THE PERFORMANCE OF THREE FITTING CRITERIA FOR MULTIDIMENSIONAL SCALING

MARION McGLYNN

Department of Psychology

McGill University, Montreal

July, 1990

A Thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for
the degree of Master of Science

ABSTRACT

A Monte Carlo study was performed to investigate the ability of **MSCAL** to recover by Euclidean metric multidimensional scaling (MDS) the true structure for dissimilarity data with different underlying error distributions. Error models for three typical error distributions: normal, lognormal, and squared normal are implemented in **MSCAL** through data transformations incorporated into the criterion function. Recovery of the true configuration and true distances for i) single replication data with low error levels and ii) matrix conditional data with high error levels was studied as a function of the type of error distribution, fitting criterion, and dimensionality. Results indicated that if the data conform to the error distribution hypotheses, then the corresponding fitting criteria provide improved recovery, but only for data with low error levels when the true dimensionality is known.

RESUME

La méthode de Monte Carlo a été utilisée pour examiner la capacité du programme **MSCAL** à recouvrer par le MDS (*multi-dimensional scaling*), la structure spatiale réelle des données de dissimilarités, lorsque différentes distributions d'erreur sont utilisées. On peut utiliser trois distributions d'erreur courantes (normale, log-normale, et normale-carrée) avec **MSCAL**, par le biais d'un processus de transformation des données inclut dans la fonction critère. L'obtention de la configuration et des distances réelles pour i) des données de matrices simples avec un taux d'erreur faible, et ii) des données de matrices conditionnelles avec un taux d'erreur élevé, a été examinée selon le type de distribution de l'erreur, la fonction critère, et la dimension de l'espace de la solution. Les résultats indiquent, que lorsque les données se conforment aux hypothèses de distribution de l'erreur, les fonctions critères correspondantes améliorent le recouvrement, mais seulement dans le cas de données à faible taux d'erreur, et lorsque la dimension de l'espace est connue.

ACKNOWLEDGEMENTS

Special thanks must be given to my advisor, Professor Jim Ramsay, whose help and encouragement have been invaluable in the preparation of this thesis, and who first suggested the topic.

The author would also like to thank D. Clarkson (IMSL, Inc.) for granting permission to use the **MSCAL** program for this study.

The author recognizes the financial assistance received from the Natural Sciences and Engineering Research Council of Canada.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The program **MSCAL** (Clarkson, 1988a & b) from IMSL, Inc. embodies a number of different multidimensional scaling (MDS) models, using both metric and non-metric procedures. However, this investigation focuses on the error models provided by the program which are associated with Euclidean metric MDS. Three typical error distributions are available to the user: normal, lognormal, and squared normal, corresponding to the data transformations used in the criterion function for least-squares estimation. Data of known configuration and error will be used to evaluate the performance of **MSCAL** in satisfactorily recovering configurations and distances for dissimilarity data, using the different error models.

1.1 Description of the MSCAL Program

MSCAL provides a number of distance models, as well as various forms of the stress function to be optimized for the different types of dissimilarity data. There are a large number of possible models due to the many options available in the program. Careful consideration is required to choose those which give appropriate measures and models for the data being analyzed.

Input parameters allow one to specify the level of measurement of the data as nominal, ordinal or interval. As well, the initial configuration may be either input or computed. The distance transformations used to compute the criterion function permit squared distances, distances, or the log of the distances. It is these transformations which are under scrutiny in this study because of their relationship to MDS error models. The various models and parameters of **MSCAL** are discussed in the next chapter, in the context of the design of this study.

1.2 Theoretical Aspects of Scaling Models

According to Takane (1981) a scaling procedure should do more than just scale the data. It should also represent the data by an appropriate model. Takane distinguishes between the representation model, error model and response model.

The representation model indicates the kind of perceptual relationship ascribed to the data. Thus, the scaling or transformation of the data should reflect some reasonable assumptions about one's concept of dissimilarity. Data representation models typically use Euclidean distance in MDS to represent the dissimilarity of two stimuli in space. Although distances, other than straight-line, are also used.

Various types of data may require different error models. Not only is the magnitude of the measurement error important for modelling subjects' judgments of dissimilarity, but also its distribution and associated characteristics. A single error distribution is generally assumed.

Different types of judgments require different mental operations by the subject. Reasonable assumptions are needed about the "psychological processes involved in a specific task situation which generates a specific type of data" (Takane, 1981, p.11). This is the response model which specifies the main characteristics of the data generating processes and gives an explicit account, in mathematical terms, of the subjects' transformations in the psychological space that produce the dissimilarity judgments. Hence, the response model includes the representation and error models, as well as their parameters, as a complete description of the observed data.

1.2.1 The Representation Model

Most MDS models for proximity (i.e. dissimilarity) data use some form of the Minkowski distance function given as:

$$d_{ij} = \left(\sum_{a=1}^{n} |x_{ia}-x_{ja}|^{r}\right)^{\frac{1}{r}}$$

to refer to the theoretical distance between coordinates for a set of stimuli i and j in a space of given dimensionality, n. Euclidean distance (r=2) is commonly used for perceptually "unitary" stimuli such as homogenous colours. The city-block metric (r=1) is preferred for "analyzable" stimuli such as geometrical shapes which differ in a number of dimensions (i.e. size, orientation, etc...) or for stimuli judged with more than one sense (Shepard, 1980; Schiffman, Reynolds, & Young, 1981). **MSCAL** provides only weighted or unweighted

Euclidean distance models, thus favouring perceptually unitary stimuli for dissimilarity data models.

1.2.2 Error Models

Data typically contain appreciable amounts of measurement error. Estimation methods are affected by various sources of error in the data, such that large enough error will obscure the underlying structure of the data and lead to worthless solutions. Yet, as Ramsay (1978) points out, the distribution of residuals or errors is usually only implicitly defined in most MDS models. An explicit form provides an opportunity to tune the analysis by subsequent revision of the choice of error model.

The additive and multiplicative error models are two typical examples given by Takane (1981):

$$\delta^*_{ijm} = d_{ij} + e_{ijm},$$

where $e_{ijm} \sim N(0, \sigma^2_m)$ and

$$\delta^*_{ijm} = d_{ij}\, e_{ijm},$$

where $\ln e_{ijm} \sim N(0, \sigma^2_m)$. The d_{ij} is the true distance between stimuli i and j, δ^*_{ijm} denotes the corresponding perturbed distance for the m-th subject, and e_{ijm} is the random normal deviate representing the error. These models allow individual differences in the variance. If subjects are taken as replications then there is a single common variance.

The additive error model assumes that the error is normally distributed over replications of judgments with constant variance for each subject. The multiplicative model given above is equivalent to the lognormal distribution of the error explicit in **MULTISCALE** (Ramsay, 1977). The lognormal model reflects two features associated with dissimilarity data: values are naturally positive and variability increases with magnitude (Ramsay, 1982a).

There is empirical evidence to support a non-linear monotonic relation between dissimilarity and distance. In psychophysics, subjective assessments of physically measurable properties with a rational origin often display a power relation between judgments of dissimilarity and distance (Ramsay, 1982a).

MSCAL provides three separate error models through its data transformation option, when a least-squares fit is used. The first two models has normally distributed error with constant within-subject variance about the distances and squared distances, respectively. Similarly, the third model has constant within-subject variance for the logarithm of the distances as in multiplicative model mentioned previously.

An important benefit derived from making the error an explicit part of the model is the opportunity it gives to identify the sources of variation in the data. The major dividing line between two-way and three-way MDS models is how they view subject differences: two-way models treat individual variation as incorrect judgments, so that individual differences are part of the error model; three-way models, on the other hand, address these differences as reflections of different personal judgments or psychological measures in the representation model by using a weighted distance model. Meulman (1986) suggested using the subject weight matrices "as a means of filtering out undesired variability" (p.176), presumably when the individual differences are not important or interesting.

Building good models requires identifying the sources of variation in the data which reflect the nature of the judgments subjects make and their ability to do so, from a psychological perspective. Subjects will be more or less adept at making judgments and some measure of individual idiosyncrasies should be reflected in the error model. Also, stimuli may have greater or lesser distinctiveness or salience for a given dimension or attribute. Thus, by varying the types of weights and distance models used, one can produce response models with different variance components so as to recognize the many sources of variation in subjects' judgments apart from the "true" dissimilarities between stimuli.

Three different types of variance components comprise **MULTISCALE** (Ramsay, 1982b): pair-wise, stimulus-wise and subject-wise. Pair-wise variance allows for separate estimates for each pair of stimuli, assuming they will be different. Stimulus-wise variance takes into account the varying familiarity of stimuli to the subject. **MSCAL** permits two components of variation: subject weights and stimulus weights for each dimension, so that the metric can vary from subject to subject and/or stimulus to stimulus. Schiffman, Reynolds and Young (1981) refer to the stimulus standard error weights, used in **MULTISCALE**, as measures of "cognitive uncertainty" of subjects' perceptions of each stimulus. **MSCAL**'s stimulus weights might be viewed as measures of the perceptual differences in the stimuli themselves.

1.2.3 Individual Differences

As pointed out by Takane, Young, and DeLeeuw (1977), individual differences models are of three psychologically distinct types: those that arise from response bias, those that result from a judgmental process (either perceptual or cognitive), and some combination of the two. One can allow for each type through model weights and/or data conditionality, regardless of the measurement level of the data. Individual differences due to response bias are effected by assuming replicated data are conditional; individual differences in the judgmental processes are reflected by the weights in the weighted distance model. Thus, even if an unweighted distance model is used, individual differences can still be represented by allowing different response transformations for each subject. The **MSCAL** program has an option for the level of stratification in the data permitting unconditional, matrix conditional, and column conditional (i.e. appropriate only for the asymmetrical case) data.

On the other hand, the non-individual difference model "assumes that replications arise from subjects with identical judgmental and response processes" (Takane, Young, & DeLeeuw (1977), p.52). This applies to single matrix data as well, as replicated data, by treating the data as unconditional.

1.3 Estimation Methods

Generally, MDS procedures use one of two methods for measuring the fit of a model: least squares and maximum likelihood. The former relies on minimizing the discrepancies for the chosen criterion function; hence, some kind of badness-of-fit index is used. Maximum likelihood methods require statistical assumptions about the random variation in the data and estimates of the model parameters are those which give the largest probability or likelihood of occurring with the observed data for a given set of the model values.

MSCAL minimizes a loss function which allows the user to specify the power of the estimates. The only restriction is that the power be at least one. Users would normally select one of the following types of estimates: least-squares, minimum absolute deviation, or the sum of the 1.5 power of the errors, a typical exponent for dissimilarity data (Ramsay, 1982a, p.288). The maximum likelihood method is not used in **MSCAL**. However, metric scaling of dissimilarities assumed to be independently normally distributed with constant residual variance produces normal distribution theory maximum likelihood estimates.

Different MDS programs fit different transforms of the data. **MULTISCALE** fits distances directly. Because **ALSCAL** (Takane, Young, & DeLeeuw, 1977) fits squared distances instead of distances, the program tends to fit mostly large discrepancies, sacrificing fit in the smaller and moderate discrepancies. **ALSCAL** performs best when error is least and distributed mostly in the smaller proximities; only when error in the larger squared distances approximates that in the smaller squared distances is **ALSCAL**'s performance unaffected by error (Ball, 1982). While **INDSCAL** (Carroll & Chang, 1972) uses the same weighted Euclidean model as **ALSCAL**, it fits scalar products that have been converted from squared proximities. Ball's (1982) study notes that **INDSCAL**'s performance is much less affected by the level of error, its distribution and the underlying error model, than when **ALSCAL** is employed.

Using least squares estimation, programs should perform best when error is normally distributed about the measures being fit. Thus, when squared distances are fit, the best results should occur when error is normally distributed about the squared distances. Similarly, the same reasoning should apply when fitting distances, scalar products and the logarithm of the distances. Ball's (1982) evaluation of **MULTISCALE** (Ramsay, 1977) attributes the lack of such findings for lognormal data to the program's frequent failure to converge rather than to how the response model interacts with the data.

The alternating least-squares method approach used in **MSCAL** sequentially fits the distance component and the spatial component. The expected effects of error and its distribution on recovery of the distance component has been discussed above; however, the effects of error on the spatial component need to be addressed.

1.4 Initial Configuration Estimation

Converting distances to scalar products and then using a decomposition of the scalar products matrix to estimate the coordinates of the configuration leads to a number of problems, as Ramsay (1982a) points out: dissimilarity must be measured to within a scale factor and random disturbances, as found in real data, are exaggerated by squaring in the double-centring transformation used to convert distances to scalar products. **MSCAL** uses this classical MDS approach, which double-centres the squared distances, and applies matrix factoring through eigen-analysis of the product moment matrices to arrive at the initial configuration. As Spence and Lewandowsky (1989) show, the Young-Householder-Torgerson (Young & Householder, 1938; Torgerson, 1958) procedure is sensitive to the effects of outliers and provides a "poorer starting position than a random configuration" (p.505). Thus, local minima are a possible problem to be considered when evaluating **MSCAL**.

1.5 Use of Monte Carlo Methods

Monte Carlo methods have come to be used "in any situation where a complete mathematical analysis of a problem is difficult or intractable." (Spence, 1983, p.406) Computer simulations using the Monte Carlo method are designed experiments, requiring thoughtful planning and precise execution. A factorial design becomes increasingly complex as the number of factors and/or levels of factors are added. For this reason, most Monte Carlo investigations have been limited in their scope and their results are only suggestive of the underlying relationships between the independent and dependent variables. Nevertheless, the practical demands of MDS techniques, require some empirical guidelines for the use of the various programs. Studying the effects of various controlled factors using Monte Carlo simulations helps to assess the solutions and measures of fit which these scaling procedures provide.

1.6 MDS and Monte Carlo Studies

Inevitably, real psychological data contain error or "noise". The capability of various programs to recover the true structure of empirical data will depend upon how successfully they deal with noise. Monte Carlo studies using known stimulus configuration with data perturbed by random error, have been important in determining the behaviour of various programs and in evaluating the goodness of fit of metric, as well as non-metric, methods. The Monte Carlo approach is ideally suited to this study because of the ease with which error models can be imposed upon the data.

1.7 Review of Monte Carlo Studies for Two-way Scaling

Wagenaar and Padmos (1971) studied the behaviour of Kruskal's non-metric **M-D-SCAL** (1964a, b) program, by measuring the stress values for random configurations of a small number of points with various levels of error and true dimensionality. Their results were intended as bases of comparison to estimate true dimensionality and measurement error of empirical data. Unlike previous studies, Wagenaar and Padmos used a different error model. Random normal deviates were multiplicatively applied to the error-free Euclidean distances instead of the coordinate points. However, negative distances were possible with such a method, and negative random elements had to be rejected.

Isaac and Poor (1974) also investigated the problem of determining the true underlying dimensionality of error-perturbed data using **M-D-SCAL** (Kruskal, 1964a, b). Three separate measures of recovery were used as the dependent variables: Kruskal's Stress Formula 2, the index of metric determinacy, and a new measure introduced as *Constraint*. This measure was the difference between the expected stress value for random data and the stress value of a typical configuration in the same dimensionality with the same number of points. The manipulated factors were typically: the number of points, the true dimensionality, the amount of error, and the dimensionality of the solution. The error generation was similar to Young's (1970) study, where random error was added to the coordinates. However, Isaac and Poor used higher levels of error. The results were consistent with those of Young and confirmed the inadequacy of stress in identifying the true dimensionality of data with error.

Sherman (1972) produced an extensive study of metric determinacy of the non-metric scaling program, **TORSCA** (Young, 1968), employing a large number of factors. In addition to studying the effects of the number of points in the configuration, true dimensionality, and error level as had Young (1970), Sherman included the Minkowski constant, r, in the distance function and the number of dimensions used in scaling. Young had considered only Euclidean distances. Sherman varied the Minkowski r-metric (r=1,2,&3) but scaled the data assuming it had been measured with a Euclidean distance function (r=2) to determine the effects of mis-estimating the distance function.

As Sherman noted, the overall trends found in the two studies were the same. Differences notwithstanding, the results related to the additional factors in Sherman's study indicated that there was good metric determinacy: 1) if the dimensionality of the recovered configuration was not less than the dimensionality of the true structure and 2) the accurate estimation of the Minkowski constant led to a better model only if the dimensionality has been properly estimated. Sherman further qualified the choice of the Minkowski constant, *r*. If the stimuli under study had perceptually distinct attributes (for example, height) then using city-block distance provided a better model. Otherwise, for stimuli with interacting attributes (i.e. hue and brightness) *r* values greater than one lead to better models.

One major objection raised by Cohen and Jones (1974) was that previous studies, such as Sherman (1972), had based their findings of the effects of dimensionality on the recovery of distances and not of the true configuration itself. Their results qualified the implications of dimensionality such that "partial information on any dimension will be recovered to the extent that it is available in the data when the solution is in a sufficiently high dimensionality." (p.88) Underestimation of dimensionality is associated with loss of information and possible distortion of the solution dimensions.

Spence (1972) employed a comparative approach among three popular non-metric algorithms, **M-D-SCAL** (Kruskal, 1964a, b), **SSA-I** (Lingoes, 1965), and **TORSCA** (Young, 1968) to assess their performance. Using Ramsay's model (1969), Spence added random error independently to each of the randomly generated coordinates. This method of generating dissimilarities was the same as that used by other investigators (Sherman, 1972; Young, 1970).

The performance of the three procedures in obtaining initial and final configurations was assessed by Kruskal's stress measure and the index of metric determinacy, *M*, the correlation between true and recovered distances. Thus, recovery of configurations were not measured directly. A complicated design was used involving 18 distinct configurations, with varying number of dimensions and datapoints, four separate error levels and five levels of recovered dimensionality.

The results showed relatively small differences between the solutions of the different algorithms. However, each of the algorithms were susceptible to sub-optimal solution problems. **M-D-SCAL** produced the largest number of deviant solutions, while very few of the **SSA-I** and **TORSCA** solutions were unsatisfactory. For all algorithms problems were severe in one dimension.

