

ePub^{WU} Institutional Repository

Patrick Mair and Ingwer Borg and Thomas Rusch

Goodness-of-Fit Assessment in Multidimensional Scaling and Unfolding

Article (Accepted for Publication)
(Refereed)

Original Citation:

Mair, Patrick and Borg, Ingwer and Rusch, Thomas (2016) Goodness-of-Fit Assessment in Multidimensional Scaling and Unfolding. *Multivariate Behavioral Research*, 51 (6). pp. 772-789. ISSN 1532-7906

This version is available at: <http://epub.wu.ac.at/5354/>

Available in ePub^{WU}: December 2016

ePub^{WU}, the institutional repository of the WU Vienna University of Economics and Business, is provided by the University Library and the IT-Services. The aim is to enable open access to the scholarly output of the WU.

This document is the version accepted for publication and — in case of peer review — incorporates referee comments.

Goodness-of-Fit Assessment in Multidimensional Scaling and Unfolding

Patrick Mair

Harvard University

Ingwer Borg

University of Münster

Thomas Rusch

WU Vienna University of Economics and Business

Abstract

Judging goodness-of fit in multidimensional scaling requires a comprehensive set of diagnostic tools instead of relying on stress rules of thumb. This article elaborates on corresponding strategies and gives practical guidelines for researchers to obtain a clear picture of the goodness-of-fit of a solution. Special emphasis will be placed on the use of permutation tests. The second part of the article focuses on goodness-of-fit assessment of an important variant of multidimensional scaling called unfolding, which can be applied to a broad range of psychological data settings. Two real-life datasets are presented in order to walk the reader through the entire set of diagnostic measures, tests, and plots. R code is provided as supplementary information which makes the whole goodness-of-fit assessment workflow, as presented in this article, fully reproducible.

Goodness-of-Fit Assessment in Multidimensional Scaling and Unfolding

Introduction

Multidimensional scaling (MDS; Torgerson, 1952; Kruskal, 1964a, 1964b; Cox & Cox, 2001; Borg & Groenen, 2005) is a technique that represents proximities among objects as distances among points in a low-dimensional space (with given dimensionality). It allows researchers to explore similarity structures among objects in a multivariate dataset.

The standard way to apply MDS is of an exploratory nature: The empirically obtained measures of proximity amongst some objects of interest are mapped as distances amongst corresponding points in a multidimensional geometric space such that a target function called *stress* is minimized. The smaller the stress, the better the fit. As in many other exploratory techniques, the number of dimensions needs to be fixed a priori. In MDS, the larger the number of dimensions, the smaller the stress. The number of dimensions is an important, albeit not exclusive determinant for MDS goodness-of-fit assessment. There are other factors such as replicability, stability, theoretical meaningfulness, and interpretation which need to be taken into account. Along these lines it is urgent to emphasize that researchers should not base their fit judgement solely on stress rules of thumb (e.g., stress should be lower than 0.20). Goodness-of-fit assessment in MDS requires in-depth examination. In the following sections we elaborate on corresponding strategies.

The first part of the article focuses on standard MDS models. We discuss the limited usefulness of common random stress norms and elaborate on permutation tests, which provide sharper stress norms and allow for significance testing. Graphical goodness-of-fit displays such as Shepard diagrams and bubble plots are discussed. Resampling approaches are presented in order to examine the stability of a solution, as well as stress-per-point strategies to determine influential outliers responsible for misfit. In addition, we discuss aspects of interpretability of a particular configuration as an important criterion to achieve a final solution, and put this into the context of confirmatory MDS models where researchers have an underlying theory regarding the configurations.

In the second part of the article we focus on unfolding, an MDS technique that uses persons-by-objects dissimilarity matrices as input. All goodness-of-fit tools considered for MDS will be discussed in relation to their applicability in unfolding models.

Using two real-life examples, we give good practice guidelines in terms of goodness-of-fit assessment for MDS and unfolding solutions. All the computations in this article are performed in R (R Development Core Team, 2016) in conjunction with the **smacof** package (De Leeuw & Mair, 2009). The supplementary materials contain an R code file which makes all analyses presented in the article fully reproducible.

Multidimensional Scaling: Basics

MDS takes a symmetric proximity matrix as input. Proximity is an umbrella term for both *similarities* and *dissimilarities*; it simply refers to the closeness of objects in a certain domain. From now on, for simplicity, all further elaborations are based on dissimilarities¹.

