

# A Monte Carlo Examination of External Unfolding

Paul Thompson  
Ohio State University

Monte carlo techniques were used to examine regression approaches to external unfolding. The present analysis examined the technique to determine if various characteristics of the points are recovered (such as ideal points). Characteristics of the situations were manipulated, including number of stimuli, error level, and measurement level. Generally, monotonic analyses resulted in good recovery. Recovery was poor when the data were generated by a weighted Euclidean process. Negative weights were commonly encountered, apparently due to error in the data. In this approach, estimation is done by statistical techniques, and some statistical concerns should be taken into account when examining the results. *Index terms: external unfolding, monte carlo simulation, multidimensional scaling, nonmetric methods, parameter estimation, unfolding.*

Unfolding is one of a number of techniques used for the analysis of rectangular or two-way, two-mode data (Bennett & Hays, 1960; Carroll, 1972, 1980; Coombs, 1964; Gold, 1973; Heiser, 1981; Schönemann, 1970). The entries in the data matrix are measures of relation between the row and column stimuli. For clarity, the row stimuli will be referred to as "persons," while the column stimuli will be referred to as "objects." These could be preferences, ratings, or dominance values. The matrices can be analyzed to recover the parameters for persons and objects (Coombs, 1964; Heiser,

1981; Schönemann, 1970), or used to interpret stimulus spaces (Schiffman, Reynolds, & Young, 1981) in an objective fashion.

## Models for Unfolding

Several models have been used in the unfolding situation. These models generally represent persons and objects as points in a multidimensional metric (usually Euclidean) space (Bennett & Hays, 1960; Heiser, 1981; Schönemann, 1970). Relations between persons and objects are often expressed as a squared Euclidean distance:

$$d_{ij}^2 = \sum_{k=1}^K w_{ik}(x_{jk} - y_{ik})^2 \quad (1)$$

where  $k$  is the index for  $K$  dimensions,  
 $x_{jk}$  are the coordinates for object  $j$ ,  
 $y_{ik}$  are the coordinates of person  $i$  on dimension  $k$ ,  
 $w_{ik}$  is the weight for person  $i$ , and  
 $d_{ij}^2$  is the relation between person  $i$  and object  $j$ . ( $d_{ij}^2$  will represent this theoretical relationship, while squared distances will be assumed throughout. If preferences or similarities are used, they must be converted to dispreferences.)

Equation 1 implies elliptical iso-preference contours; that is, those points  $x_j$  which are equally preferred by ideal point  $y_i$  form a hyper-ellipse.

Relations between persons and objects may also be modeled with a linear model (Tucker, 1960):

$$d_{ij}^2 = \sum_{k=1}^K b_{ik}x_{jk} \quad (2)$$

where  $b_{ik}$  is the importance of dimension  $k$  for person  $i$ . This approach does not postulate an "ideal point" but rather a preferred direction, along which preference increases without limit. Iso-preference contours are straight lines, perpendicular to the preferred direction.

### External Unfolding

In external unfolding, coordinate values for the objects are known. These values are either stated a priori from theory, or recovered from the rectangular proximity matrix during an earlier step, such as the successive unfolding process (Rodgers & Young, 1981; Thompson, 1988). The approach incorporates the persons into the space by determining the appropriate model for the relationships between persons and objects as well as the specific coordinates for the objects.

Carroll's (1972, 1980) hierarchy of nested models may be used for this process. This hierarchy is examined here, with the exception of Model I. Although Model I is a logical generalization of Model II, it has seldom been found useful above and beyond the simpler models.

*Model IV: The vector model.* This is a scalar products model (Equation 2). In this model, preference increases without limit, which may not be a reasonable description of actual behavior. There is no ideal point for persons, although an infinitely distant ideal point can be postulated (Carroll, 1980); such a concept represents a mathematical rather than an interpretable notion.