Spence identified the quality of the initial configuration as the major factor in avoiding local minima in the configuration. **TORSCA**'s success was attributed to its ability to generate a good initial configuration, one close to the global optimum. Also, not all algorithms improved the fit; sometimes, it was worsened. It is believed that such non-convergent programs display local optimum problems more frequently than convergent programs (Schiffman, Reynolds, & Young, 1981).

Graef and Spence (1979) investigated whether small, medium, or large distances were most important in determining the recovery performance of **TORSCA** (Young, 1968). An important feature of this Monte Carlo study was the inclusion of two separate error models: the Ramsay (1969) model and the Wagenaar-Padmos (1971) model, each with two levels of error. The two error models were used principally because they were thought to represent the extremes encountered with real data. Results indicated that large distances were critical to good performance in recovery of the true distances; whereas small and medium distances had a less crucial role, independent of the error models used.

Spence and Lewandowsky (1989) studied the effects of outliers on various MDS procedures, including their procedure **TUFSCAL**, using Monte Carlo simulations with a lognormal error model. Their results indicated that metric procedures such as Young-Householder-Torgerson metric scaling (Young & Householder, 1938; Torgerson,1958), **KYST-2** (Kruskal, Young, & Seery, 1978) in metric mode, and **MULTISCALE** (Ramsay, 1977; 1982b) generally perform worse than the non-metric procedures, **ALSCAL** (Takane, Young, & DeLeeuw, 1977; Young & Lewyckyj, 1981) in the non-metric model and **KYST-2** in non-metric mode, when outliers were present and, with the exception of **MULTISCALE** when the data also had moderate background error.

Despite general resistance to the effects of outliers, **TUFSCAL**'s performance was less good in some cases than others, for which higher percentages of outliers gave perfect recovery. Such instances were attributed to "particular patterns of outliers" for which perfect recovery was not possible, rather than to occurrences of local minima. From this viewpoint, it is sometimes the particular nature of the error in data that may be limiting recovery, rather than the algorithm.

By using a "cross"-shaped true configuration, Spence and Lewandowsky (1989) provided yardstick measures of what constitutes good recovery correlations: configurations whose recovery correlations were less than 0.7 had little resemblance to the true configuration and generally those configurations associated with correlations less than or equal to 0.9 were judged to be unsatisfactory.

In general, Monte Carlo studies have shown that two-way non-metric procedures perform best when the true dimensionality is low, the number of points is large and the error level is moderate. The minimum number of points suggested per dimension is six to avoid the problem of degenerate solutions (Spence, 1983). These results are assumed to apply equally to metric and non-metric models. Various error models have been considered, including those which add error to the distances. Also, it is important to measure both recovery of distances and configuration to get a true picture of performance.

Depending on the nature of the error in the data, various interactions are expected for the different error models.

1.8 Interaction of True and Predicted Error Models

It is natural to expect that **MSCAL** will perform best when the predicted error model matches the error distribution of the data. For a least-squares fit it is assumed that the absolute errors are distributed uniformly across the data. Specifically, random normal error is distributed uniformly across distances for a normal error model and across squared distances, if a squared normal error distribution is present in the data. However, for lognormal error, it is the relative error (i.e. relative to the magnitude of the distances) that is scattered uniformly across the range of distances.

When the data has a normal error distribution and a log transformation is used, as in the case where the predicted error model is lognormal, a least squares fit will tend to minimize the error in the smaller distances over that in the larger distances. On the other hand, if a squared transformation is used, estimation will concentrate on the larger distances, since errors in the larger proximities are expanded by squaring. However, if the actual values of the distances are more or less equal, then all transformations should give the similar results.

When the data has error with a squared normal distribution, a least squares fit assuming a normal error model will tend to overfit the larger and smaller distances. If a lognormal transformation is used, it will tend to overfit the smaller distances. However, as in the case of normal error, if the actual distances are about equal and close to one then all transformations should give approximately equal recovery.

In the case of data with lognormal error, larger absolute errors are concentrated in the larger distances. If the predicted error model assumes a normal distribution, least-squares estimation will concentrate on fitting the larger distances and the smaller distances will have less of an influence. If a squared normal error model is assumed, squared distances are fit to the squared disparities. Again, there is a tendency to fit error in the larger error-perturbed values, since errors in the larger distances are expanded by squaring.

1.9 Simulating Typical Users' Needs

One last point needs to be made with respect to the design of Monte Carlo studies. The needs of a typical user should be considered. Spence (1972) in evaluating non-metric MDS programs suggested the following guidelines to simulate circumstances encountered in an experimental situation: 1) The range of stimuli should be realistic. 2) Standard default options should be used to simulate what the average user typically selects. 3) The maximum number of iterations should be 15-30 to allow an algorithm to reach a minimum according to experience of some investigators. 4) Spaces of differing true dimensionality should be explored, as well as the effect of error in the data. 5) One should look for interactions between the programs and the error level, the number of stimuli, and the true and recovered dimensionalities. (Spence, 1972, p.468)

Following this advice, the next chapter provides the detailed methodology for the present Monte Carlo study.

Chapter 2

DESIGN OF THE STUDY

The performance of the **MSCAL** program will evaluated in terms of two separate measures: 1) recovery of the true configuration and 2) recovery of the true distances. Starting from a true configuration, the corresponding true proximities are calculated. Using the Monte Carlo method, random error is added to the true distances according to the error models established by the three fitting criteria in the **MSCAL** program. The study will include two parts: the first analyzes data with a single matrix for which no individual differences are possible, and the second allows for individual differences in response style using a replicated analysis.

Since **MSCAL** is a new multidimensional scaling program, this study will concentrate on its basic capabilities, and leave evaluation of other aspects for future consideration. Metric multidimensional scaling will be used for square symmetric data with ratio measurement level, the strictest condition on optimal scaling, using an unweighted model and least-squares estimation. Also, a simple linear model relating the true distances and the dissimilarity data of the form $f(\delta^*_{ijm}) = f(d_{ijm}) + e_{ijm}$ is assumed to make the simulations as similar as possible across the different error distributions imposed on the data.

A weighted distance model, with either subject or stimulus weights, would add more complexity to the problem of assessing performance of the fitting criteria. However, **MSCAL** does permit individual differences in response style for an unweighted Euclidean distance model, depending on the conditionality of the data, through the stratification level option. A stratification level indicating matrix conditional data (i.e. each stratum will correspond to a single subject, reflecting individual differences in variance) will be used to improve the fit of replicated data. Accordingly, single matrix dissimilarities data will be treated as unconditional by **MSCAL**.

The general stress function optimized by metric **MSCAL** is given as:

$$\phi = \sum_{h} v_h \sum_{i,j} |f(\delta^*_{ijm}) - \alpha_h - \beta_h f(\delta_{ijm})|^p$$

where δ denotes the predicted distances, δ^* denotes the dissimilarities, v_h is the stratum weight, f is one of the data transformations ($f(x)=x, f(x)=\ln(x)$, or $f(x)=x^2$), α_h is the stratum intercept, β_h is the scaling factor and h indexes the strata (for matrix conditional data, $h=m$, the index for subjects) and $p=2$ for least squares estimation.

2.1 Choice of MSCAL Parameters

The type of the data transformation specified for the stress function is the main variable of interest in this study. The number of options available for the program parameters are many and their selection must be compatible within the framework of these fitting criteria, as well as the data.

Given the choice of representation model and the measurement level of the data, the stratum intercepts α_h were not necessary. To eliminate overparameterization, **MSCAL** places restrictions related to the fitting criteria on the scaling factors, β_h. A normal or squared normal fitting criterion requires: 1) $\beta_h = 1$ for unconditional data and 2) $\sum_h \beta_h^2 = \eta$ for matrix conditional data, where η is the number of subjects. To avoid differences in the criterion function optimized for the different error models, only the configuration coordinates will be estimated.

MSCAL permits three different stratum weights in the stress function: stratum variance, the sums of the squared disparities, or unconditional variance. The normalizing factor which corresponds to the sum of the squared disparities corresponds to the stress function, Φ_1 given by:

$$\Phi_1 = \sum_h \nu_h \sum_{i,j} (f(\delta^*_{ijm}) - f(\delta_{ijm}))^2,$$

where $\nu_h = \sum_{i,j} (f(\delta^*_{ijm}))^2, h=m$ for matrix conditional data. For a metric model, Φ_2 the stress function weighted by unconditional stratum variance, is related to Φ_1 as follows:

$$\Phi_1 = \Phi_2 \frac{\sum_h \sum_{i,j} (f(\delta^*_{ijm}) - \bar{f}(\delta^*_{...}))^2}{\sum_h \sum_{i,j} (f(\delta^*_{ijm}))^2};$$

where $\bar{f}(\delta^*_{\ldots})$ is the mean transformed dissimilarities in the stratum. For simplicity, the stress function ϕ_1 was chosen. Moreover, when a squared transformation is used for the fitting criterion, ϕ_1 is proportional to SSTRESS in **ALSCAL** (Takane, Young, & DeLeeuw, 1977). The third stratum weighting option, conditional variance, was not chosen because it resulted in the following criterion function:

$$\phi_0 = \sum_h n_h \ln [\sum_{i,j} (f(\delta^*_{ijm}) - f(\delta_{ijm}))^2] ,$$

where n_h is the number of disparities in stratum h. It was unclear how the optimization would be affected by the logarithm of the stress function when the underlying error distribution and fitting criterion do not match.

2.2 Generation of the True Data

The true stimulus configurations and the corresponding true proximities were selected from configurations output by **MULTISCALE II** for two sets of data: 1) data for 32 stimuli of occupations (Coxon, 1982), and 2) dissimilarity ratings on a 9-category rating scale of 14 emotions from 15 members of an MDS Workshop (J.O.Ramsay, personal communication, 1990). The original datasets represent two basic types of applications: the Coxon data have only a single matrix of dissimilarities, and the emotions data are replicated.

The chosen configuration populations are not artificial, unlike other Monte Carlo studies whose results have uncertain applicability to empirical data. The number of datapoints in each dataset should be sufficient for analysis and are typical in size. Inasmuch as these datasets have been previously analyzed, they are reasonable choices for this study. There is no claim that the conclusions of this study will apply to all datasets, but rather only to datasets of comparable size.

2.2.1 True Stimulus Configurations

The output configuration from a **MULTISCALE II** analysis using an underlying normal error distribution and scale transformation of the Coxon data was used as the common stimulus space for occupations. The configuration derived from an analysis using a lognormal error distribution and scale transformation was used for the emotions common stimulus space for the 15 subjects. For both sets of data, three dimensional solutions were chosen, since previous **MULTISCALE II** analyses indicated the data sets were both three-dimensional. Figures 1 and 2 give the corresponding true configurations. The corresponding population configuration matrices are given in Appendix A Tables 13 and 14.

2.2.2 True Proximities

The true proximities were the associated $n(n-1)/2$ inter-stimulus distances for the two true configurations derived using the unweighted Euclidean distance model in the original analyses. The corresponding population proximities matrices are given in Appendix A Tables 15 and 16.

The True Configuration for the Coxon Data

Figure 1a: Dim. 2 vs Dim. 1

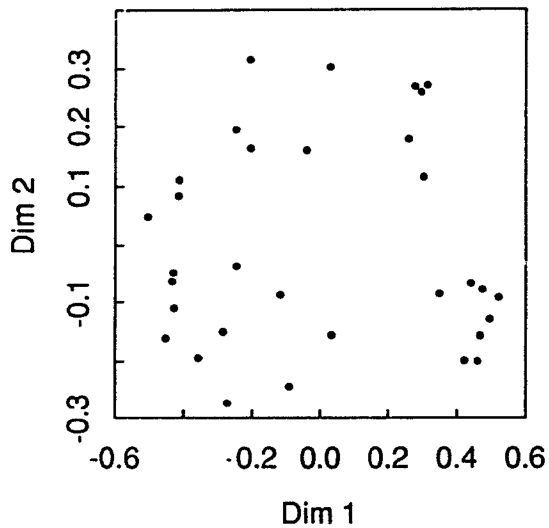


Figure 1b: Dim. 3 vs Dim. 1

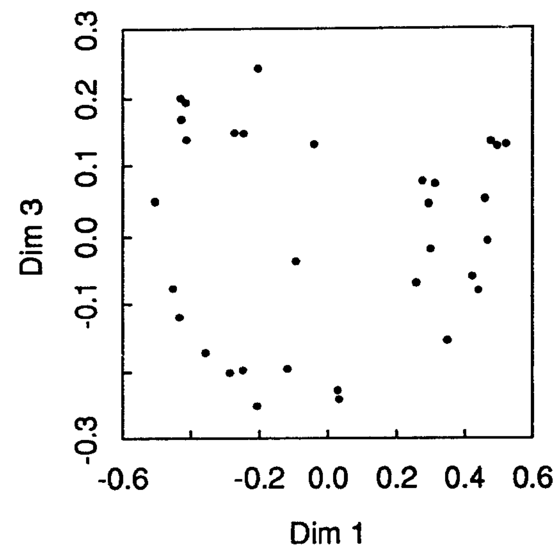
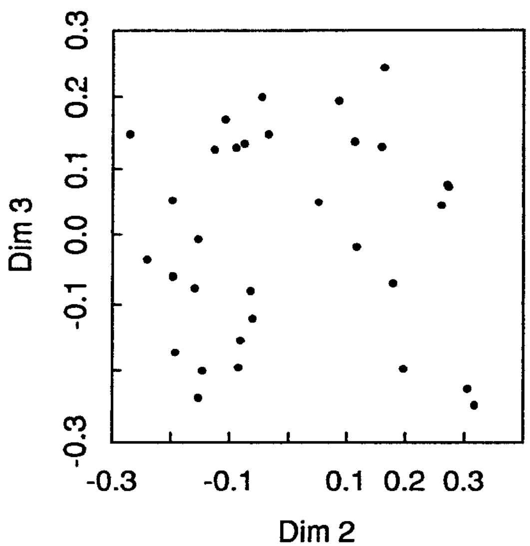


Figure 1c: Dim. 3 vs Dim. 2



The True Configuration for the Emotions Data

Figure 2a: Dim. 2 vs Dim. 1

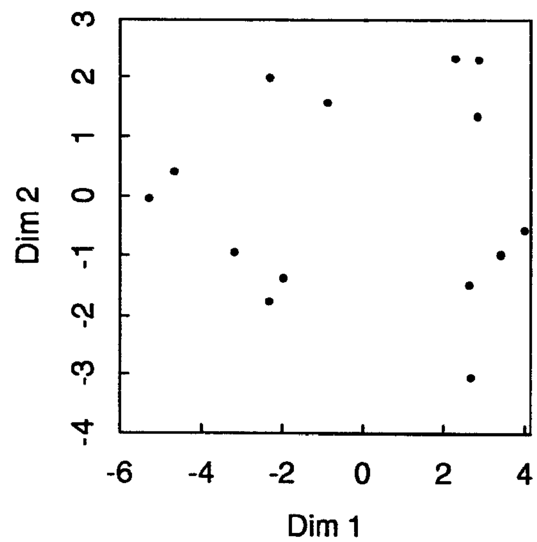


Figure 2b: Dim. 3 vs Dim. 1

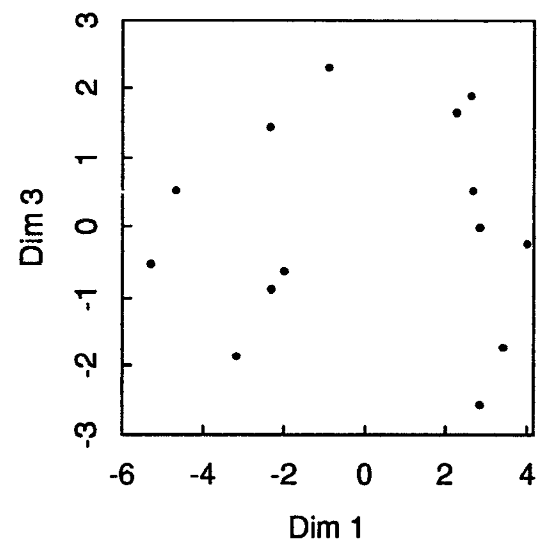
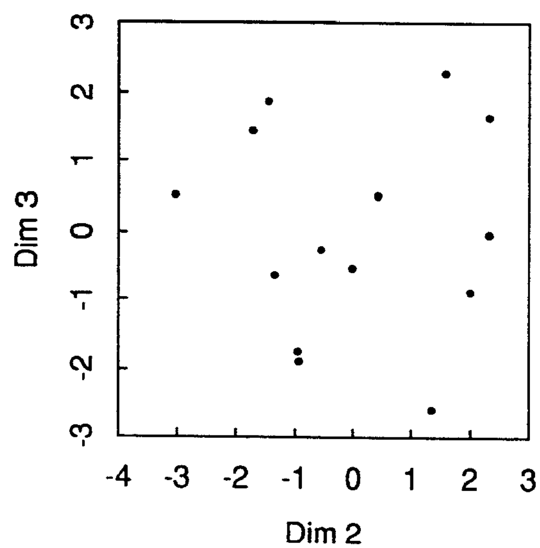


Figure 2c: Dim. 3 vs Dim. 2



2.3 Error Models

Error was generated according to Monte Carlo techniques. The random normal number generator, **RNNOR**, from **IMSL** STAT/LIBRARY (1987) was used to produce the error.

Random normal deviates, with a mean of zero and constant variance for each subject, were added to each of the true proximities in accordance with the three error models. A replication factor of 25 was chosen so that the study would be of reasonable size. This number of replications was expected to be sufficient to produce the kinds of effects that are important to sample variability.

2.3.1 Normal Error Model

Random normal deviates e_{ijm} were added to each of the $n(n-1)/2$ true distances d_{ijm} for m subjects and n stimuli to give the perturbed proximities δ_{ijm} reflecting an underlying normal error distribution as follows:

$$\delta_{ijm} = d_{ijm} + e_{ijm}.$$

In this case, there is no relationship between the magnitude of the distance and error, with the exception that for error-perturbed values less than or equal to zero are set to the value of 0.01 for the Coxon data and 0.1 for the emotions data. These values were chosen after inspection of the data, to be small relative to the distances generated. This will have some effect on the distribution of error for small distances and on the expected values of the estimated configuration matrix.

2.3.2 Lognormal Error Model

Random lognormal error was added to the true inter-stimulus distances as follows:

$$\delta_{ijm} = \exp[\ln(d_{ijm}) + e_{ijm}].$$

The typical error is proportional to the true distances. No zero perturbed distances were generated.