Let Δ be a symmetric $n \times n$ input dissimilarity matrix with elements δ_{ij} ($i, j = 1, \dots, n$). The problem we solve in MDS is to locate n points in a low-dimensional Euclidean space of dimension p such that their distances are as close as possible to the given input dissimilarities δ_{ij} (or, actually, some transformation of it, as described below). In other words, we aim to find an $n \times p$ matrix \mathbf{X} of the points' coordinates where the distance between any points i and j is

$$d_{ij}(\mathbf{X}) = \sqrt{\sum_{s=1}^p (x_{is} - x_{js})^2}. \quad (1)$$

We typically aim for a low p (i.e., much smaller than n). If $p = 2$ or $p = 3$, we can easily plot the solution.

Several variants of MDS with different target criteria have been proposed in the literature. In this article we focus on the state-of-the-art approach called SMACOF (Stress MAjorization of a COmplicated Function) which goes back to De Leeuw (1977). It solves

¹Note that there is no restriction in using dissimilarities only since similarities can be transformed into dissimilarities by a proper transformation function (for a detailed discussion see Gower & Legendre, 1986).

the MDS problem numerically by the majorization approach and, compared to other MDS variants such as classical scaling (Torgerson, 1952), it offers a tremendous amount of flexibility. As we will see, this flexibility is highly valuable in relation to goodness-of-fit assessment. SMACOF uses Kruskal's *stress* as target criterion:

$$\sigma(\mathbf{X}) = \sum_{i < j} w_{ij} (\hat{d}_{ij} - d_{ij}(\mathbf{X}))^2, \quad (2)$$

with the constraint $\sum_{i < j} w_{ij} \hat{d}_{ij}^2 = n(n-1)/2$.

The w_{ij} 's are optional a priori weights for the dissimilarities and are especially useful in cases where some of the δ_{ij} 's are missing (i.e., $w_{ij} = 0$ if δ_{ij} is missing, and $w_{ij} = 1$ otherwise). The \hat{d}_{ij} are called *disparities* or, simply, *d-hats*. Disparities are optimally scaled dissimilarities. That is, they are subject to transformations that are admissible on the assumed scale level of the δ_{ij} 's and, at the same time, minimize Eq. (2) for any given \mathbf{X} .

The most popular optimal scaling transformation in MDS applications is a monotone regression transformation of the form $\hat{d}_{ij} = f(\delta_{ij})$, where f is a monotone step function. This version is called *ordinal* (or nonmetric) MDS (Kruskal, 1964b). It takes dissimilarities on an ordinal scale so that transformations that preserve the order of the data do not change the interpretable information of the data. In terms of linear metric MDS strategies, *ratio* MDS uses $\hat{d}_{ij} = b\delta_{ij}$ (with $b > 0$) as transformation, whereas *interval* MDS uses $\hat{d}_{ij} = a + b\delta_{ij}$. In terms of a nonlinear metric transformation we can also think of a monotone spline transformation $\hat{d}_{ij} = f(\delta_{ij})$ where f is an *I-spline* (integrated spline) transformation (Ramsay, 1988), or power transformations $\hat{d}_{ij} = \delta_{ij}^\lambda$ (Mair, Rusch, & Hornik, 2014; Rusch, Mair, & Hornik, 2015).

MDS software typically prints out a standardized version of the raw stress, called *stress-1*, such that the resulting stress value does not depend on the absolute magnitude of the dissimilarities. Using the normalization above, the stress-1 is

$$\sigma_1(\mathbf{X}) = \sqrt{\frac{\sum_{i < j} w_{ij} (\hat{d}_{ij} - d_{ij}(\mathbf{X}))^2}{\sum_{i < j} \hat{d}_{ij}^2}}. \quad (3)$$

As with other exploratory techniques, we must decide how many dimensions p we need in order to achieve an acceptable fit. We also need to clarify what “acceptable” means within an MDS fit context. In addition, since a broad variety of MDS models can be specified within the SMACOF framework, we might also ask which type of MDS model fits our data well (e.g., ordinal vs. interval model, exploratory vs. confirmatory model).

Over the years, several MDS extensions have been proposed. The most notable extensions are models for scaling individual differences and unfolding models. Models for individual differences (e.g. INDSCAL, IDIOSCAL; Carroll & Chang, 1970) take K dissimilarity matrices $\Delta_1, \dots, \Delta_K$ as input. The *stress* formulation in Eq. (2) uses an additional index k and involves individual configuration matrices $\mathbf{X}_1, \dots, \mathbf{X}_K$, correspondingly. The base of these matrices is a *group space*. Each \mathbf{X}_k is generated from this group space by individually weighting a common set of dimensions of the group space (INDSCAL), or an individually chosen set of dimensions of the group space (IDIOSCAL). Different individual weight specifications lead to different (more or less restrictive) MDS models (see, e.g., De Leeuw & Mair, 2009).