*Model III: The circular ideal point model.* The remaining models assume that preferences are single-peaked (Coombs, 1964). The "ideal point" is the location which is most preferred by the person. If  $w_{jk} = 1$ , this model is stated in Equation 1. The coordinates for the inserted points may be estimated:

$$y_{ik} = \frac{-b_{ik}}{2a_i} \quad (3)$$

where

$$\delta_{ij}^2 = \sum_{k=1}^K b_{ik}x_{jk} + a_i \left( \sum_{k=1}^K x_{jk}^2 \right) + c_i \quad (4)$$

(Carroll, 1980, pp. 250–251). Here,  $\delta_{ij}^2$  represents the observed dispreference value, and  $c_i$  includes an error term. In Model III, all dimensions are equally weighted, making iso-preference contours circular.

*Model II: The weighted ideal point model.* In Model II, the assumption of equally weighted dimensions is relaxed, and iso-preference contours are elliptical. The weighted ideal point model is stated in Equation 1. Estimates for the unknown coordinates may be obtained:

$$y_{ik} = -\frac{b_{ik}}{2a_{ik}} = -\frac{2w_{ik}y_{ik}}{2w_{ik}} \quad (5)$$

where

$$\delta_{ij}^2 = \sum_{k=1}^K b_{ik}x_{jk} + \sum_{k=1}^K a_{ik}x_{jk}^2 + c_i \quad (6)$$

*Comparisons and contrasts.* These models are simple approaches to the estimation of parameters to interpret preferences/dissimilarities. One simple linear model can represent Models IV, III, and II. Equation 6 represents Model II, while the other models are obtained by restriction in the parameter estimates for the model. Model III may be obtained from Model II by requiring

$$a_{ik} = a_{ik'} \in k, k' \quad (7)$$

Model IV may be obtained from Model II by requiring

$$a_{ik} = 0 \in k \quad (8)$$

### Parameter Estimation and Model Selection

These different approaches to external unfolding may be implemented in a number of ways. Standard least-squares methods (regression approaches) may be used to estimate parameters (Carroll, 1980). In some cases, negative weights are obtained for Model II. Negative weights may be interpreted as "anti-ideal" points (Carroll, 1972), or least-preferred points.

The linear regression approach is simple to implement, easy to use, and understandable. These methods have been implemented as strictly metric approaches; that is, a linear relation between data and disparity values is assumed. The linear regression approach may be easily generalized to the nonmetric situation, using an alternating least-squares approach to linear regression (Young, de Leeuw, & Takane, 1976). Thus, the regression approach is very general.

If negative weights are not desired, other approaches must be used. Davison (1976a, 1980) used a constrained quadratic programming technique to obtain a solution with the weights restricted to be positive, while Srinivasan and Shocker (1973) used a linear programming method (LINMAP) to accomplish much the same purpose. While these are interesting alternatives, the standard linear approach is probably the method most often used by persons interested in external unfolding.

#### Estimation Using Linear Regression

*Quadratic regression.* Equation 6 represents a quadratic or polynomial regression problem. In the full model for Equation 6, each dimension (represented by values  $x_{ik} \in k$ ) is presented both as a single value and as a squared value. The collinearity problems inherent in such polynomial regression problems are well known (Neter, Wasserman, & Kutner, 1985, p. 301). Collinearity causes several difficulties. First, the standard errors of coefficients are much larger. Second, the signs of coefficients may be affected by the inherent variability. Because the matrix  $\mathbf{X}$  is usually columnwise centered, this should be a very minor problem.

*Ratios of parameter estimates.* A second problem arises in obtaining the estimates of  $y_{ia}$ , using Equation 5. Both coefficients involved in such computations are point estimates with inherent degrees of variability. Point estimates must be examined in the context of their standard errors (Neter et al., 1985), and this is difficult to do when ratios are formed from other point estimates. The problem is magnified by placing the term with the largest standard error (the coefficient of the squared

term) as the denominator of the ratio. In such cases, the ratio is likely to be quite unstable.