2.3.3 Squared Normal Error Model

Random normal deviates were added to the squared distances and the square root of each sum gave the following perturbed proximities:

$$\delta_{ijm} = \sqrt{d_{ijm}^2 + e_{ijm}}.$$

It should be noted that when the error added to the squared true distance gave a negative or zero value the perturbed distance was set as described for the normal error model.

In total, 3 (error model types) X 3 (data transformation types) X 25 (replications) proximities matrices for m subjects were generated for the Coxon (m=1) and emotions (m=15) data.

2.4 Subject Standard Error

Within-subject variances were assumed to be constant; however for the emotions data, error variance varied across subjects, as is typical for observed data. Subject standard errors for the Coxon and the emotions data were taken from the **MULTISCALE II** estimates that provided the true configurations and proximities described above. The within-subject standard errors estimated by **MULTISCALE II** for these configurations are presumed to reflect realistic standard error weights for each subject for these datasets. The standard error for the Coxon data, 0.091, with a normal error distribution represents a low error condition, while the average standard error for the emotions data, 0.397, with a lognormal error distribution indicates a high level of error.

Population standard error values for the two remaining error models were determined as follows: Regression with no intercept of the squared dissimilarities on the squared true distances for each subject gave residual standard errors which were used to generate the appropriate perturbed proximities with squared normal error for both sets of data. Similarly, for the emotions data with normal error, the original dissimilarities data for each subject were regressed on the true distances. The unbiased standard error estimated by **MULTISCALE II**, for a lognormal error distribution and scale transformation analysis of the original Coxon data, was used to generate Coxon data with lognormal error. Thus, for both the Coxon and emotions data, three sets of subject standard error weights were provided for the corresponding error models.

2.5 Standardization of Error Levels

Comparison of the results from the various **MSCAL** analyses required standardization of the error levels in the generated data. It was otherwise unclear how to take into account the standard error levels. The measure chosen to equate error levels in the generated data was the proportion-of-variance of error (PVE), one minus the squared correlation between the $n(n-1)/2$ true proximities and the corresponding perturbed proximities. This procedure was used by Ball (1982).

After adjustment the mean PVE for the Coxon data was 0.17, using standard deviations of: 0.111, 0.186 and 0.096 for normal, lognormal, and squared normal error, respectively. However, the emotions data required 15 subject standard error weights for each type of error. The sets of initial subject standard error weights, given above, were scaled to produce roughly equivalent global levels of error as measured by the proportion-of-variance statistic. Multiple runs for the emotions data were necessary to produce mean PVE values of 0.50 for squared normal error, 0.53 for lognormal error and 0.54 for normal error. It was difficult to obtain exact equivalence.

Also, it is important to note where the error tends to be distributed for the various distributions. The mean correlat-ion between the absolute magnitude of error and the true distances (FCE) is reported for the three types of error distributions: 0.03, 0.46, and -0.43 for the Coxon data, and 0.05, 0.37, and -0.31 for the emotions data, for normal, lognormal, and squared normal, respectively. A positive FCE value indicates the extent to which larger error is present in larger distances whereas a negative value indicates more error in the smaller distances. Thus, error is distributed quite differently for lognormal and squared normal data.

2.6 MSCAL Program Parameter Values

The version provided by D. Clarkson is a test copy for batch processing of the **MSCAL** program from **IMSL**, Inc. **MSCAL** (Clarkson, 1988a & b) has a number of options available which determine the type of distance, error and overall response models in use for any given multidimensional scaling analysis, as previously noted.

For the two sets of data in the study, the *NSUB*, *NROW* and *NCOL* parameters were set to the appropriate values for the number of subjects and stimuli for these datasets. The following parameters and chosen values were used in the analyses:

IFORM=0 - square symmetric matrices;
ICNVRT=0 - dissimilarity matrices are input and no conversion is necessary;
ISTRAT=1 - data are matrix conditional (a single matrix is treated as matrix unconditional);
IDISP=0 - ratio or interval level data;
IMOD1=3 - requests initial estimates of the configuration;
IEST=0 - indicates ratio level data are used;
ISTRS=1 - selects the stress criterion weighted by the inverse of the sum of the p-th powered disparities and is related to the use of matrix conditional data;
POWER=2 - indicates least-squares estimation; and
EPS=0.001 - the default convergence criterion.

Levels of two of the parameters manipulated according to the design of the study were set according to the following:

NTRANS - the selected transformation function:
0 - squared distances,
1 - distances (no transformation),
2 - log of the distances); and
NDIM - the dimensionality of the solution.

Dimensionality was treated as a repeated measure, such that each simulation dataset were analyzed by **MSCAL** three times corresponding to: a) the true dimensionality or dimensionality of the true configuration, *NDIM=3;* b) the underfit case, *NDIM=2;* and c) the overfit case, *NDIM=4*. It was included in the design because a number of Monte Carlo studies have indicated a relationship between the criterion or stress function that is minimized and the dimensionality of the solution, as mentioned earlier.

2.7 Convergence of the Algorithm

The default convergence criteria in **MSCAL** is 0.001. Moreover, the documentation of the **MSCAL** program indicates that "iterations in MSCAL can have linear convergence properties. For this reason a relatively large value (say 0.001) should be used." (p.5 MSCAL documentation, 1988) Therefore, this default convergence value was used.

2.8 Dependent Variables

For metric scaling, **MSCAL** incorporates the dissimilarities directly into the criterion or stress function according to the data transformation or fitting criterion selected. Therefore, recovery of the proximities depends upon the estimation of the stimulus coordinates and optimization of the stress function. Both coordinates and proximities recovery were measured to assess performance of the **MSCAL** program. Mean-square error and bias were also calculated for true dimensionality estimates, as additional information with which to assess recovery.

2.8.1 Recovery of the Stimulus Configuration

Some means of measuring configuration recovery is necessary for dimensionalities other than the true one. **MATFIT** (Ramsay, 1989) allows the comparison of two matrices, using linear mappings into the necessary subspace of common variation. Specifically, **MATFIT** provides a correlational measure for any two configuration matrices without requiring either to be a fixed target.

The correlational measure used to optimally compare: **A**, the true configuration matrix of dimensionality a and **B**, the recovered configuration matrix of dimensionality b, is given by:

$$\rho^2_{AB} = \frac{tr(S'A'BT)^2}{tr(S'A'AS)\ tr(T'B'BT)},$$

where **S** and **T** are the two transformation matrices of dimensions a by s and b by s, respectively. The size s of the subspace is determined by: $1 \leq s \leq \min(a,b)$. Since MDS configurations can only be rotated, the two transformation matrices were constrained to be column orthonormal. **S** and **T** are defined by the singular value decomposition $A'B = SDT'$.

The mean-square error of the stimulus coordinates recovery when $b = a$ was calculated as follows:

$$MSE_{cf} = \frac{\sum_{r=1}^{25}\sum_{i=1}^{n}\sum_{k=1}^{a}(x_{ik}-y_{ikr})^2}{25na}$$

and

$$Y_r = BTS'N_r'$$

where **X** is the normalized true configuration matrix, N_r, the normalization matrix for replication r. Y_r is the associated matrix of standard scores of the transformed recovered configuration, and n is the number of stimuli. For both **X** and Y_r, there were independent normalizations of each dimension.

Similarly, squared bias was calculated as follows:

$$Bias^2{}_{cf} = \frac{\sum_{i=1}^{n}\sum_{k=1}^{a}(x_{ik}-\bar{y}_{ik})^2}{na}$$

where $\bar{Y} = \sum_{r=1}^{25} Y_i/25$.

2.8.2 Recovery of Proximities

The recovery of proximities was measured by correlating the recovered proximities, determined from each of the unnormalized, untransformed recovered configurations, with the *n(n-1)/2* true proximities. As for the coordinates recovery, mean-square error and squared bias were calculated. The mean square error for proximities is given by:

$$MSE_{pr} = \frac{\sum_{r=1}^{25}\sum_{i=2}^{n}\sum_{j=1}^{i-1}(d_{ij}-\delta_{ijr})^2}{r(n-1)n/2}$$

and the squared bias is as follows:

$$Bias^2{}_{pr} = \frac{\sum_{i=2}^{n}\sum_{j=1}^{i-1}(d_{ij}-\bar{\delta}_{ij})^2}{(n-1)n/2}$$

where **D** is the matrix of true inter-stimulus distances, $\bar{\Delta}$ is the matrix of mean recovered distances across *r=25* replications.

2.9 Summary of the Design

The design is mixed with two independent-groups factors: error model and predicted error model or fitting criterion, and one repeated-measures factor, dimensionality. The levels of both the error model and fitting criteria factors are the same: normal, lognormal and squared normal. The three dimensionalities are 2, 3 (the true dimensionality) and 4. Two sets of experiments were run using this design, corresponding to two different populations: the Coxon and the emotions configurations.

The distances used to simulate dissimilarities data were calculated from the corresponding true configuration using a unweighted Euclidean model. Error was added to each of the $n(n-1)/2$ inter-stimulus distances according to one of three error models. The levels of error for the Coxon and emotions data were approximately 17% and 52% error, respectively.

A replication factor of 25 was used to generate both sets of data samples, for each of the 9 treatment conditions, giving a total of 225 in each set. Each dataset in the Coxon batch with only one lower triangular 32 X 32 matrix of dissimilarities provided simulated input to **MSCAL** corresponding to a single subject. For the emotions set of data, there were 15 lower triangular 14 X 14 matrices of perturbed distances for each simulation of replicated data corresponding to 15 imaginary subjects' dissimilarities.

Ratio level dissimilarities were indicated for each **MSCAL** run, and estimation of the initial configuration using metric scaling was requested with optimization of the specified criterion function. Recovery estimates of the configuration and proximities were assessed according to their correlation to the true values. Mean-square error and squared-bias were provided as additional measures for estimates with true dimensionality.

Chapter 3

RESULTS FOR THE COXON DATA

Both of the dependent variables, coordinates recovery and proximities recovery, were analyzed with a 3 X 3 X 3 ANOVA, using a two-between and one-within design.

The distributions of cell means for the correlations between the true and recovered stimulus configurations indicated some non-normality and heterogeneous variances among the between-groups. Therefore, a transformation was used to give the following badness-of-fit index of recovery:

$$BOF = \log(1-r^2)$$

where r is the correlation measuring recovery.

3.1 Recovery of Stimulus Coordinates

According to Table 1, there was a triple interaction between dimensionality, true error model and the predicted error model. The interaction of the true and predicted error models was significant at each of the three dimensionalities. The results of the ANOVA are difficult to interpret since all the simple main effects were significant. Tukey post hoc tests were performed to discover any significant interactions of true and predicted error models for each dimensionality.

3.1.1 Recovery in 3-Dimensions

For the true dimensionality of the data, the best fit was produced when the criterion function matched the appropriate

Table 1

Stimulus space recovery Analysis of Variance
for Coxon Data on 32 Occupations

Effect	df	MS	F-ratio	
Between-datasets-effects				
True error models	2	1.02	3.7	$p<0.03$
Predicted error models	2	12.84	45.8	$p<0.01$
True X predicted models	4	22.20	79.2	$p<0.01$
"Between" denominator	216	0.28		
Within-datasets-effects				
Dimensions	2	14.17	99.1	$p<0.01$*
Dims X true error models	4	2.00	14.0	$p<0.01$*
Dims X pred. err. models	4	3.06	21.4	$p<0.01$*
Dims X true X pred. err.	8	0.98	6.8	$p<0.01$*
"Within" denominator	432	0.14		
Simple interaction effects				
True err X pred. err models				
at dim=2	4	4.84	66.9	$p<0.01$
Denominator	216	0.07		
at dim=3	4	13.10	54.8	$p<0.01$
Denominator	216	0.24		
at dim=4	4	6.21	24.4	$p<0.01$
Denominator	216	0.25		

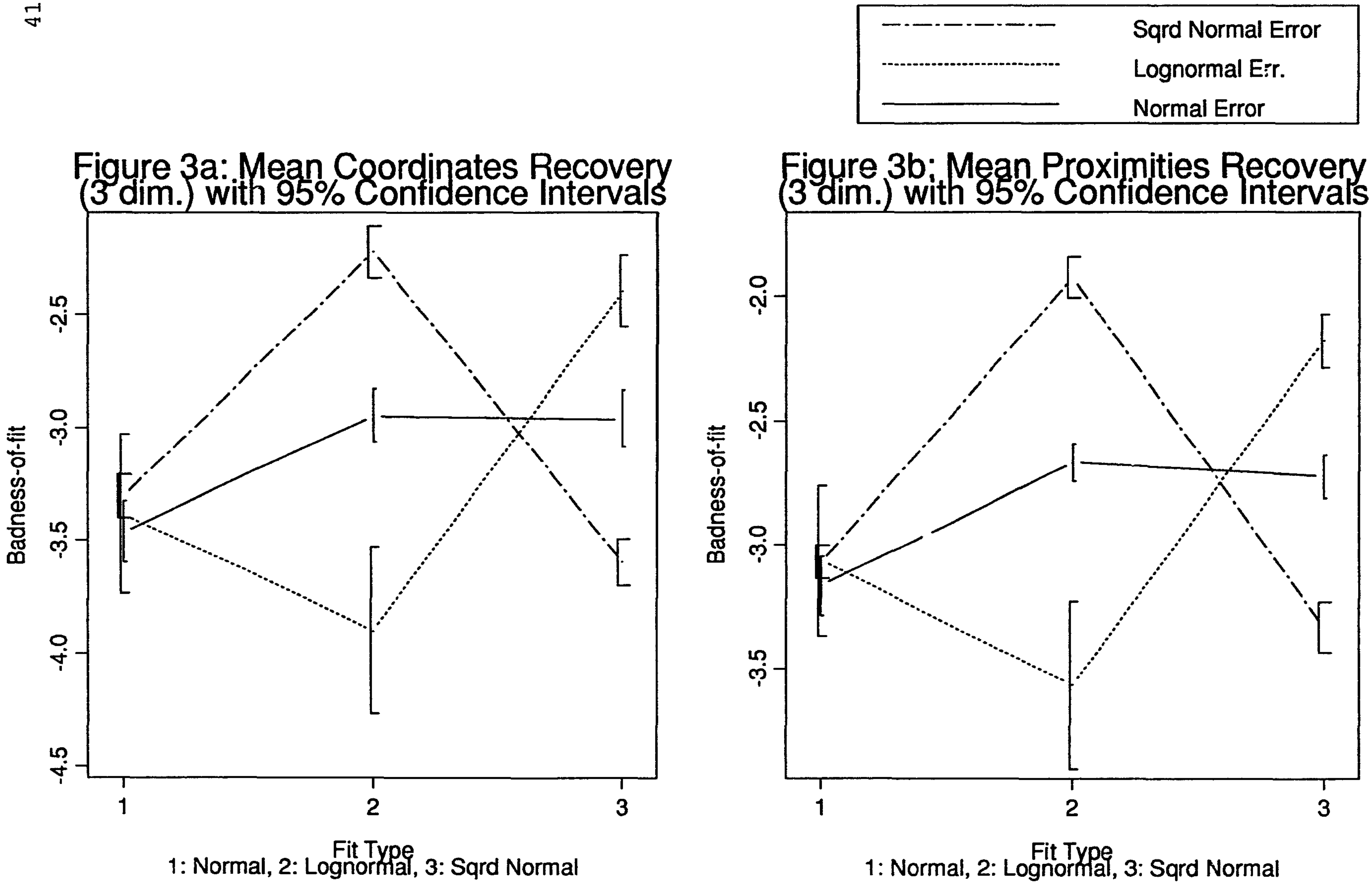
* using Greenhouse-Geisser conservative test.

underlying error model. However, Figure 3a shows that the recovery is about the same for all three underlying error distributions using a normal fit. In fact, there were no significant differences between these interactions ($p>.05$) in the context of the other post hoc pair-wise comparisons. However, the lognormal and squared normal fitting criteria, when used inappropriately, gave significantly worse ($p<.05$) recovery compared to that obtained when the fitting criterion matched the underlying error distribution.