A second important MDS development is unfolding models (Coombs, 1964) and corresponding variants and extensions. Unfolding models can be regarded as a special case of MDS and will play an important role in the second part of this paper. For the moment, the main difference is that the input dissimilarity matrix Δ is not square. Additionally, its elements are usually preference scores (such as rank-orders of preference) of different individuals for a set of choice objects (Borg & Groenen, 2005). A stress expression as in Eq. (2) can be established in an analogous way. Before we consider unfolding, let us focus first on standard MDS.

Goodness-of-Fit Strategies in MDS

Multiple Random Starts

Before we discuss the stress magnitude in relation to MDS goodness-of-fit, the issue of *local minima* needs to be addressed. The stress target function in Eq. (2) is typically very bumpy. It can easily happen that the algorithm ends up in a local minimum and not in the global minimum (see De Leeuw, 1984, 1988; Groenen, 1993; Vera, Heiser, & Murillo, 2007). The final solution is affected by the starting solution for the configurations.

By default, most MDS programs use a classical scaling solution as starting configuration. Classical scaling (Torgerson, 1952) is an old MDS technique which tackles the problem analytically, however it is way less flexible than numerical MDS approaches. Classical scaling often provides a good initial guess. However, the user should always check the fit based on multiple random starts or provide a starting solution based on theoretical considerations (e.g., based on an underlying psychological theory). Subsequently, one can either pick the MDS solution with the lowest stress value or a solution which has an acceptably low stress value but may be more attractive from an interpretive point of view. Examples of the effects of choosing different initial configurations on the MDS solutions and their interpretability are given, with both artificial and real-life data, are given in Borg and Mair (2016). More detailed discussions on this topic follow in the section on interpretability.

Scree Plot, Stress Rules of Thumb, and Stress Norms

The larger the number of dimensions p , the lower the stress (as long as the MDS algorithm does not get trapped in suboptimal local minima). Similar to principal components analysis (PCA) or exploratory factor analysis (EFA), a *scree plot* can be produced with the number of dimensions on the x-axis and the respective stress values on the y-axis. Subsequently, the elbow criterion (i.e., drop all further dimensions/solutions after the one starting with the elbow) can be applied to get an initial picture of conceivable

solutions.

Regarding stress rules of thumb, Kruskal (1964a, p. 3) stated about stress-1: “Our experience with experimental and synthetic data suggests the following verbal evaluation: 20% poor, 10% fair, 5% good, 2.5% excellent, 0% perfect.” There are several problems with applying such rules of thumb. First of all, it is easy to see from Eq. (2) that for an increasing n , the stress naturally becomes larger since the number of terms of the sum increases (p fixed). This issue is important in modern MDS applications where n can be very large. In addition, the upper bound for Kruskal’s stress (i.e., the largest stress value we can realistically get) is not trivial to obtain (see De Leeuw & Stoop, 1984). This fact also implies that reporting stress values on a percentage scale, as often seen in publications, is not feasible.

The fact that stress depends on n was recognized early and studied in various Monte Carlo simulations in the 1970’s (for an overview, see Spence & Young, 1978). Stress norms from random arrangements were derived by Stenson and Knoll (1969) and Spence and Ogilvie (1973). Such stress norms represent a “bad fit” benchmark; the observed stress value should clearly be smaller.

The norms established in these early simulations were limited to ordinal MDS only. Within an environment like R it is easy to derive random stress norms for arbitrary MDS models. We illustrate the stress properties for varying n and p in relation to three MDS settings: ratio MDS, interval MDS, and ordinal MDS. We draw the dissimilarities from a random uniform $U(0, 1)$ distribution for varying n and p . For each scenario we perform 500 replications, fit the MDS model using the **smacof** package, and report the average stress. The underlying R code is provided in the supplementary materials and allows readers to perform their own simulations.

[INSERT FIGURE 1 HERE]

Figure 1 shows the resulting stress norms for varying p across the panels. We see how

the stress norms become larger for increasing n and that different MDS models lead to different stress norms. In addition, one can compute the standard deviation sd of the random stress distribution. An observed stress value is considered “significant” in the literature if it is smaller than the lower $2 \times sd$ boundary. Note that, as Spence and Ogilvie (1973) observed, the dispersion of the random stress norms is in general very small.

Even though such simulation studies certainly give one a better insight into the stress behavior depending on various model circumstances, at the end of the day they are only of limited use. This is because we create fully random configurations under the null hypothesis, what Cliff (1973) calls the “the nullest of all null hypotheses”, which almost always leads to a significant result in real life applications. This is due to the fact that empirical data typically have some sort of distance-like structure and the resulting stress is nowhere near the random stress norm. Thus, rejecting the random null is almost trivial since what is actually tested is whether the data can be represented “significantly” better than random data.