*Model selection.* Methods for selecting the correct model are not often discussed, but PREFMAP (Carroll, 1972, 1980) compares models using incremental  $R^2$  tests. Such tests assume an additive error term (see Equation 9). It must also be assumed that this error term is normally distributed. These assumptions are strong and have some problems (e.g., error-perturbed distances may be negative). However, in considering this problem, Ramsay (1977, p. 245) stated that "this is usually very small" in the usual case.

Takane (1981) also considered normal errors for distances. Normal errors for distances were also used for the current analysis. Because this assumption is made in the Gauss-Markov approach to the general linear model (Timm, 1975), standard tests of significance for the values of  $R^2$  may be used to determine fit.

#### Conclusions

The external unfolding models are very interpretable models for preferences. However, the method for fitting the models depends on a number of assumptions. Additionally, the effect of error on the process is not known. As the technique is frequently used in data analysis (Carroll, 1980; Green & Wind, 1973; Schiffman et al., 1981), answers to these questions are of practical interest.

The inherent variability of parameter estimates must be considered in any evaluation of the meaning of the estimates. In general, users of regression are often counseled not to interpret parameter estimates, but rather to examine the significance of the predictor variables, the size of the standard error, and the confidence intervals for estimating the parameter. Do these concerns for regression estimates apply to external unfolding using regression methods?

#### Monte Carlo Examinations of Unfolding

This question was evaluated using monte carlo methods. Other evaluations have examined the model

selection problem for extant datasets (Davison, 1976b); such examinations have face validity, but are not really helpful in examining a technique, as the true situation is unknown. A monte carlo analysis affords the opportunity to study a technique while knowing all of the important information.

### Error in Monte Carlo Evaluations

*Monte carlo evaluations.* It is often difficult to interpret the error variances in monte carlo analyses, as error is often defined in an arbitrary or ad hoc fashion. In order to interpret error variance, the variance of the data before the error is added should be compared to the variance after the error is added. One interpretable method for examining error is as a reliability, defined as the ratio of "true" to "observed" (error-perturbed) variance (Lord & Novick, 1968).

*An overall model for error.* The different forms of error involved in a dissimilarity may be viewed as

$$d_{ij}^{2*} = \sum_{k=1}^K w_{ik}(y_{ik} - x_{jk})^2 + \varepsilon_{mij} \quad (9)$$

where  $\varepsilon_{mij}$  is an error associated with the dissimilarity between person  $i$  and object  $j$  for the  $m$ th comparison. This is a variant of the Hefner (1958) model for multidimensional dissimilarity judgments (Zinnes & MacKay, 1983). For this discussion,  $\varepsilon_{mij} \approx N(0, \sigma^2)$ . The assumption of normality simplifies the error addition process, and it is needed to ensure the correctness of significance testing using incremental  $R^2$  approaches.

### Method

A monte carlo experiment was run to examine the ability of the external unfolding analysis to recover important characteristics of the situation.

*Data generation.* Data were generated and analyzed using the Statistical Analysis System (SAS Institute, 1986). Each dataset consisted of three types of data: the object points, the person points, and weights for the person points. There were varying numbers of object points. The person points consisted of 20 points to be inserted into or related

to the object points. Person and object points (or importance directions, as the case may be) were generated using standard normal random numbers. Weights were generated using uniform (0,1) random numbers. All three components (inserted points, fixed points, weights) were initiated from different seeds. A SAS matrix macro which performs external unfolding (PREF; Thompson, 1986) was used for estimation.

*Model type.* The data for analysis were generated using either Model IV, III, or II. For Model IV, the data were scalar products; for Models III and II, the data were squared dissimilarities. The discussion will concern estimates obtained under the same model from which the data were generated (the Model Type factor, 3 levels—numbered 2, 3, and 4 below).

*Number of fixed points.* Matrices were generated with either 8, 16, 25, 35, or 50 fixed points (the Fixed Points factor, 5 levels).