3.1.2 Recovery in 2-Dimensions

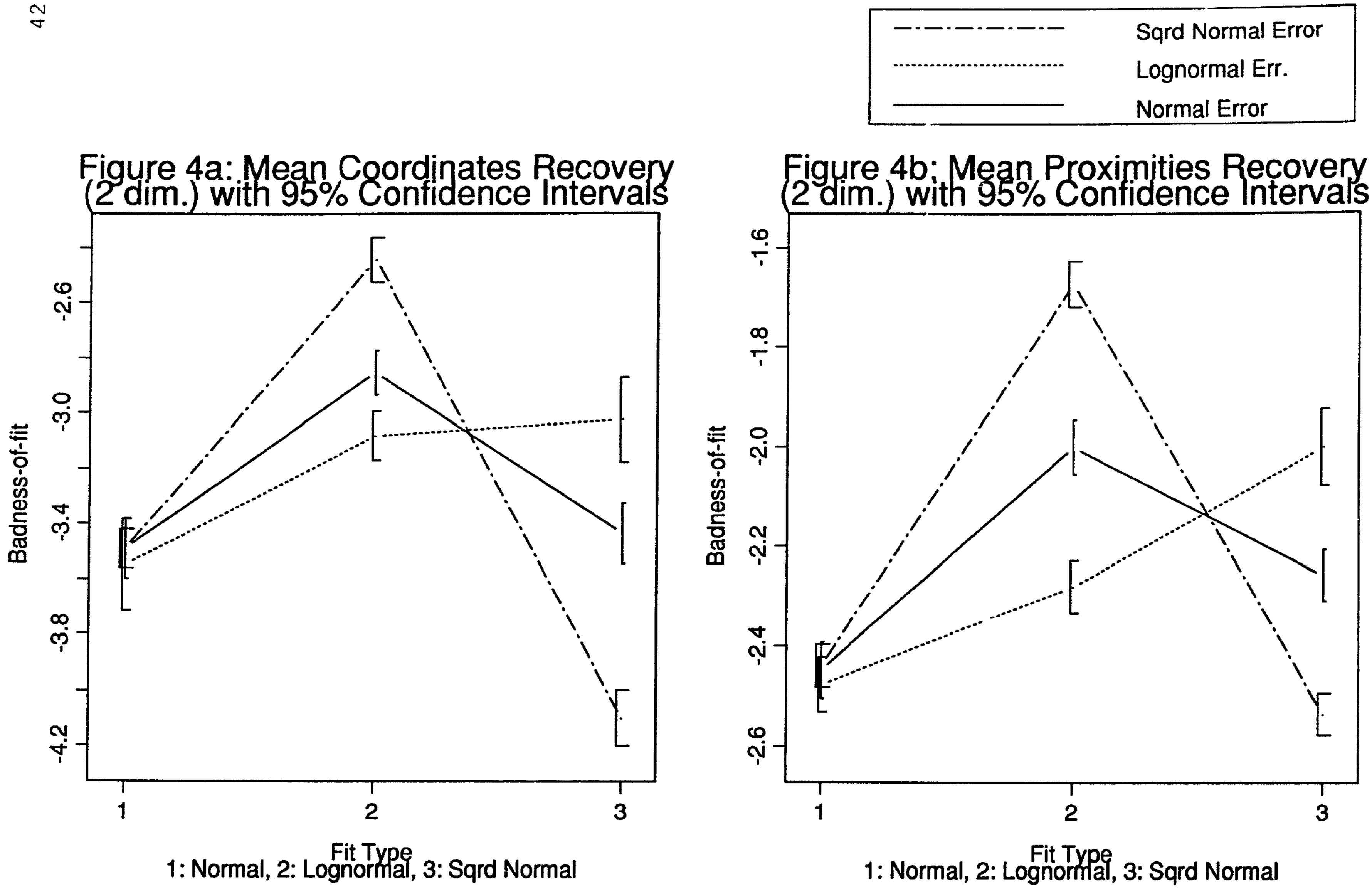
When the data were underfit, again a normal fit provided good recovery which was not significantly different for the three error distributions. The lognormal criterion gave the worst results of the three types of fit (see Figure 4a). In particular, it produced the worst fit in 2 dimensions, for squared normal data. A lognormal fit of lognormal data was not significantly better ($p>.05$) than when a normal or squared normal fit was used. Also, the normal fit of normal data was not significantly better ($p>.05$) than when a squared normal fit was used. Only the squared normal fit produced recovery that was significantly better ($p<.05$) for the appropriate error model. In fact, this combination gave the best fit in any of the 3 dimensionalities. Squared normal data showed a large reduction in recovery when a lognormal transformation was used.

Coxon Data



Sqrd Normal Error
Lognormal Err.
Normal Error
Figure 3a: Mean Coordinates Recovery (3 dim.) with 95% Confidence Intervals
Badness-of-fit
-2.5
-3.0
-3.5
-4.0
-4.5
1
2
3
Fit Type
1: Normal, 2: Lognormal, 3: Sqrd Normal
Figure 3b: Mean Proximities Recovery (3 dim.) with 95% Confidence Intervals
Badness-of-fit
-2.0
-2.5
-3.0
-3.5
1
2
3
Fit Type
1: Normal, 2: Lognormal, 3: Sqrd Normal

Coxon Data


Sqrd Normal Error
Lognormal Err.
Normal Error
Figure 4a: Mean Coordinates Recovery (2 dim.) with 95% Confidence Intervals
Badness-of-fit
-2.6
-3.0
-3.4
-3.8
-4.2
1
2
3
Fit Type
1: Normal, 2: Lognormal, 3: Sqrd Normal
Figure 4b: Mean Proximities Recovery (2 dim.) with 95% Confidence Intervals
Badness-of-fit
-1.6
-1.8
-2.0
-2.2
-2.4
-2.6
1
2
3
Fit Type
1: Normal, 2: Lognormal, 3: Sqrd Normal

3.1 3 Recovery in 4-Dimensions

From Figure 5a, it is apparent that normal and lognormal error distributions were equally poorly fit, regardless of the fitting criterion used. There was one exception: normal data were significantly better fit ($p<.05$) by a squared normal criterion than were the lognormal data. The squared normal model gave the best fit for the appropriate error distribution, comparable to that given for dimensionality three. However, it was not a significant improvement over a normal fit of the squared normal data. Only the lognormal fitting criterion gave poor recovery with squared normal error.

Using higher dimensionality produced solutions with the worst recovery for normal and lognormal data, relative to the other two dimensions used. The recovery of squared normal data was relatively insensitive to dimensionality. In addition, a lognormal fit of squared normal data consistently gave the worst recovery in each of the three solution dimensions.

3.1.4 Mean-Square Error for 3-Dimensions

It is apparent from Figure 6a that amount of the stimulus recovery in three dimensions was satisfactory for the amount of error in the data, mean PVE=0.17. The maximum mean squared discrepancy was approximately 0.24 between the normalized true and recovered configurations. However, the largest mean-square error for a normal fit was 0.12, indicating much less discrepancy for this particular fit. On average, the variance of the

Coxon Data

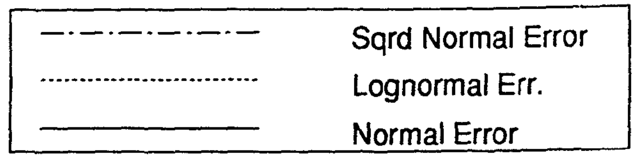


Figure 5a: Mean Coordinates Recovery (4 dim.) with 95% Confidence Intervals

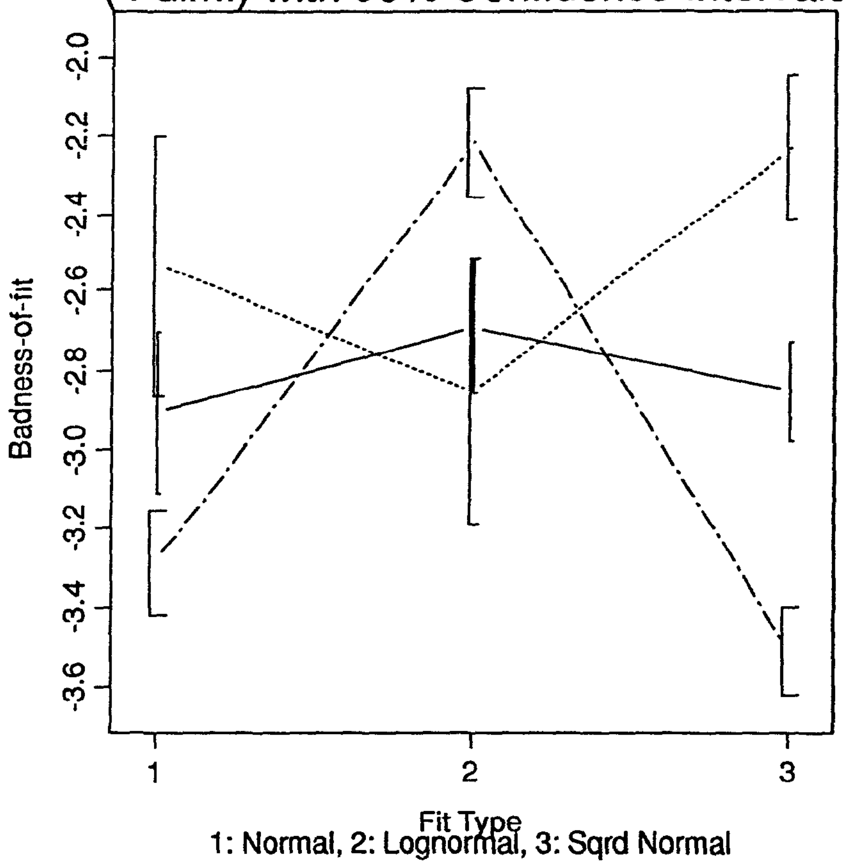
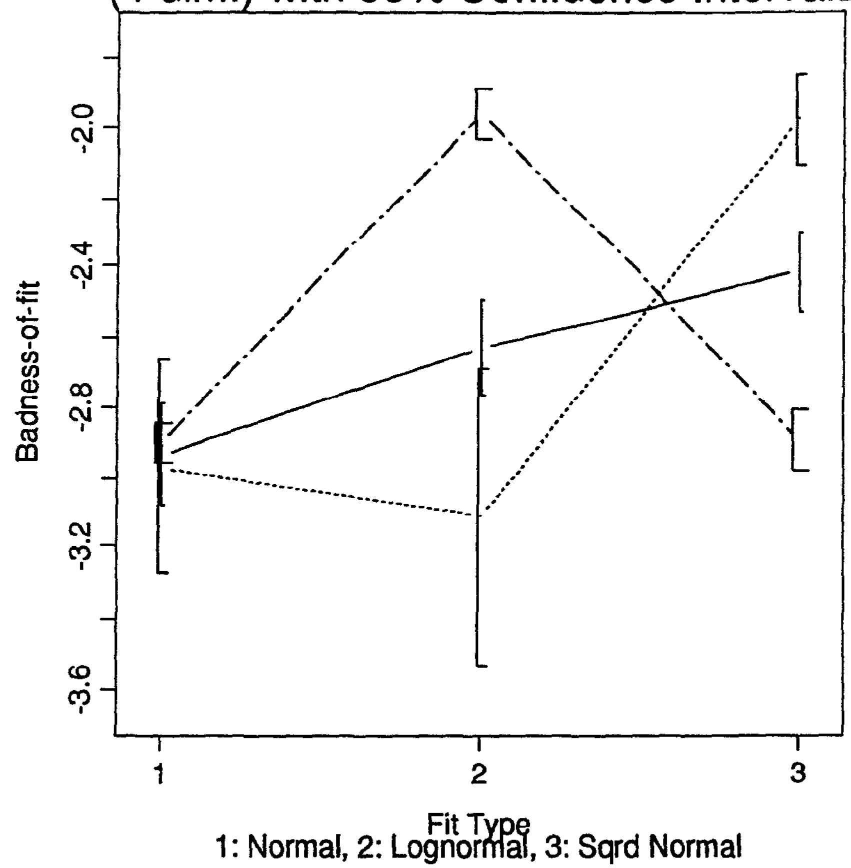


Figure 5b: Mean Proximities Recovery (4 dim.) with 95% Confidence Intervals



Coxon Data

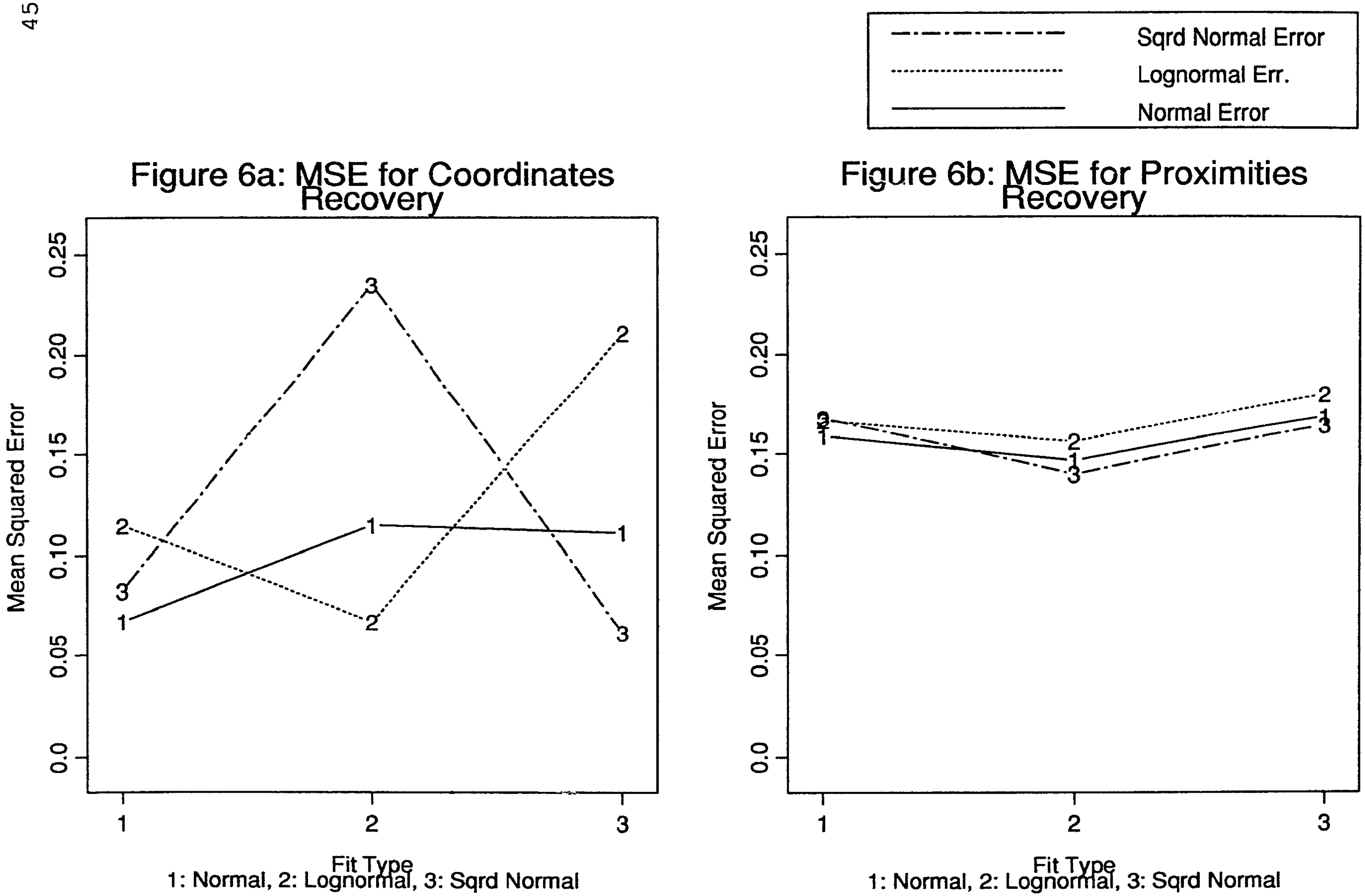


Figure 6a: MSE for Coordinates Recovery

Figure 6b: MSE for Proximities Recovery

Coxon Data

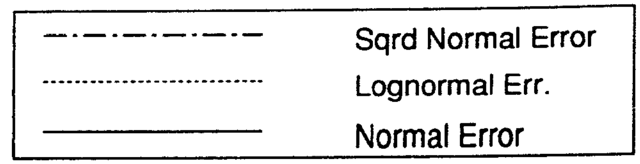


Figure 7a: Variance of Est./MSE for Coordinates

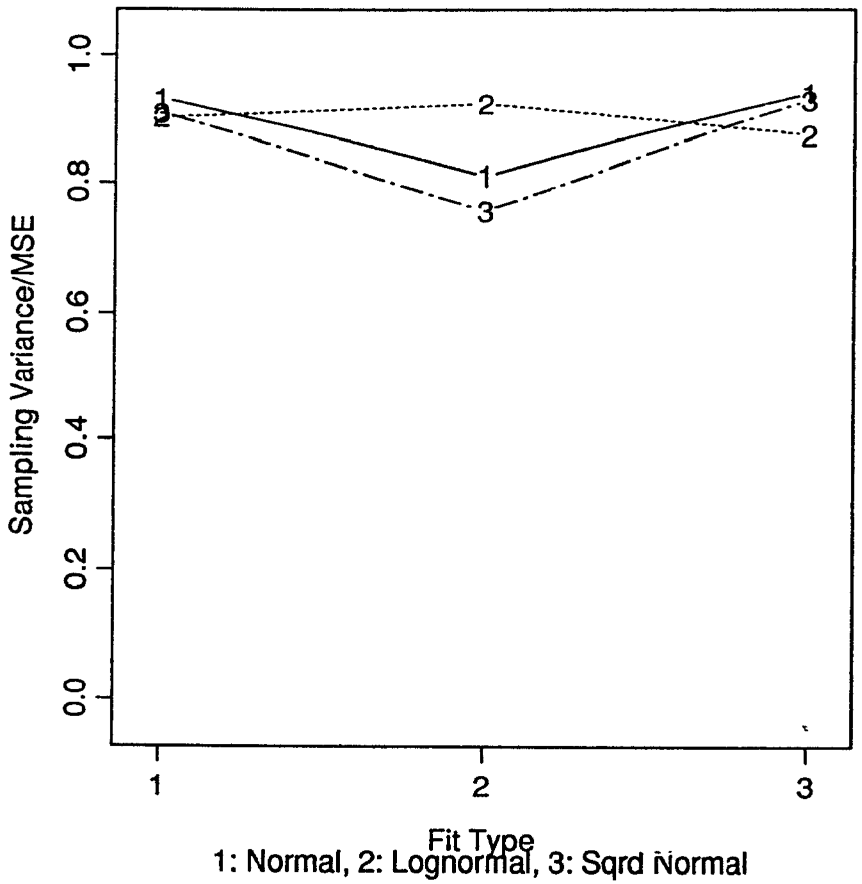
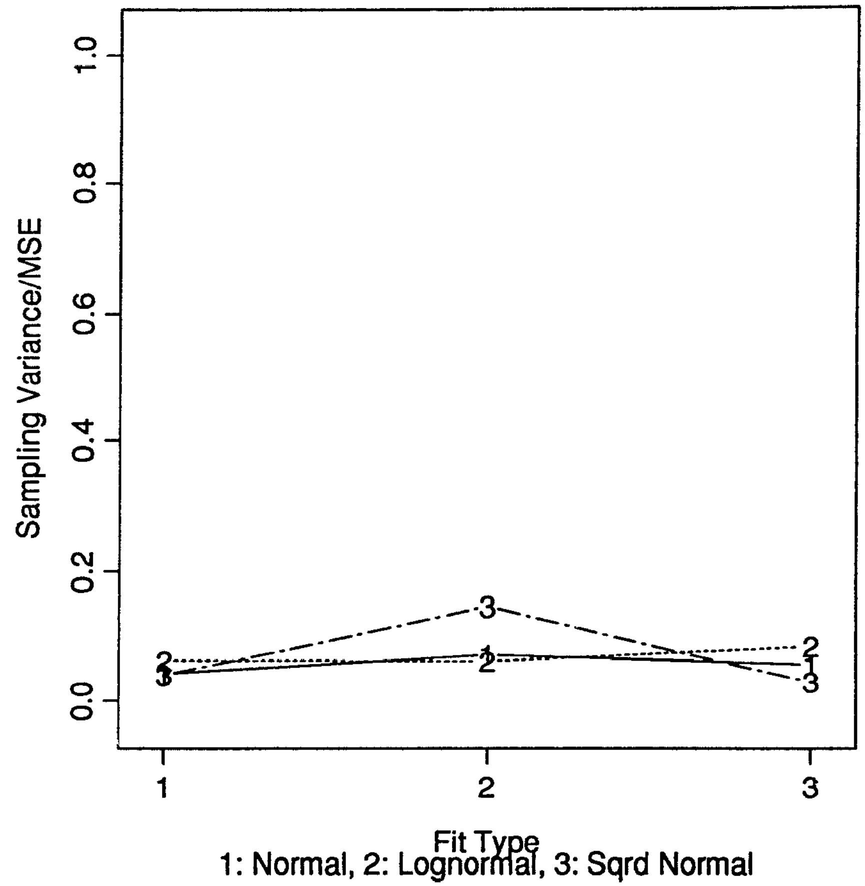


Figure 7b: Variance of Est./MSE for Proximities



estimate accounted for approximately 90% of the mean-square error (see Figure 7a). The largest ratio of squared bias to mean-square error was 0.24 for squared normal data fit by a lognormal model, the worst recovery condition as can be seen in Figure 3a.

