzich and D. R. L. Worthington, A comparison of graphical representations of multidimensional psychiatric diagnostic data. In *Graphical Representation of Multivariate Data* (Edited by P. C. C. Wang). Academic Press, New York (1978).
13. W. M. Rand, Objective criteria for the evaluation of clustering methods. *J. Am. statist. Ass.* **66**, 846-850 (1971).
14. L. C. Morey and A. Agresti, The measurement of classification agreement: An adjustment to the Rand statistic for chance agreement. *Classif. Soc. Bull.* **5**, 9-10 (1981).



15. E. B. Fowlkes and C. L. Mallows, A method for comparing two hierarchical clusterings. *J. Am. statist. Ass.* **78**, 553–569 (1983).
16. W. J. Dixon (Ed.) *BMDP: Statistical Software*. Univ. of California Press, Berkeley, Calif. (1981).
17. C. H. Romesburg and K. Marshall, CLUSTAR and CLUSTID: Computer programs for hierarchical cluster analysis. *Am. Statist.* **34**, 186 (1980).
18. F. E. Tidmore and D. W. Turner, On clustering with Chernoff-type faces. *Commun. Statist.* **A12**, 381–396 (1983).

## APPENDIX

We have generated all of the graphical representations of the multivariate data contrasted in this paper using the specialized interactive package GRAPHPAK. Turner and Hall [5] described this package in detail. GRAPHPAK, written by Danny Turner and Keith Hall, is available through IMSL Distribution Services in Houston, Texas. CALCOMP hardware and software is the GRAPHPAK option chosen for plotting the graphical displays in this article.