*Error level for data.* Error was added to the preferences using a normal error process. Let  $\sigma_e$  represent the standard deviation of the errors, and let  $\sigma_o$  represent the standard deviation of the observed values prior to error perturbation. The factor was defined by setting  $\sigma_e/\sigma_o = 0.0, .05, .10, .20, .40, \text{ or } .80$  (the Error Level factor, 6 levels).

*Data categorization.* After error was added to the data, the data were either categorized or not. The dissimilarities were categorized by rounding the values. No effort was made to retain equal numbers of categories, or categories with equal numbers of values. Values more than 3 standard deviations from the mean were truncated to 3 standard deviations. This categorization represents a monotonic function of the original dissimilarity, precisely the type of function that monotonic regression approaches (Kruskal, 1964) are designed to recover. Cases were called Uncategorized or Categorized (the Categorization factor, 2 levels).

*Measurement level.* The data were analyzed using an interval-level (linear) or ordinal-level (monotonic) analysis. In the latter, analysis was done using Kruskal's (1964) primary (ties untied optimally) approach (the Measurement Level factor, 2 levels).

### Design Summary

Two replications were run for each of the cells defined by the factors given above. Overall, there were 2,160 treatments, as the design consisted of 3 (levels for the Model Type factor)  $\times$  5 (Fixed Points)  $\times$  6 (Error Variance)  $\times$  2 (Categorization)  $\times$  2 (Measurement Level)  $\times$  2 (Replication) cells. For the analysis, the Model Type, Fixed Points, Error Variance, and Categorization factors were completely randomized factors, while the Measurement Level factor was a randomized block factor. For some analyses, each matrix was analyzed under all models in order to determine some effects on model selection and frequency of negative weights. However, this was the exception.

### Interpretation of Error

Mean reliability values for the different levels of Error Level (in the Uncategorized cases) ranged from .997 (Error Level = .05) to .612 (Error Level = .80). With the exception of Error Level = .80, the reliability values are roughly equivalent to their theoretical values (obtained from  $\rho = \sigma_t/\sigma_x$ , where  $t$  represents the true score and  $x$  the observed score). The data included a range of situations encountered in well-designed experiments.

### Statistics for Recovery and Fit

*Correlation  $r_3$ .* The matrix correlation  $r_3(\mathbf{R}, \mathbf{T})$  (Ramsay, ten Berge, & Styan, 1984) between true inserted set points (represented by matrix  $\mathbf{T}$ ) and recovered coordinates (represented by matrix  $\mathbf{R}$ ) is defined as

$$r_3 = \text{tr}(\mathbf{V}_R \mathbf{U}_R' \mathbf{U}_T \mathbf{V}_T'), \text{ s.t. } \mathbf{R} = \mathbf{U}_R \mathbf{D}_R \mathbf{V}_R', \quad (10)$$

where  $\mathbf{U}_R$ ,  $\mathbf{D}_R$ , and  $\mathbf{V}_R'$  are components of the singular-value decomposition of  $\mathbf{R}$ . This correlation is not sensitive to dispersion (as measured by singular values), but does assess similarity in point shape (Ramsay et al., 1984). The correlation is close to 1 if the matrices have similar orientations and shapes. In this case, recovery is assessed with a fixed orientation because the stimulus points are fixed.

*Model identification.* The best model was selected using incremental  $R^2$  tests between different models. The simplest model which was significant at  $\alpha = .01$  (and for which no better models existed, determined by the incremental  $R^2$  test using  $\alpha = .01$ ) was chosen as the best model. The  $R^2$  of the correct model was also examined.

*Number of negative weights.* For the analysis of each dataset, Model II was used to examine the frequency of negative weights. Both the number of such negative weights and their size are interesting, and will be examined.