3.1.5 Loose Convergence

Solutions which were poor relative (i.e. recovery correlations < 0.9) to the other replications were rerun using a criterion of 0.0001, an order of magnitude higher than the default, to ensure that instances of poorer recovery were not due to a loose convergence. Since two measures of recovery were used, reruns were made for any replication whose coordinate and/or proximity recovery correlations were less than 0.9. Both correlations were recalculated; however, only coordinates recovery are reported here.

A total of 33 replications were redone, seven of which had both coordinates and proximities recovery correlations less than 0.9. Table 2 indicates the number of error model/fitting criterion conditions which had coordinates recovery correlations less than 0.9. Rerunning these analyses with a tighter convergence criterion made little difference. The largest increase in recovery correlation for a given error and fitting criterion combination in any dimension was a 0.001.

Table 2

Number of MSCAL Coordinates Recovery Correlations
Less Than 0.9 for Coxon Data
by Error/Fitting Criteria Condition and Dimensionality

Condition Error/Fitting Criteria	Dimensionality 2	3	4
Norm./Norm.	0	0	0
Norm./Lognorm.	0	0	0
Norm./Sq. Norm.	0	0	0
Lognorm./Norm.	0	2	5
Lognorm./Lognorm.	0	1	1
Lognorm./Sq. Norm.	0	0	2
Sq. Norm./Norm.	0	0	0
Sq. Norm./Lognorm.	0	0	1
Sq. Norm./Sq. Norm.	0	0	0

3.1.6 Starting Configurations

A good starting configuration is generally accepted as one means for improving recovery. To this end, the analyses of replication 5, the worst case coordinates recovery in three dimensions, were repeated as before but the true configuration was the starting configuration. Table 3 indicates that the starting configuration had a large effect on recovery. For all three dimensionalities the recovery correlations were almost 1.0 with a perfect start. This particular set of data was perhaps less well-defined for some stimuli, given that the Coxon data is not replicated.

3.2 Recovery of Proximities

There was a similar picture produced for the recovery of proximities. Figures 3b and 4b for two- and three- dimensional recovery of proximities closely parallel their counterparts in coordinates recovery. Yet, the magnitude of the recovery in these two dimensionalities is less for proximities. However, when the data were overfit the results differed from those corresponding to coordinates recovery. In particular, the lognormal criterion did not fare as badly, nor squared normal fitting fare as well, as in the case of coordinates recovery (see Figures 5a & 5b). Furthermore, recovery was comparable in magnitude to that of the coordinates. Lognormal fit gave the best recovery of lognormal data which was significantly better ($p<.05$) than for either of the other two types of error. This

Table 3

MSCAL Coordinates Recovery Correlations of a Single Replication for the Worst Case Error/Fitting Criteria Condition of Dimensionality 3 with Default and Perfect Starts

Data	Start	Dimensionality		
		2	3	4
Emotions*	Default	0.97	0.89	0.87
	Perfect	0.87	1.00	1.00
Coxon+	Default	0.98	0.85	0.91
	Perfect	0.99	0.99	0.99

Note: Results are rounded to 2 decimal points.

* Worst case is replication 19 for lognormal error/lognormal fit for dimensionality 3.
+ Worst case is replication 5 for lognormal error/normal fit for dimensionality 3.

Table 4

Proximities Recovery Analysis of Variance
for Coxon Data on 32 Occupations

Effect	df	MS	F-ratio	
Between-datasets-effects				
True error models	2	0.57	2.8	p<0.06
Predicted error models	2	10.83	52.6	p<0.01
True X predicted models	4	18.82	91.5	p<0.01
"Between" denominator	216	0.21		
Within-datasets-effects				
Dimensions	2	22.53	211.7	p<0.01*
Dims X true error models	4	0.08	0.7	p<0.56*
Dims X pred. err. models	4	0.91	8.5	p<0.01*
Dims X true X pred. err.	8	1.14	10.7	p<0.01*
"Within" denominator	432	0.11		
Simple interaction effects				
True err X pred. err models				
at dim=2	4	2.05	120.6	p<0.01
Denominator	216	0.02		
at dim=3	4	12.35	72.0	p<0.01
Denominator	216	0.17		
at dim=4	4	6.70	29.2	p<0.01
Denominator	216	0.23		

* using Greenhouse-Geisser conservative test.

contrasts with the corresponding case in coordinates recovery. Moreover, for overfit data, only the lognormal criterion produced significantly better recovery ($p<.05$) when the appropriate model was used to fit the data. These results were very similar to those of proximities recovery in three dimensions, although recovery was better in three dimensions. Overall, recovery of proximities was the worst in two dimensions.

3.2.1 Mean-Square Error for 3-Dimensions

Figure 6b indicates that the recovery of proximities was good. The mean squared discrepancy showed little variation across the various treatment conditions and was approximately 0.16 for the unnormalized proximities (i.e. proximity magnitudes were less than 2.6). Squared bias accounted for most of the discrepancy, which was approximately 90% of the mean-square error (see Figure 8b), unlike the recovery of coordinates which had little squared bias with respect to the mean-square error.

3.2.2 Loose Convergence

Solutions with a proximities recovery correlation less than 0.9 were rerun using a tighter convergence criteria of 0.0001, as in the case of coordinates recovery. Also, the corresponding proximities recovery correlations recalculated for coordinates recovery that were rerun are included here. Table 5 shows the number of replications with recovery

Coxon Data

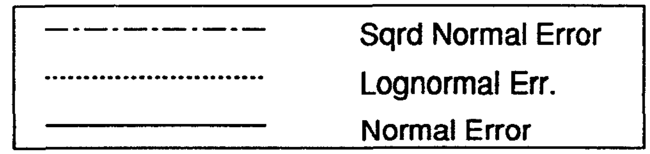


Figure 8a: Squared Bias/MSE for Coordinates

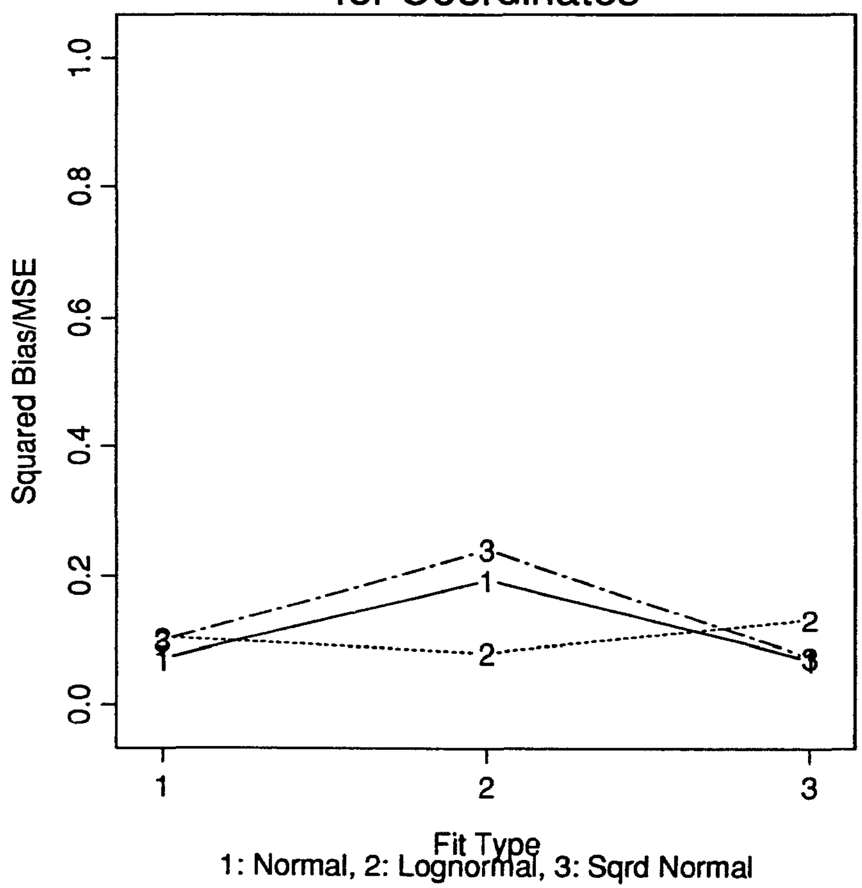


Figure 8b: Squared Bias/MSE for Proximities

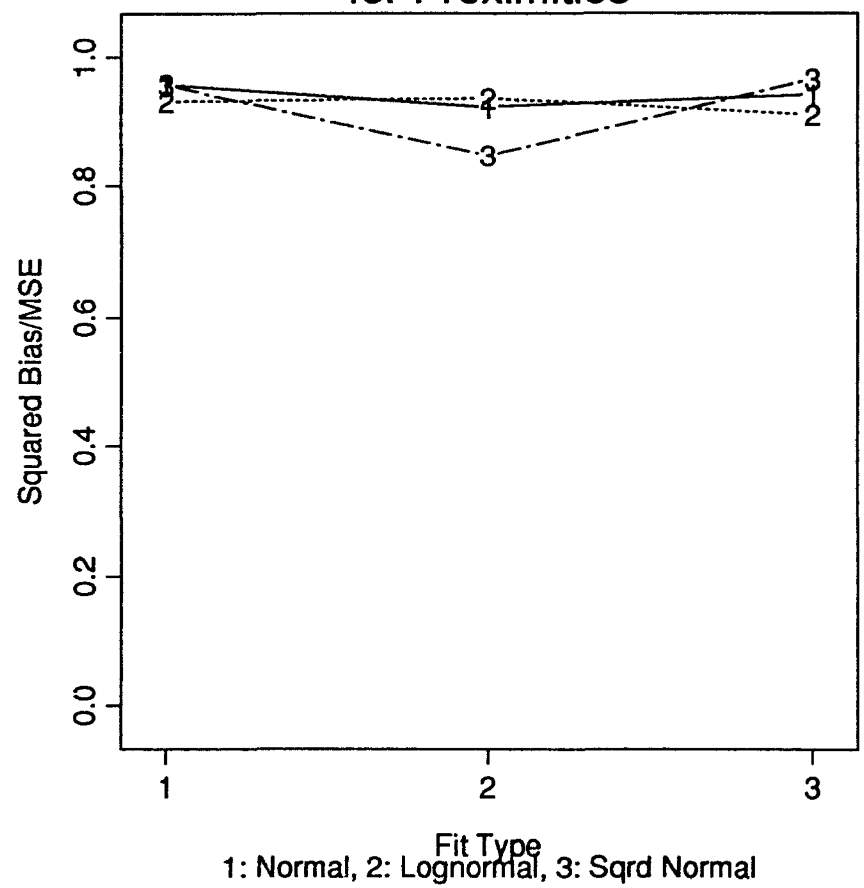


Table 5

Number of MSCAL Proximities Recovery Correlations
Less Than 0.9 for Coxon Data
by Error/Fitting Criteria Condition and Dimensionality

Condition Error/Fitting Criteria	Dimensionality		
	2	3	4
Norm./Norm.	0	0	0
Norm./Lognorm.	0	0	0
Norm./Sq. Norm.	0	0	0
Lognorm./Norm.	0	2	2
Lognorm./Lognorm.	0	1	3
Lognorm./Sq. Norm.	1	1	5
Sq. Norm./Norm.	0	0	0
Sq. Norm./Lognorm.	3	0	0
Sq. Norm./Sq. Norm.	0	0	0

correlations less than 0.9. The use of a tighter convergence criterion gave little improvement in recovery. In fact, the correlations often worsened. The single largest increase for a given combination of error and fitting criterion was a 0.01.

3.2.3 Starting Configurations

Results were consistent with those obtained for stimulus coordinates recovery using a perfect starting configuration for the same replication and are shown in Table 6. However, only for dimensionality 3 was there an improvement in the corresponding proximities recovery correlation. It appears that the starting configuration has less effect on proximities recovery than for configuration recovery in this case. It is reasonable to expect that providing an initial configuration will benefit coordinates recovery more than proximities recovery.

3.3 MSCAL Performance

Table 7 contains some overall measures of performance. The squared correlations for coordinates recovery were good in all 3 dimensions; they were slightly less so for proximities recovery. The mean elapsed times were mostly less than 1 CPU second. There are two notable exceptions: lognormal fit and squared normal fit, of normal data in 4 dimensions. This is surprising because these conditions did not show the worst recovery. Nevertheless, these times do indicate that the

Table 6

MSCAL Proximities Recovery Correlations of a Single Replication for the Worst Case Error/Fitting Criteria Condition of Dimensionality 3 with Default and Perfect Starts

Data	Start	Dimensionality		
		2	3	4
Emotions*	Default	0.92	0.91	0.90
	Perfect	0.61	1.00	1.00
Coxon+	Default	0.96	0.88	0.98
	Perfect	0.96	0.98	0.98

Note: Results are rounded to 2 decimal points.

* Worst case is replication 19 for lognormal error/lognormal fit for dimensionality 3.

+ Worst case is replication 5 for lognormal error/normal fit for dimensionality 3.

Table 7

MSCAL Mean Performance on 27 datasets
across 25 replications
for Coxon Data for 32 Occupations

True error model	Fit model	Dim	Squared Correlations* Coordinates	Proximities	Elapsed Time** in secs.	No.+ Iter.
Normal	Normal	2	0.97	0.91	0.63	3.8
Normal	Lognormal	2	0.94	0.86	0.84	9.6
Normal	Sq.normal	2	0.97	0.89	0.51	3.2
Lognormal	Lognormal	2	0.95	0.90	0.30	8.8
Lognormal	Normal	2	0.97	0.92	0.05	5.0
Lognormal	Sq.normal	2	0.95	0.86	0.08	3.3
Sq.normal	Sq.normal	2	0.98	0.92	0.36	3.1
Sq.normal	Normal	2	0.97	0.91	0.22	3.9
Sq.normal	Lognormal	2	0.91	0.81	0.54	9.6
Normal	Normal	3	0.97	0.96	0.78	3.9
Normal	Lognormal	3	0.95	0.93	0.80	10.9
Normal	Sq.normal	3	0.95	0.93	1.37	3.2
Lognormal	Lognormal	3	0.97	0.96	0.56	7.1
Lognormal	Normal	3	0.95	0.94	0.77	4.5
Lognormal	Sq.normal	3	0.90	0.88	0.57	3.6
Sq.normal	Sq.normal	3	0.97	0.96	0.85	3.0
Sq.normal	Normal	3	0.96	0.95	1.02	3.6
Sq.normal	Lognormal	3	0.89	0.85	0.16	9.6
Normal	Normal	4	0.94	0.94	0.28	5.2
Normal	Lognormal	4	0.93	0.92	0.15	14.0
Normal	Sq.normal	4	0.94	0.91	0.49	3.1
Lognormal	Lognormal	4	0.92	0.92	0.12	8.0
Lognormal	Normal	4	0.89	0.93	2.65	4.8
Lognormal	Sq.normal	4	0.88	0.93	0.20	3.9
Sq.normal	Sq.normal	4	0.97	0.94	1.50	2.9
Sq.normal	Normal	4	0.96	0.94	2.22	4.1
Sq.normal	Lognormal	4	0.88	0.86	1.89	9.6

* These are mean R-Squared values for coordinate & proximities recovery.
**These are mean elapsed times of execution of the DMSCAL subroutine.
\+ These are mean number of iterations.

algorithm worked well on the SUN-3/60 without use of the floating-point accelerator. It is interesting to note for all 3 dimensionalities that larger mean number of iterations were necessary for lognormal fit; however, this was not always reflected in the mean elapsed times.

The average ratio of mean elapsed time to mean number of iterations across the nine conditions was 0.09, 0.19 and 0.26 for dimensionality of 2, 3, and 4 respectively, increasing in an essentially uniform manner with dimensionality, as expected.

Chapter 4

RESULTS FOR THE EMOTIONS DATA

The dependent variables, coordinates recovery and proximities recovery, were analyzed with a 3 X 3 X 3 ANOVA, using a two-between and one-within design.

The distributions of cell means for the correlations between the true and recovered stimulus indicated some non-normality and heterogeneous variances among the between-groups. Therefore, a transformation was employed, using the same badness-of-fit index of recovery as the Coxon data.

4.1 Recovery of Stimulus Coordinates

There was a triple interaction between dimensionality, true error model and the predicted error model (see Table 8). The simple interactions of the true and predicted error models were significant for each of the three dimensions used. Tukey post hoc tests were done on these two-way interactions and the results are explained below.

4.1.1 Recovery in 3-Dimensions

For the true dimensionality, the best fit was produced when the criterion function matched the appropriate underlying error model in only the squared normal case. The best recovery using a normal fitting criterion was given for squared normal data but it was not significantly different ($p>.05$) from that

Table 8

Stimulus space recovery Analysis of Variance for Emotions Data from MDS Workshop Members

Effect	df	MS	F-ratio	
Between-datasets-effects				
True error models	2	12.64	37.1	$p<0.01$
Predicted error models	2	11.48	33.7	$p<0.01$
True X predicted models	4	3.28	9.6	$p<0.01$
"Between" denominator	216	0.34		
Within-datasets-effects				
Dimensions	2	5.31	37.2	$p<0.01$*
Dims X true error models	4	2.62	18.3	$p<0.01$*
Dims X pred. err. models	4	2.83	19.8	$p<0.01$*
Dims X true X pred. err.	8	0.38	2.7	$p<0.01$*
"Within" denominator	432	0.14		
Simple main effects				
True err X pred. err models				
at dim=2	4	0.68	6.1	$p<0.01$
Denominator	216	0.11		
at dim=3	4	2.53	8.4	$p<0.01$
Denominator	216	0.30		
at dim=4	4	0.84	4.0	$p<0.01$
Denominator	216	0.21		

* using Greenhouse-Geisser conservative test.

for normal data (see Figure 9a). There were no significant differences ($p>.05$) between the lognormal fit of the three data types.

The worst recovery was for lognormal data using a squared normal fit; however, there were no significant improvements ($p>.05$) in fit for lognormal data using a normal or lognormal error model (see Figure 9a). It appears that squared normal data were recovered equally well using the distance and squared distance transformation but recovery was significantly worse ($p<.05$) using the lognormal transformation. Normal data were best fit by the normal model but not significantly ($p>.05$) better than by the lognormal or squared normal models.

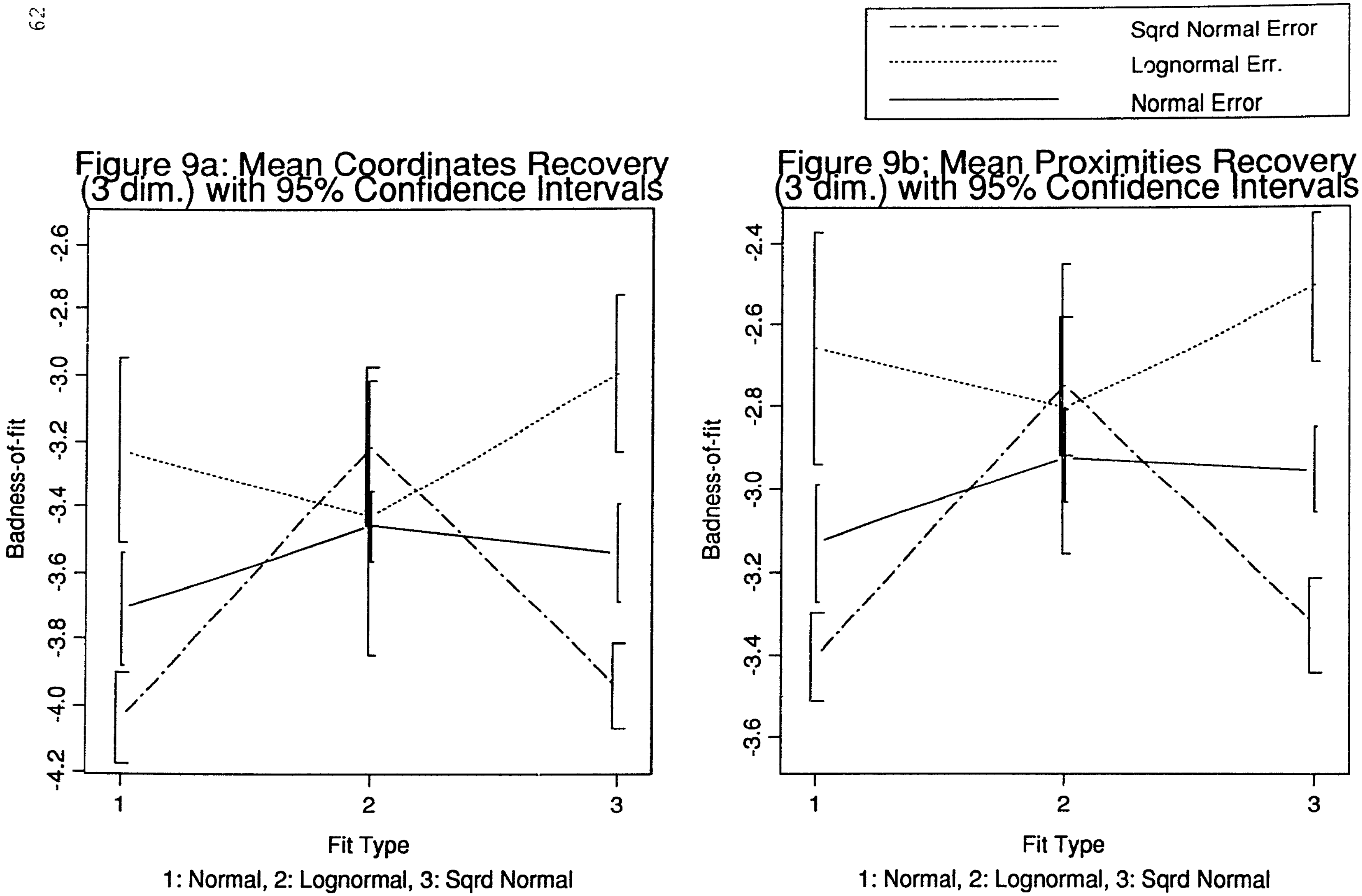
4.1.2 Recovery in 2-Dimensions

When the data were underfit, the lognormal fitting was the worst of the three types of fit and recovery was significantly worse ($p<.05$) than for the other two fitting criteria. Also, all three types of data were equally poorly fit ($p>0.5$) by the lognormal model (see Figure 10a). There were no significant differences in recovery using a normal fit for the three data types. Squared normal fit gave the best fit overall in two dimensions. The fit was significantly better ($p<.05$) for squared normal data than lognormal data but not better than for normal data.

4.1.3 Recovery in 4-Dimensions

Overfitting in higher dimensionality produced clearcut results for the three data types. From Figure 11a, it is

Emotions Data



Emotions Data

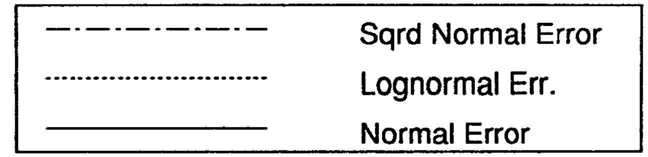


Figure 10a: Mean Coordinates Recovery (2 dim.) with 95% Confidence Intervals

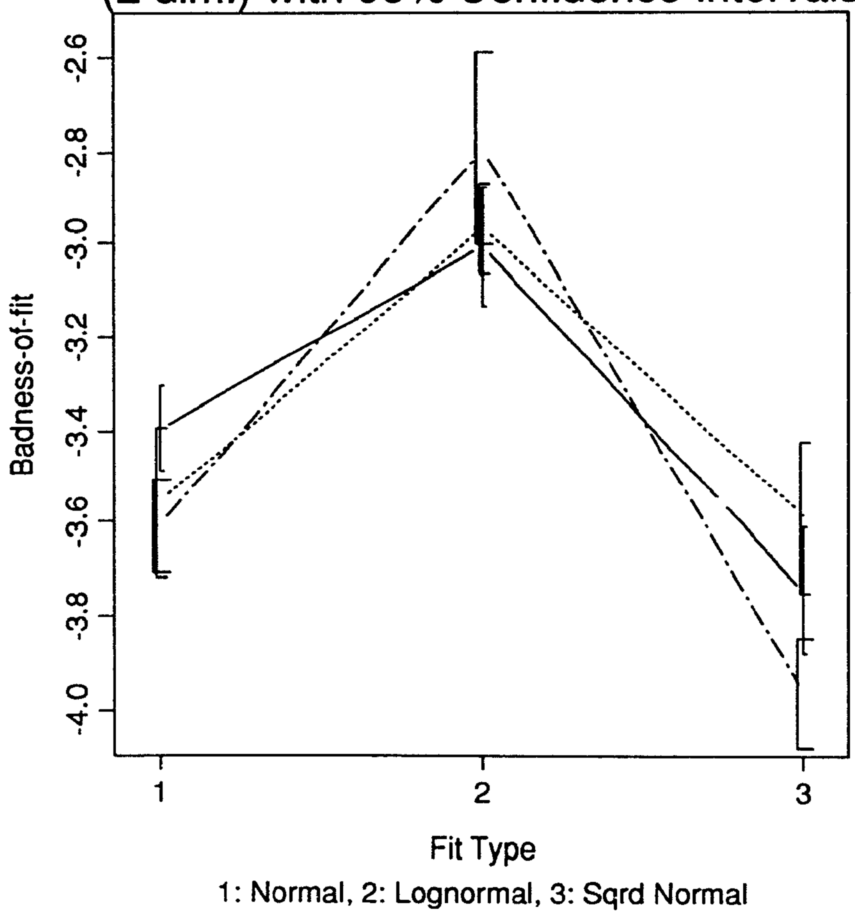
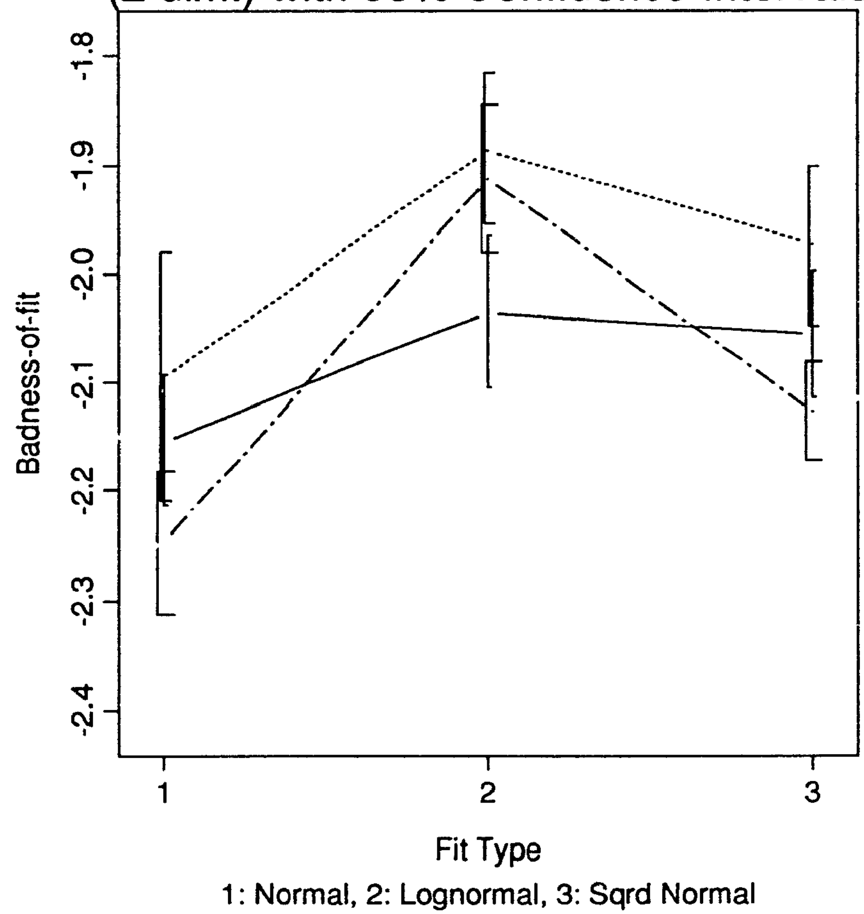


Figure 10b: Mean Proximities Recovery (2 dim.) with 95% Confidence Intervals



apparent that lognormal data had poor recovery regardless of the fitting criteria used, and differences were nonsignificant ($p>.05$). Similarly, there were no significant differences ($p>.05$) in recovery for normal data using the three fitting criteria. Squared normal data was the best recovered in comparison to the other data types. Squared normal data were equally well fit using the normal model as the squared normal model ($p>.05$) but showed significantly better recovery than when a lognormal fit was used. Matching fitting criteria to the error distributions did not improve recovery in four dimensions.

4.1.4 Mean-Square Error for 3-Dimensions

Figure 12a indicates that the amount of stimulus recovery was more than adequate given the amount of error in the data, mean PVE=0.52. The maximum discrepancy was approximately 0.12 for the normalized recovery configurations. Also, the mean-square error for normal data was constant at 0.05 for the three fitting criteria. Squared normal data had the lowest MSE, 0.03, using either a normal or squared normal fitting criterion. This discrepancy was slightly higher using a lognormal fit.

Most of the mean-square discrepancy was accounted for by the variance, which was approximately 90% of the mean-square error for all error and fitting criterion conditions. Since