## Results

### Recovery of Structure

The analysis of the simulation results concentrated primarily on presentation of means. The factors manipulated were examined primarily to determine appropriate places for collapsing means. Several different multivariate analyses of variance (multivariate split-plot analysis; O'Brian & Kaiser, 1985; Timm, 1975) were used to examine the measures of recovery and fit (Equation 10 and  $R^2$ , respectively). The analysis examined all factors and interactions. The multivariate five-way interaction was not statistically significant, nor were any of the four-way interactions. Only 2 of the remaining interactions (1 of 6 three-way, and 1 of 4 two-way interactions) of any variable with Measurement Level were significant. The other two- and three-way interactions (of the remaining four variables) were almost all significant.

In order to present means, the Categorization and Measurement Level factors were collapsed into a single factor called the Measurement/Categorization factor. Analyses discussed below are the analysis of the uncategorized data using the interval approach (the Uncategorized/Linear level) and the categorized data using the monotonic approach (the Categorized/Monotonic level). Means are presented in the simplest possible manner, collapsing over some significant (but practically unimportant) interactions.

*Fit measure:  $R^2$ .* The mean  $R^2$  for the 20 inserted points, using the same model to analyze as

was used to generate the data, is examined first. Table 1 presents mean  $R^2$  values for the Uncategorized/Linear cases (Categorized/Monotonic results were similar). These values are mean adjusted  $R^2$  values, to correct for overfitting when few stimuli are used.  $R^2 = 1.0$  when Error Level = .00, thus validating the analysis. The substantial break in  $R^2$  occurs when Error Level drops to .20 from .40. These adjusted  $R^2$  values may be compared to the reliability values, as reliability is a limit to the valid variance in a measure. In this case, the roots of the adjusted  $R^2$  values are roughly equivalent to the reliability values, indicating that the procedure is extracting all meaningful variance from the data.

*Recovery measure: Correlation  $r_3$ .* Results for correlation  $r_3$  (Equation 10) are presented in Table 2. This table presents results for analyses using the same model for analysis as was used to generate the data. Table 2 presents results for the Uncategorized/Linear cases (Categorized/Monotonic results were similar). Additionally, because  $r_3 = 1.00$  when Error Level = .00, these cases are excluded from the table. Other categories are combined as necessary for clarity and conciseness. This correlation ignores scale and is appropriate for all levels of Model Type.

This measure is sensitive to small deviations from an optimal relationship (Thompson, 1988). Table 2 shows that Error Level had a large effect on recovery (depending on Model Type). When Model Type = 2, recovery degraded very quickly as Error Level increased. For the other levels of Model Type, recovery did become degraded, but not to the same

Table 1  
 $R^2$  for Correct Model, Averaged  
 Over Model Type and Adjusted for  
 Number of Parameters, for  
 Uncategorized/Linear Cases Only

Error Level	Number of Stimuli				
	8	16	25	35	50
.00	1.00	1.00	1.00	1.00	1.00
.05-.10	.996	.992	.992	.989	.987
.20	.966	.956	.958	.945	.936
.40	.895	.878	.881	.882	.864
.80	.739	.708	.710	.681	.661

Table 2  
 Correlation  $r_3$  for Correct Model,  
 for Uncategorized/Linear Cases Only

Model Type	Error Level	Number of Stimuli	
		8, 16	25, 35, 50
2	.05	.599	.603
	.10	.362	.498
	.20	.223	.485
	.40	-.066	.405
	.80	.151	.197
3	.05, .10	.993	.994
	.20	.961	.968
	.40	.795	.892
	.80	.269	.604
4	.05, .10, .20	.993	.992
	.40	.874	.942
	.80	.806	.825

degree. Recovery improves as Number of Points increases; this result is intuitive and reasonable. The increase in recovery reaches its asymptote when Number of Points = 25; there seems little reason to believe that the addition of more objects will substantially change the results.

*Recovery measure: Selection of correct model.* In the external unfolding analysis, the parameter estimates obtained using different models for analysis of a given dataset may be compared. The differential fit of these estimates may be examined using the incremental  $R^2$  test of fit (Neter et al., 1985). This approach is especially appropriate for external unfolding, as the models are nested within one another. The proportions of cases in which the correct model was selected are given in Table 3.