Emotions Data

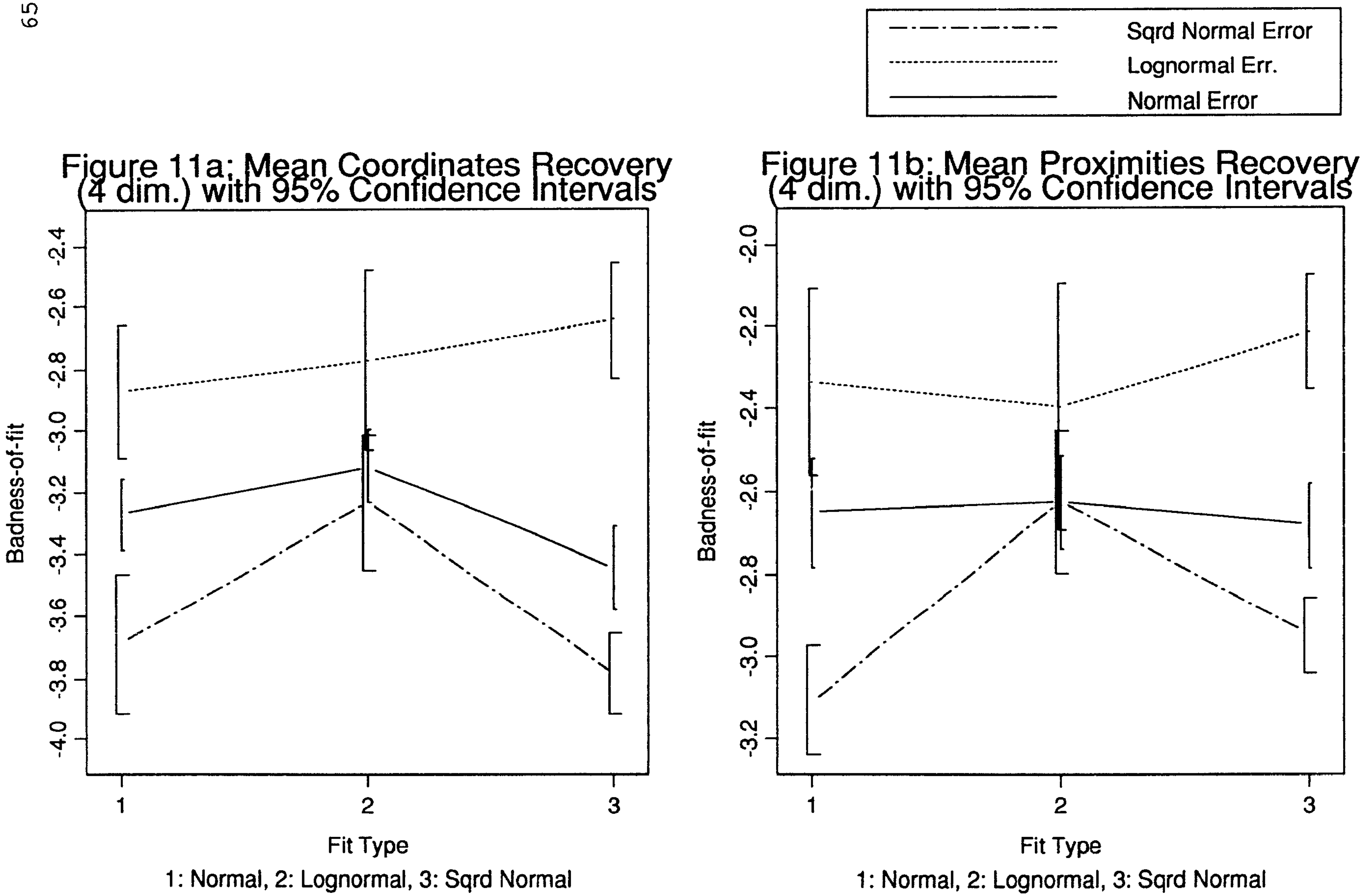


Figure 11a: Mean Coordinates Recovery (4 dim.) with 95% Confidence Intervals

Figure 11b: Mean Proximities Recovery (4 dim.) with 95% Confidence Intervals

Emotions Data

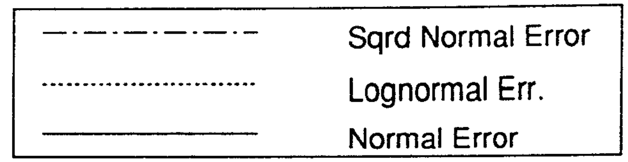


Figure 12a: MSE for Coordinates Recovery

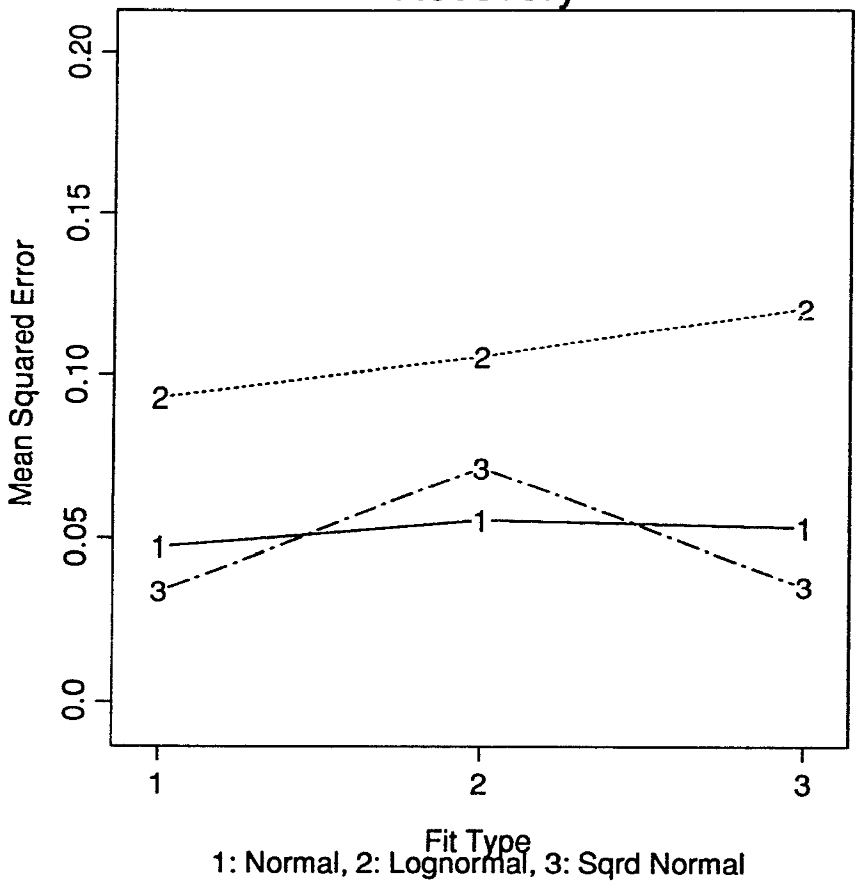
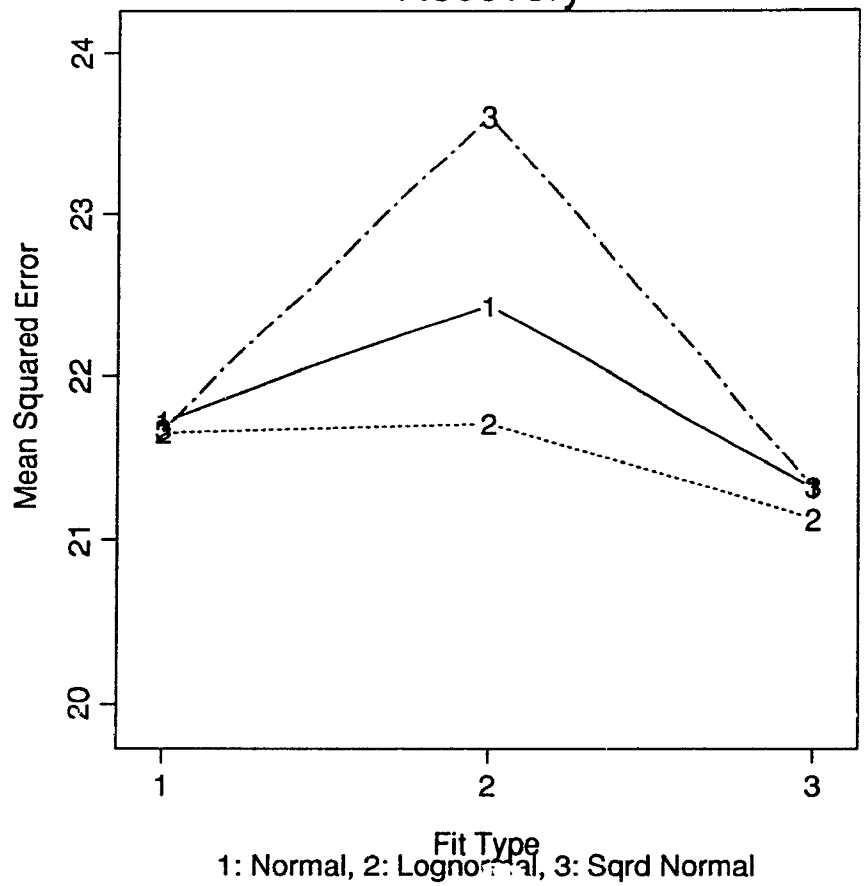


Figure 12b: MSE for Proximities Recovery



the mean-square error was small, the amount of bias was considered negligible in the coordinates recovery in three dimensions.

4.1.5 Loose Convergence

Solutions with recovery correlations, for either coordinates or proximities, less than 0.9 were rerun using a convergence criterion of 0.0001. For dimension 2, there were no coordinates recovery correlations less than 0.9. Table 9 contains the number of replications for which the recovery correlations were below 0.9. Rerunning using a tighter convergence criteria made little difference. The largest increase for a given combination of error and fitting criterion was 0.009. Correlations often decreased rather than increased.

4.1.6 Starting Configurations

Improvement in recovery was sought by providing a "good" starting configuration. The analyses of replication 19 for the lognormal error and fit condition were repeated with the true configuration input as a starting configuration. This replication was chosen because it gave the worst coordinates recovery correlation for dimensionality 3.

A perfect starting configuration improved recovery for dimensionalities 3 and 4 such that perfect correlations resulted (see Table 3). However, the recovery correlation in two dimensions worsened when a perfect starting configuration was provided. This is in sharp contrast to the improvement a

Table 9

Number of MSCAL Coordinates Recovery Correlations
Less Than 0.9 for Emotions Data
by Error/Fitting Criteria Condition and Dimensionality

Condition Error/Fitting Criteria	Dimensionality		
	2	3	4
Norm./Norm.	0	0	0
Norm./Lognorm.	0	0	0
Norm./Sq. Norm.	0	0	0
Lognorm./Norm.	0	0	0
Lognorm./Lognorm.	0	1	1
Lognorm./Sq. Norm.	0	1	2
Sq. Norm./Norm.	0	0	0
Sq. Norm./Lognorm.	0	0	1
Sq. Norm./Sq. Norm.	0	0	0

perfect starting configuration gave to the worst case replication of the Coxon data. It seems unlikely that the starting configuration itself accounted for the reduction in recovery.

It was not unexpected that the worst case of recovery in three dimensions, for both the Coxon and emotions data, involved lognormal data since error is proportional to the magnitude of the true distances. Presumably it was the error distribution for the particular replication that led to a sub-optimal local minimum.

Figures 13 and 14 plot the relative error against the true distances for the worst case replications of the Coxon and emotions data, respectively; the error distances were the differences between the true distances and the error-perturbed distances. The two statistics provided for each plot are the proportion-of-variance of error (PVE) and the correlation between the absolute magnitude of error and the true distances (FCE). A positive FCE value indicates the extent to which larger error is present in larger distances. The FCE values indicate that more error was present in the larger distances in the emotions data, for the two replications considered.

It is apparent from Figure 13 that the relative error for replication #5 of the Coxon data was less than 1 and evenly distributed across all true distances. On the other hand, Figure 14 of the emotions data indicates that replication 19, with lognormal error and the matching fitting criteria, had

Plots of Relative Error to True Distances for Single Replications

Figure 13: Replication #5 of the Coxon Data for 1 Subject with Lognormal Error/Normal Fit

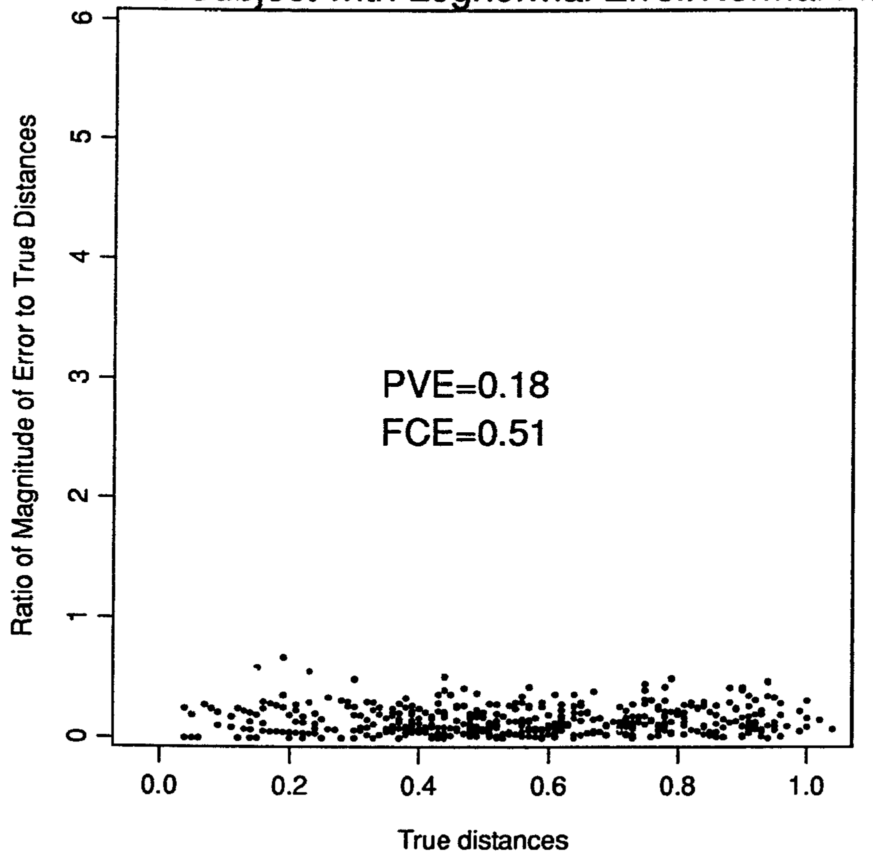
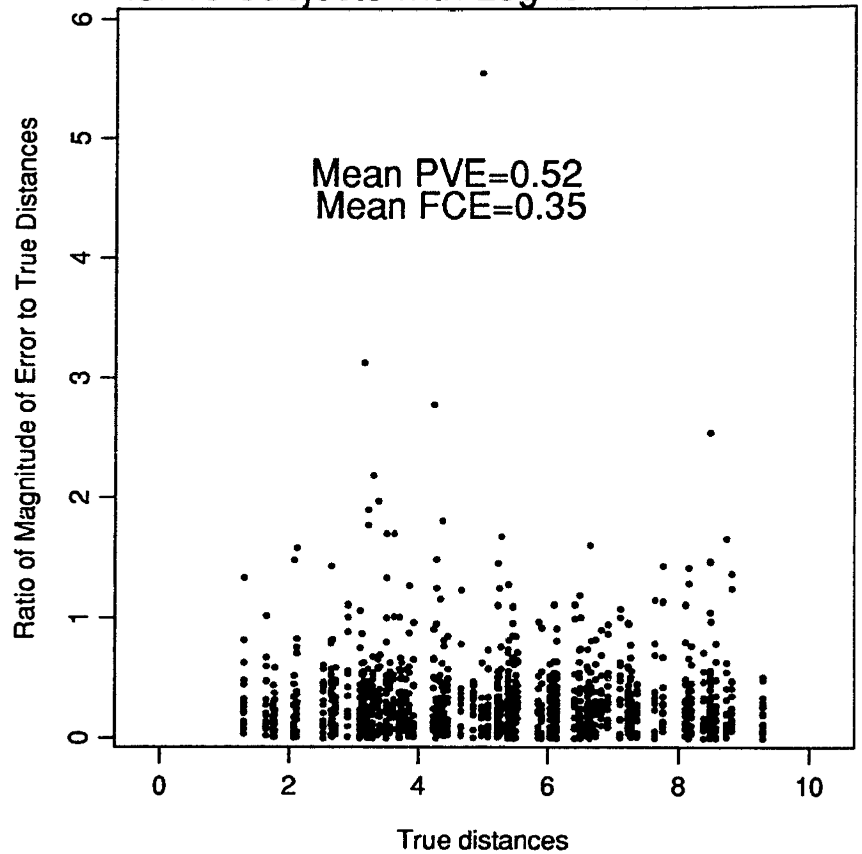


Figure 14: Replication #19 of the Emotions Data for 15 Subjects with Lognormal Error & Fit



very large error components: the largest relative error was approximately 5.5 times the true distance of about 5. There were also a number of other datapoints with large relative error. In this case, they were probably outliers in the data due to the simulation which might not occur in actual data.

Since it was only for dimensionality 2, the underfit case, that recovery worsened dramatically given a perfect starting configuration, it is probably due to "the particular pattern of outliers", suggested by Spence and Lewandowsky (1989) for somewhat similar results, that made a perfect recovery unattainable for this replication in a given dimensionality. This assumes very different patterns of error were possible in different dimensionalities. It is to **MSCAL**'s credit that it managed to avoid local minima problems in the default case, if indeed outliers were present. Results by Spence and Lewandowsky (1989) showed that programs using the Young-Householder-Torgerson procedure as a starting configuration, as **MSCAL** does, were badly affected by outliers.

4.2 Recovery of Proximities

The results produced for the recovery of proximities were similar to those for coordinates recovery, although recovery magnitudes were smaller for proximities in all solution dimensions. Further differences were noted. When the data were underfit, recovery was not very good relative to the other two dimensions for any error distribution and fitting

Table 10

Proximities Recovery Analysis of Variance
for Emotions Data from MDS Workshop Members

Effect	df	MS	F-ratio	
Between-datasets-effects				
True error models	2	9.17	42.7	p<0.01
Predicted error models	2	2.30	10.7	p<0.01
True X predicted models	4	1.64	7.6	p<0.01
"Between" denominator	216	0.21		
Within-datasets-effects				
Dimensions	2	45.17	457.9	p<0.01*
Dims X true error models	4	1.20	12.2	p<0.01*
Dims X pred. err. models	4	0.05	0.5	p<0.75*
Dims X true X pred. err.	8	0.30	3.1	p<0.01*
"Within" denominator	432	0.10		
Simple main effects				
True err X pred. err models				
at dim=2	4	0.09	3.0	p<0.02
Denominator	216	0.03		
at dim=3	4	1.49	7.0	p<0.01
Denominator	216	0.22		
at dim=4	4	0.67	4.0	p<0.01
Denominator	216	0.17		

* using Greenhouse-Geisser conservative test.

criterion combination (see Figure 10b). Also, there was no significant improvement ($p>0.5$) when the appropriate model was used to fit any of the three data types. For the case of overfitting, the results parallel those of the corresponding coordinates recovery (see Figures 11a & b).

4.2.1 Mean-Square Error for 3-Dimensions

Figure 12b indicates that the amount of recovery of proximities was less than satisfactory, given that MSE ranged from approximately 21 to 24 for unnormalized proximities whose magnitudes were less than 2.0. The mean-square error for normal and squared normal fitting criteria were approximately 22 and 21, respectively, for the three data types. In all cases, 100% of the mean-square discrepancy was accounted for by squared bias, unlike the results for coordinates recovery which had little bias.

4.2.2 Loose Convergence

The number of solutions with recovery correlations less than 0.9 which were rerun using a convergence criterion of 0.0001 are given in Table 11. Stricter convergence criterion made little difference. The largest single increase in correlation for a given replication was 0.02.

4.2.3 Starting Configurations

Results were similar to those obtained for stimulus coordinates recovery using a perfect starting configuration for the corresponding worst case replication in dimensionality 3 and are shown in Table 6. Once more, the disastrous effect

Table 11

Number of MSCAL Proximities Recovery Correlations
Less Than 0.9 for Emotions Data
by Error/Fitting Criteria Condition and Dimensionality

Condition Error/Fitting Criteria	Dimensionality		
	2	3	4
Norm./Norm.	0	0	0
Norm./Lognorm.	1	0	0
Norm./Sq. Norm.	0	0	0
Lognorm./Norm.	2	1	1
Lognorm./Lognorm.	2	2	3
Lognorm./Sq. Norm.	3	0	0
Sq. Norm./Norm.	0	0	0
Sq. Norm./Lognorm.	1	0	1
Sq. Norm./Sq. Norm.	0	0	0

a perfect starting configuration had on the corresponding coordinates recovery for dimensionality 2 was magnified for proximities.

4.2.4 MSCAL Performance

Table 12 contains several overall measures of the program's performance. The squared correlations for coordinates recovery were good in all 3 dimensions; they were less so for proximities recovery in 2 dimensions and for lognormal error models in 4 dimensions. This agrees with previous results. The mean elapsed times were almost all less than 1 CPU second on the SUN-3/60 without the floating-point accelerator. The average ratios of elapsed time to number of iterations were 0.10, 0.16 and 0.21 for dimensionality of 2, 3, and 4 respectively. This ratio increased by about 0.05 as the dimensionality increased, corresponding to the increased number of datapoints to be estimated with each increase in dimensionality. It should be noted that a single iteration included all the steps associated in estimating the various parameters as defined in the documentation (Clarkson, 1988a).

Table 12

MSCAL Mean Performance on 27 datasets
across 25 replications
for Emotions Data from MDS Workshop Members

True error model	Fit model	Dim	Squared Correlations* Coordinates	Proximities	Elapsed Time** in secs.	No.+ Iter.
Normal	Normal	2	0.97	0.88	0.22	3.6
Normal	Lognormal	2	0.95	0.87	0.47	3.9
Normal	Sq.normal	2	0.98	0.87	0.06	3.1
Lognormal	Lognormal	2	0.95	0.85	0.46	5.2
Lognormal	Normal	2	0.97	0.87	0.41	3.3
Lognormal	Sq.normal	2	0.97	0.86	0.36	3.1
Sq.normal	Sq.normal	2	0.98	0.88	0.35	3.2
Sq.normal	Normal	2	0.97	0.89	0.42	3.4
Sq.normal	Lognormal	2	0.93	0.85	0.50	3.6
Normal	Normal	3	0.97	0.95	0.43	3.0
Normal	Lognormal	3	0.97	0.94	0.33	3.2
Normal	Sq.normal	3	0.97	0.95	0.35	2.5
Lognormal	Lognormal	3	0.95	0.92	0.43	3.6
Lognormal	Normal	3	0.95	0.91	0.68	3.0
Lognormal	Sq.normal	3	0.94	0.91	0.34	2.8
Sq.normal	Sq.normal	3	0.98	0.96	0.49	2.8
Sq.normal	Normal	3	0.98	0.97	0.23	2.8
Sq.normal	Lognormal	3	0.95	0.93	0.81	3.1
Normal	Normal	4	0.96	0.93	0.05	2.5
Normal	Lognormal	4	0.95	0.92	0.33	2.9
Normal	Sq.normal	4	0.97	0.93	0.15	2.3
Lognormal	Lognormal	4	0.92	0.89	1.37	3.2
Lognormal	Normal	4	0.94	0.89	0.61	2.8
Lognormal	Sq.normal	4	0.92	0.89	0.53	2.6
Sq.normal	Sq.normal	4	0.98	0.95	0.58	2.4
Sq.normal	Normal	4	0.97	0.95	0.63	2.6
Sq.normal	Lognormal	4	0.95	0.92	0.60	2.8

* These are mean R-Squared values for coordinate & proximities recovery.
**These are mean elapsed times of execution of the DMSCAL subroutine.
+ These are mean number of iterations.

Chapter 5

DISCUSSION

The interaction between the fitting criteria and error distributions in **MSCAL** have been explored. The implications of this study are confined to analyses with a least squares fit of ratio level data using a unweighted distance model. Also, the effects of different levels of error were not considered. Nevertheless, an attempt was made to study situations modelled on two applications of high quality and scope.

The fitting criteria provide improved fit for the appropriate error distributions when the correct number of solution dimensions is used in the analysis. Lognormal fitting of squared normal data and squared normal fitting of lognormal data gave poorer recovery, as expected. The assumptions of both models, with respect to where error is distributed most heavily, were at odds with the actual error in the data, as measured by the FCE values.

In the lower dimensionality case, a lognormal fitting criterion generally gives poorer recovery, while a squared normal fit produces better recovery, especially for the emotions data which had greater level of error. The squared normal fitting criteria was expected to do better when most of the error was in the smaller distances. This corresponds with the expectation that larger distances are decreased in lower dimensionality. When overfitting, normal and squared normal fitting criteria generally do better except for data with lognormal error.

Given that the error distribution and dimensionality of the data in practice are unknown, an assumption of normal error or rather, no transformation of dissimilarities and distances in the fitting criterion, is a cautious and safe initial step. Under this assumption, both coordinates and proximities estimates are quite satisfactory for the underfit and overfit dimension solutions, in addition to those for the true dimensionality.

The configuration estimates for the true dimensionality appear to be relatively unbiased. However, the opposite is seen for the proximities which tend to be highly biased. The algorithm gives close correspondence between configuration and distance recovery for true dimensionality and the overfit case, but proximities recovery is less satisfactory for lower dimensionality. This is explained by the associated loss of information in lower dimensionality which directly affects the distances. For those cases tested, a stricter convergence criterion seems to have little effect on improving the solutions.

Although direct comparison of single replication and multiple replication data analyses cannot be made since two different datasets were used, there do not appear to be any major differences in their performance. Results seem to parallel each other for the two sets of data. Although the replicated data seem to show more variability in the estimates, more error was present in the data which was magnified for data with lognormal error.

5.1 Further Research

As mentioned earlier there are many more options to be tested in **MSCAL**. Certainly non-metric methods need to compared with metric. Further research for the current models studied might include investigation of the role of the starting configuration in performance. In particular, random starts would indicate whether **MSCAL**'s performance is particular to certain configurations, number of dimensions, and/or error levels.

It is also of interest whether the choice of weighting of the criterion function would give different results, since different functions would be optimized. Specifically, allowing the stress function to be weighted by the stratum variance could have interesting consequences for the optimization procedure when error and fitting criteria do not match.

BIBLIOGRAPHY

Ball, A.D. (1982). *Individual differences multidimensional scaling and the nature of error in human judgments*. Unpublished Ph.D. dissertation, Ohio State University, .

Carroll, J.D., & Chang, J. (1972). Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckhart-Young" decomposition, *Psychometrika*, **35**, 283-319.

Clarkson, D. (1988a). *Multidimensional scaling*. [Draft of IMSL User's Manual Stat/Library]. Houston, Texas: IMSL,Inc.

Clarkson, D. (1988b). *MSCAL/DMSCAL*. [Draft of IMSL User's Manual Stat/Library, MSCAL/DMSCAL routine]. Houston, Texas: IMSL,Inc.

Cohen, H.S., & Jones, L.E. (1974). The effects of random error and subsampling of dimensions on recovery of configurations by non-metric multidimensional scaling, *Psychometrika*, **30**, 69-90.

Coxon, A.P.M. (1982). *The user's guide to multidimensional scaling: with special reference to the MDS(X) library of computer programs*. London: Heinneman Educational Books.

Davies, P.M., & Coxon, A.P.M. (1982). *Key texts in multidimensional scaling*. London: Heinneman Educational Books.

Davison, M.L. (1983). Introduction to multidimensional scaling and its applications, *Applied Psychological Measurement*, **7**, 373-379.

Girard, R.A., & Cliff, N. (1976). A Monte Carlo evaluation of interactive multidimensional scaling, *Psychometrika*, **41**, 43-64.

Graef, J., & Spence, I. (1979). Using distance information in the design of large multidimensional scaling experiments, *Psychological Bulletin*, **86**, 60-66.

IMSL, Inc. (1987). *IMSL STAT/LIBRARY User's manual for FORTRAN Subroutines for Statistical Analysis*, Version 1.0, Houston, Texas.

Isaac. P.D., & Poor, D.D.S. (1974). On the determination of appropriate dimensionality in data with error, *Psychometrika*, **39**, 91-109.

Kruskal, J.B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, *Psychometrika*, **29**, 1-27.

Kruskal, J.B. (1964b). Nonmetric multidimensional scaling: a numerical method, *Psychometrika*, **29**, 115-129.

Kruskal, J.B., Young, F.W., & Seery, J.B. (1978). *How to use KYST-2, a very flexible program to do multidimensional scaling and unfolding*, Murray Hill, NJ: Bell Laboratories.

Lingoes, J.C. (1965). An IBM 7090 program for Guttman-Lingoes smallest space analysis-1, *Behavioral Science*, **10**, 183-184.

MacCallum, R.C., & Cornelius, E.T. (1977). A Monte Carlo investigation of recovery of structure by ALSCAL, *Psychometrika*, **42**, 401-428.

Meulman J. (1986). *A distance approach to nonlinear multivariate analysis*, Leiden: DSWO Press.

Morgan, B.J.T., Woodhead, M.M., & Webster, J.C. (1976). On the recovery of physical dimensions of stimuli, using multidimensional scaling, *Journal of Acoustical Society of America*, **60**, 186-189.

Ramsay, J.O. (1969). Some statistical considerations in multidimensional scaling, *Psychometrika*, **34**, 167-182.

Ramsay, J.O. (1973). The effect of number of categories in rating scales on precision of estimation of scales values, *Psychometrika*, **38**, 513-532.

Ramsay, J.O. (1977). Maximum likelihood estimation in multidimensional scaling, *Psychometrika*, **42**, 241-266.

Ramsay, J.O. (1978). Confidence regions for multidimensional scaling analysis, *Psychometrika*, **43**, 145-160.

Ramsay, J.O. (1980a). Some small sample results for a maximum likelihood estimation in multidimensional scaling, *Psychometrika*, **45**, 139-144.

Ramsay, J.O. (1980b). The joint analysis of direct ratings, pairwise preferences, and dissimilarities, *Psychometrika*, **45**, 149-165.

Ramsay, J.O. (1982a). Some statistical approaches to multidimensional scaling data, *Journal of the Royal Statistical Society*, **145**, 285-312.

Ramsay, J.O. (1982b). *MULTISCALE II manual*. [Updated 1986]. Unpublished manuscript, McGill University, Montreal.

Ramsay, J.O. (1989). *MATFIT: A FORTRAN subroutine for comparing two matrices in a subspace*. Unpublished manuscript, McGill University, Montreal.

Schiffman, S.S., Reynolds, M.L. & Young, F.W. (1981). *Introduction to multidimensional scaling theory, methods, and applications*. New York: Academic Press.

Shepard, R.N. (1980). Multidimensional scaling, tree-fitting and clustering. In P.M. Davies & A.P.M. Coxon (Eds.), *Key texts in multidimensional scaling*. London: Heinneman Educational Books, 1982.

Sherman, C.R. (1972). Nonmetric multidimensional scaling: a Monte Carlo study of the basic parameters, *Psychometrika*, **37**, 323-355.

Spence, I. (1972). A Monte Carlo evaluation of three nonmetric multidimensional scaling algorithms, *Psychometrika*, **37**, 461-486.

Spence, I., & Young, F.W. (1978). Monte Carlo studies in nonmetric scaling, *Psychometrika*, **43**, 115-117.

Spence, I. (1979). A simple approximation for random rankings stress values, *Multivariate Behavioral Research*, **14**, 355-365.

Spence, I. (1983). Monte Carlo simulation studies, *Applied Psychological Measurement*, **7**, 405-425.

Spence, I., & Lewandowsky, S. (1989). Robust multidimensional scaling, *Psychometrika*, **54**, 501-513.

Takane, Y., Young, F.W., & DeLeeuw, J. (1977). Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features, *Psychometrika*, **42**, 7-67.

Takane, Y. (1981). Multidimensional successive categories scaling: a maximum likelihood method, *Psychometrika*, **46**, 9-28.

Torgerson, W.S. (1958). *Theory and Methods of Scaling*. New York: Wiley.

Torgerson, W.S. (1986). Scaling and PSYCHOMETRIKA: spatial and alternative representations of similarity data, *Psychometrika*, **51**, 57-63.

Wagenaar, W.A., & Padmos, P. (1971). Quantitative interpretation of stress in Kruskal's multidimensional scaling technique, *British Journal of Mathematical and Statistical Psychology*, **24**, 101-110.

Young, F.W. (1968). *TORSCA-9*, A FORTRAN IV program for nonmetric multidimensional scaling, *Behavioral Science*, **13**, 343-344.

Young, F.W. (1970). Nonmetric multidimensional scaling, recovery of metric information, *Psychometrika*, **35**, 455-473.

Young, F.W., & Null, C.H. (1978). Multidimensional scaling of nominal data: the recovery of metric information with ALSCAL, *Psychometrika*, **43**, 367-379.

Young, F.W., & Lewyckyj, R. (1981). *ALSCAL4 user's guide*, Chapel Hill, NC: Data Analysis and Theory Associates.

Young, G., & Householder, A.S. (1958). Discussion of a set of points in terms of their mutual distances, *Psychometrika*, **3**, 19-21.

APPENDIX A

Table 13

Population Configuration Matrix of Coxon Data for 32 Occupations

Stimulus	Dimension 1	Dimension 2	Dimension 3
1	-0.429	-0.046	0.201
2	-0.451	-0.160	-0.075
3	0.295	0.261	0.046
4	0.460	-0.198	0.053
5	-0.427	-0.108	0.170
6	-0.284	-0.148	-0.199
7	0.313	0.273	0.074
8	0.349	-0.083	-0.153
9	-0.245	-0.034	0.150
10	-0.432	-0.062	-0.119
11	0.276	0.271	0.078
12	-0.118	-0.085	-0.194
13	-0.092	-0.242	-0.035
14	-0.356	-0.193	-0.171
15	0.495	-0.127	0.129
16	0.034	-0.153	-0.238
17	-0.414	0.086	0.195
18	0.029	0.305	-0.225
19	0.521	-0.090	0.132
20	0.302	0.116	-0.018
21	-0.204	0.164	0.244
22	-0.207	0.318	-0.248
23	0.258	0.179	-0.068
24	0.421	-0.197	-0.059
25	-0.412	0.112	0.140
26	-0.271	-0.271	0.150
27	0.476	-0.076	0.137
28	0.439	-0.065	-0.079
29	-0.505	0.051	0.051
30	-0.247	0.196	-0.196
31	-0.041	0.160	0.132
32	0.468	-0.155	-0.006

Table 14

Population Configuration Matrix of Emotions Data

Stimulus	Dimension 1	Dimension 2	Dimension 3
1	-5.284	-0.016	-0.516
2	-2.303	2.005	-0.868
3	-0.885	1.583	2.304
4	-2.333	-1.720	1.449
5	-4.661	0.431	0.537
6	-1.985	-1.341	-0.611
7	-3.179	-0.930	-1.859
8	2.669	-3.040	0.536
9	2.267	2.330	1.656
10	2.850	2.320	-0.002
11	2.834	1.354	-2.559
12	3.405	-0.961	-1.722
13	2.617	-1.455	1.889
14	3.987	-0.559	-0.234

Table 15

Population Proximities Matrix
(Lower Triangle) of Coxon Data

2	0.30										
3	0.80	0.86									
4	0.91	0.92	0.49								
5	0.07	0.25	0.82	0.90							
6	0.44	0.21	0.75	0.79	0.40						
7	0.82	0.89	0.04	0.49	0.84	0.78					
8	0.86	0.81	0.40	0.26	0.84	0.64	0.42				
9	0.19	0.33	0.62	0.73	0.20	0.37	0.64	0.67			
10	0.32	0.11	0.81	0.92	0.29	0.19	0.84	0.78	0.33		
11	0.78	0.86	0.04	0.50	0.80	0.75	0.04	0.43	0.61	0.81	
12	0.51	0.36	0.59	0.64	0.48	0.18	0.62	0.47	0.37	0.32	0.60
13	0.46	0.37	0.64	0.56	0.42	0.27	0.66	0.48	0.32	0.39	0.64
	0.22										
14	0.41	0.14	0.82	0.85	0.36	0.09	0.85	0.71	0.37	0.16	0.82
	0.26	0.30									
15	0.93	0.97	0.44	0.11	0.92	0.85	0.44	0.32	0.75	0.96	0.46
	0.69	0.62	0.90								
16	0.65	0.51	0.57	0.52	0.62	0.32	0.60	0.33	0.49	0.49	0.58
	0.17	0.25	0.40	0.59							
17	0.13	0.37	0.75	0.93	0.20	0.48	0.76	0.86	0.21	0.35	0.72
	0.52	0.51	0.46	0.94	0.67						
18	0.72	0.68	0.38	0.72	0.73	0.55	0.41	0.51	0.58	0.60	0.39
	0.42	0.59	0.63	0.73	0.46	0.65					
19	0.95	1.00	0.43	0.15	0.95	0.87	0.42	0.33	0.77	0.99	0.44
	0.72	0.65	0.93	0.05	0.62	0.95	0.73				
..											

Table 15 (cont'd)

Population Proximities Matrix
(Lower Triangle) of Coxon Data

..

20	0.78	0.80	0.16	0.36	0.79	0.67	0.18	0.24	0.59	0.76	0.18
	0.50	0.53	0.74	0.34	0.44	0.75	0.39	0.34			
21	0.31	0.52	0.55	0.78	0.36	0.55	0.56	0.72	0.22	0.48	0.52
	0.51	0.51	0.57	0.77	0.62	0.23	0.54	0.78	0.57		
22	0.62	0.56	0.58	0.90	0.64	0.47	0.61	0.69	0.53	0.46	0.59
	0.42	0.61	0.54	0.91	0.53	0.54	0.24	0.92	0.59	0.52	
23	0.77	0.79	0.15	0.44	0.78	0.65	0.18	0.29	0.59	0.73	0.17
	0.48	0.55	0.72	0.43	0.44	0.73	0.30	0.43	0.09	0.56	0.52
24	0.90	0.87	0.49	0.12	0.88	0.72	0.50	0.16	0.72	0.87	0.51
	0.57	0.52	0.79	0.21	0.43	0.92	0.66	0.24	0.34	0.78	0.83
	0.41										
25	0.17	0.35	0.73	0.93	0.22	0.45	0.75	0.84	0.22	0.31	0.71
	0.49	0.51	0.44	0.94	0.64	0.06	0.60	0.96	0.73	0.24	0.48
	0.70	0.91									
26	0.28	0.31	0.78	0.74	0.23	0.37	0.80	0.72	0.24	0.38	0.77
	0.42	0.26	0.34	0.78	0.51	0.39	0.75	0.81	0.71	0.45	0.71
	0.73	0.73	0.41								
27	0.91	0.95	0.39	0.15	0.90	0.83	0.39	0.32	0.72	0.94	0.40
	0.68	0.62	0.89	0.06	0.59	0.91	0.69	0.05	0.30	0.73	0.88
	0.39	0.24	0.91	0.77							
28	0.91	0.90	0.38	0.19	0.90	0.74	0.39	0.12	0.72	0.87	0.41
	0.57	0.56	0.81	0.22	0.44	0.91	0.57	0.23	0.24	0.76	0.77
	0.30	0.13	0.90	0.77	0.22						
29	0.19	0.25	0.83	1.00	0.21	0.39	0.85	0.89	0.29	0.22	0.81
	0.48	0.51	0.36	1.02	0.64	0.17	0.65	1.04	0.81	0.37	0.50
	0.78	0.96	0.14	0.41	0.99	0.96					
30	0.50	0.43	0.60	0.85	0.51	0.35	0.63	0.66	0.42	0.33	0.60
	0.31	0.49	0.40	0.87	0.45	0.44	0.30	0.88	0.58	0.44	0.14
	0.52	0.79	0.38	0.58	0.84	0.74	0.38				
31	0.44	0.56	0.36	0.62	0.47	0.51	0.38	0.54	0.28	0.51	0.34
	0.42	0.44	0.56	0.61	0.49	0.39	0.39	0.62	0.38	0.20	0.44
	0.36	0.61	0.37	0.49	0.57	0.57	0.48	0.39			
32	0.93	0.92	0.45	0.07	0.91	0.78	0.46	0.20	0.74	0.91	0.47
	0.62	0.57	0.84	0.14	0.49	0.94	0.67	0.16	0.32	0.78	0.86
	0.40	0.08	0.93	0.76	0.16	0.12	1.00	0.82	0.61		

Table 16

Population Proximities Matrix
(Lower Triangle) of Emotions Data

2	3.62												
3	5.47	3.50											
4	3.93	4.39	3.71										
5	1.30	3.16	4.33	3.30									
6	3.56	3.37	4.27	2.12	3.41								
7	2.66	3.22	5.38	3.50	3.13	1.78							
8	8.57	7.22	6.09	5.25	8.11	5.08	6.66						
9	8.20	5.23	3.30	6.13	7.27	6.06	7.26	5.50					
10	8.48	5.23	4.45	6.73	7.76	6.09	7.10	5.39	1.76				
11	8.48	5.45	6.13	7.23	8.16	5.85	6.47	5.38	4.36	2.73			
12	8.82	6.49	6.41	6.60	8.49	5.52	6.59	3.16	4.85	3.75	2.53		
13	8.38	6.62	4.65	4.98	7.64	5.24	6.92	2.08	3.81	4.23	5.27	3.73	
14	9.29	6.82	5.90	6.64	8.74	6.03	7.36	2.91	3.86	3.10	3.22	1.65	2.68