When the data were generated using Error Level = .00, the correct model was found in all Uncategorized/Linear cases; thus they are not presented. When the data were generated using Model Type = 2, the selection of correct model was very poor. When the data were generated using Model Type = 3, selection was better than Model Type = 2, but degraded as large amounts of error are added. When the data were generated using Model Type = 4, the correct model was frequently selected, provided at least 25 points were included in the fixed set. Generally, the correct model

Table 3  
 Proportion of Correct Model  
 Selections for Uncategorized/Linear  
 Cases Only

Model Type	Error Level	Number of Stimuli			
		8	16,25	35	50
2	.05	.150	.931	.925	.912
	.1-.2	.044	.559	.612	.537
	.4-.8	.025	.117	.074	.070
3	.05	.987	.994	1.00	.962
	.1-.2	.794	.988	.969	.962
	.4-.8	.062	.581	.606	.781
4	.05	.987	.969	1.00	.975
	.1-.2	.975	.972	.975	.987
	.4-.8	.381	.792	.938	.956

was frequently selected only when number of points  $\geq 25$ .

Although Categorized/Monotonic cases are not presented, results for these analyses were generally similar to Uncategorized/Linear results presented in Table 3. Two important types of differences were found in comparing Uncategorized/Linear and Categorized/Monotonic cases. Performance is poorer in the Categorized/Monotonic cases (in comparison to the Uncategorized/Linear cases) for Error Level = 0 and Model Type = 2. In this case, Model III was often selected instead of Model II. Clearly, the categorization process degraded some of the information required to discriminate between these models. On the other hand, performance was better in the Categorized/Monotonic cases when Error Levels were high and Model Type = 2 or 3.

*Performance measure: Number of negative weights.* Several discussions of external unfolding have discussed negative weights in the Model II analysis and suggested interpretations (Carroll, 1980; DeSarbo & Carroll, 1985). Such invalid values could also be due to sampling fluctuation. Considering the collinearity inherent in models such as these, this is an important consideration. Table 4 presents the proportion of negative weights for Model II analysis (for Uncategorized/Linear cases only).

Analysis of data generated under Model II resulted in several negative weights when any error at all was present. Clearly, error affects the prob-

Table 4  
 Proportion of Negative Weights,  
 Model II, for Uncategorized/Linear  
 Cases Only

Model Type	Error Level					
	.00	.05	.10	.20	.40	.80
2	.000	.078	.103	.158	.233	.300
3	.000	.003	.027	.052	.148	.257
4	.540	.432	.471	.532	.504	.514

ability of obtaining negative weights. More interestingly, specifying the wrong model results in a very large proportion of negative weights. When Model Type = 4, a large proportion (about 50%) of negative weights is encountered. In this case, the weights  $w_{ia} = 0$ . Thus, by chance, the weights should be positive or negative about 50% of the time.

## Discussion

### The Monte Carlo Experiment

*Recovery and model type.* Recovery, as measured by  $R^2$ , was strongly affected by error. When Model Type = 2, recovery was generally poor, and was strongly affected by even small amounts of error. This is a matter of concern, as this is a popular model for preference data. Recovery under Model Type = 3 was better, and under Model Type = 4 it was very good indeed. Thus, in all cases except Model Type = 2, the model performed well.

*Model identification.* The correct model was not identified consistently when error was present. The proportion of correct identifications dropped quite low when error was high, but there was an interaction with the number of stimuli. Identification was best with Model Type = 4 and worst with Model Type = 2. As the number of stimuli increased with Model Type = 2 (holding error level constant), the proportion of correct model selection did not appreciably increase when large amounts of error were present. With Model Type = 3 or 4, only small differences in correct identifications were seen from 16 to 50 stimuli. Thus there is some

reason to be confident in the performance of these models, a posteriori to model identification. When faced with observations with an unknown measurement model, it is difficult to assess the likelihood of making the correct decision. It is clear, however, that there should be serious reservations about the identification process using incremental  $R^2$  statistics.

*Negative weights for Model II.* The results obtained here suggest that negative weights generally result from either sampling fluctuation (when Model Type = 2) or inappropriate model choice. The results suggest that negative weights should be interpreted very cautiously. Interpretations of negative weights as anti-ideal points (Carroll, 1972, 1980) are not parsimonious. It is simpler to regard them as a product of sampling error or incorrect model choice. If the number of negative weights is approximately 50%, this is diagnostic information suggesting that Model IV is really appropriate. For Model II analyses, negative weights also arise frequently. If this model is considered, constrained methods (Davison, 1976a, 1980; Srinivasan & Shocker, 1973) should be employed.

*Interpretation of the estimation process.* One problem is clear in the use of regression models for estimation. In the use of Model II, the primary parameters are ratios of estimates from a regression model. The concerns noted earlier about collinearity and ratios of statistics were well-taken. In general, using estimators without attention to the standard errors can result in uninterpretable conclusions.

## Conclusions

*External unfolding.* The use of regression approaches to external unfolding must be considered with some degree of caution and a little skepticism. The external unfolding approach is clearly a useful method, but not entirely free from problems. Generally, these problems arise from neglecting the statistical nature of the estimation process. Typical problems of statistical estimation, such as the sampling fluctuation of parameter estimates, require care in their interpretation. If a single general statement is to be extracted from the current results, it

is "Simpler external unfolding models perform better. Model II should almost never be postulated." Although Model I was not examined, these results imply that it would perform worse than Model II.

*Effects of categorization.* The effect of categorization may be generally summarized as being rather mild and reasonable. The monotonic analysis resulted in good recovery for many statistics. The exception is found in the model identification probability, which was generally poorer with the monotonic approaches. Because the categorization process used here is a rather simple one (in which classes were of equal width), the conclusions drawn here certainly are not entirely general. Yet the results are encouraging.

*Choice of model.* The results suggest that moderate amounts of error make it difficult to determine the correct model for estimation. Users of external unfolding should use theory rather than data to guide them in the model selection process. In this case, using an appropriate error model and somewhat artificially clean conditions, the correct model was selected less frequently than might be desired.

*Model selection.* Traditional methods for model identification in external unfolding rely on more qualitative aspects of recovery. The vector model may be interpreted as an ideal point model with infinitely remote ideal points (Carroll, 1980). Thus, when ideal points fall outside of the stimulus space, this is a necessary (but not sufficient) diagnostic for the vector model. The current research suggests that the incremental  $R^2$  test is not a necessary and sufficient diagnostic tool for making this important decision. Davison (1976b) suggested the use of a cross-validation approach, which has the advantage of incorporating external information into the decision process.

Theoretical considerations should also be addressed. For instance, can the preference function be considered single-peaked? Such situations would lead to the consideration of Model II or III. If it is conceivable that preference increases without limit through the domain of stimuli considered, this would suggest the use of Model IV. Although the vector model has some rather peculiar assumptions in the entire universe of stimuli, it may hold locally. Finally, future research should consider incorporat-

ing a priori information about the appropriate model into the situation. Perhaps this could be done in some Bayesian fashion.

### References

- Bennett, J. F., & Hays, W. L. (1960). Multidimensional unfolding: Determining the dimensionality of ranked preference data. *Psychometrika*, 25, 27–43.
- Carroll, J. D. (1972). Individual differences and multidimensional scaling. In R. N. Shepard, A. K. Romney, & S. Nerlove (Eds.), *Multidimensional scaling: Theory and applications for the behavioral sciences. Vol. I: Theory*. New York: Seminar Press.
- Carroll, J. D. (1980). Models and methods for multidimensional analysis of preferential choice (or other dominance) data. In E. D. Lantermann & H. Feger (Eds.), *Similarity and choice*. Bern: Huber.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
- Davison, M. L. (1976a). Fitting and testing Carroll's weighted unfolding model for preferences. *Psychometrika*, 41, 233–247.
- Davison, M. L. (1976b). External analysis of preference models. *Psychometrika*, 41, 557–558.
- Davison, M. L. (1980). DACAR: A FORTRAN program for fitting linear and distance models of preference. *Applied Psychological Measurement*, 4, 419.
- DeSarbo, W. S., & Carroll, J. D. (1985). Three-way metric unfolding via alternating least squares. *Psychometrika*, 50, 275–300.
- Gold, E. M. (1973). Metric unfolding: Data requirements for unique solution and clarification of Schönemann's algorithm. *Psychometrika*, 38, 555–569.
- Green, P. E., & Wind, Y. (1973). *Multiatribute decisions in marketing*. Hinsdale IL: Dryden.
- Hefner, R. A. (1958). *Extension of the law of comparative judgment to discriminable and multidimensional stimuli*. Doctoral dissertation, University of Michigan.
- Heiser, W. J. (1981). *Unfolding analysis of proximity data*. Doctoral dissertation, University of Leiden, The Netherlands.
- Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness-of-fit to a nonmetric hypothesis. *Psychometrika*, 29, 1–28.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading MA: Addison-Wesley.
- Neter, J., Wasserman, W., & Kutner, M. H. (1985). *Applied linear statistical models*. Homewood IL: Irwin.
- O'Brian, R. G., & Kaiser, M. K. (1985). MANOVA methods for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin*, 97, 316–333.
- Ramsay, J. O. (1977). Maximum likelihood estimation in multidimensional scaling. *Psychometrika*, 42, 241–266.
- Ramsay, J. O., ten Berge, J. M. F., & Styan, G. P. H. (1984). Matrix correlation. *Psychometrika*, 49, 403–423.
- Rodgers, J. L., & Young, F. W. (1981). Successive unfolding of family preferences. *Applied Psychological Measurement*, 5, 51–62.
- SAS Institute, Inc. (1986). *SAS user's guide: Statistics, version 5 edition*. Cary NC: Author.
- Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling*. New York: Academic Press.
- Schönemann, P. H. (1970). On metric multidimensional unfolding. *Psychometrika*, 35, 349–366.
- Srinivasan, R., & Shocker, A. (1973). Linear programming techniques for multidimensional analysis of preferences. *Psychometrika*, 38, 337–369.
- Takane, Y. (1981). Multidimensional successive-categories scaling: A maximum likelihood method. *Psychometrika*, 46, 9–28.
- Thompson, P. A. (1986). *PREF: A SAS PROC MATRIX/IML macro for external analysis of preferences*. Unpublished manuscript.
- Thompson, P. A. (1988, June). *Successive unfolding*. Paper presented at the annual meeting of the Psychometric Society, Los Angeles.
- Timm, N. H. (1975). *Multivariate analysis with applications in education and psychology*. Belmont CA: Brooks/Cole.
- Tucker, L. R. (1960). Intra-individual and interindividual multidimensionality. In H. Gulliksen & S. Messick (Eds.), *Psychological scaling: Theory and applications*. New York: Wiley.
- Young, F. W., de Leeuw, J., & Takane, Y. (1976). Regression with quantitative and qualitative variables: An alternating least squares method with optimal scaling features. *Psychometrika*, 41, 505–529.
- Zinnes, J. L., & MacKay, D. B. (1983). Probabilistic multidimensional scaling: Complete and incomplete data. *Psychometrika*, 48, 27–48.

### Acknowledgments

This paper was presented to the 4th European Meeting of the Psychometric Society and the Classification Society, July 2–5, 1985, Cambridge, England. The author thanks the Editor and two reviewers, whose careful review improved the paper in a number of ways.

### Author's Address

Send requests for reprints or further information to Paul Thompson, 301 Hagerty Hall, 1775 College Rd., Ohio State University, Columbus OH 43210-1399, U.S.A.