The evaluation of stress norms is, however, more complex. Apart from n and p , and the underlying type of MDS, the following circumstances affect stress (Borg & Groenen, 2005): the more error in the data, the higher the stress; the more ties in an ordinal MDS, the lower the stress; the more missing values, the lower the stress. Therefore, random stress norms are too “soft” for modern MDS settings. We will elaborate on more flexible, “sharper” permutation tests below.

Shepard Diagram

The Shepard diagram (Shepard, 1962; De Leeuw & Mair, 2015) is a scatterplot with the dissimilarities δ_{ij} on the x-axis and the fitted MDS distances $d_{ij}(\mathbf{X})$ on the y-axis. It adds the transformation function that shows how dissimilarities and disparities \hat{d}_{ij} are related to each other. Hence, the Shepard diagram gives detailed insight into the dissimilarity transformation, which can be subject to further substantive investigation. For

instance, if a clear monotone increasing pattern is detected, one can try to parameterize the pattern through a specific function. This function could be a line, a polynomial, a spline, an exponential function, or any other function which reflects the relationship between observed dissimilarities and MDS distances.

The goodness-of-fit of the transformation function can be characterized by a residual sum-of-squares (RSS), just as in regression. In MDS this corresponds to the raw stress as given in Eq. (2). Finally, to some degree, Shepard diagrams help to detect outliers; a more sophisticated approach is called *stress-per-point*.

Stress-per-Point

In the Shepard diagram we look at each residual individually. A variant of this technique is to look at all residuals of one object in relation to all the remaining ones. From Eq. (2) we see that the stress is an additive decomposition in terms of a RSS over all pairs of objects. For a given object i , the residuals related to each of the remaining $n - 1$ objects can be extracted and averaged.

This concept is called *stress-per-point* (SPP). Expressed on a percentage scale, it tells us how much each object contributes to the misfit. Points with high SPP should be subject to special consideration. A high SPP can provide evidence of mistakes in data collection, or it can be that points with a high SPP relate to objects with certain special properties, different from the remaining objects. From an interpretation point of view one can give less emphasis to high SPP objects compared to the ones with low SPP. If the data setting permits, points with high SPP can be eliminated and the MDS model re-fitted (which leads to a fit improvement).

In terms of plots, a simple diagram with the objects on the x-axis and the SPP on the y-axis can be created. Bubble plots are an option to incorporate the SPP information into an MDS configuration plot: The larger the bubble of a point in the configuration plot, the larger the SPP of this particular object.

Confidence Regions, Stability, and Replicability

Sometimes it is interesting to look at confidence regions for the points in the MDS space. There are two basic approaches to get confidence regions. The first one is a parametric maximum likelihood approach to MDS (MULTISCALE; Ramsay, 1977) where one needs to impose a specific structure on the residuals (i.e., independently log-normally distributed). These assumptions may be too restrictive in practice, but if they hold, confidence regions can be derived easily.

The second approach is to apply resampling strategies to the dissimilarities, such as jackknife (De Leeuw & Meulman, 1986) and bootstrap (Heiser & Meulman, 1983; Jacoby & Armstrong, 2014). For each jackknife or bootstrap sample an MDS solution is fitted, subject to a subsequent Procrustes transformation. Procrustes removes irrelevant differences between two configurations by appropriate rotations, dilations (i.e., enlargements or reductions of the entire configuration), and translations (i.e., a displacement of the entire configuration relative to a fixed point). As a result, multiple MDS solutions are comparable and confidence regions for the points can be obtained (and plotted).

Jackknife and bootstrap create artificial data replications. For systematic theory construction, researchers need to collect multiple datasets or have access to data from previous studies. Such replications are desirable in MDS. First of all, it is tempting to overinterpret a particular solution based on a single dataset. This goes hand in hand with the statistical concept of “overfitting”. For example, if there is an underlying theory about possible partitions of the MDS space such as in *facet theory* (see Borg & Shye, 1995), replications allow one to study whether the MDS space can be partitioned in the same way across multiple datasets. The stability of the partitions can substantiate various “laws of formation” of the MDS configuration and it can be of relevance to explore whether these laws are still valid if additional items are included.

Another strategy to explore replicability is to split the data according to external

variables such as gender or age, fit MDS models for each subsample, perform a Procrustes transformation, and check how similar the solutions are (e.g., by using a congruence coefficient). This strategy can provide important substantive information for possible differences across subgroups. Conversely, if the MDS results for various data subsets are, in fact, quite similar, the MDS of the full sample can be considered as stable over subgroups.

Yet another procedure one can apply is to explore what happens to the stress if some of the points in the configuration are moved around slightly. This can be achieved by fitting an MDS model, extracting the coordinates, manipulating a few points of interest (e.g., points that do not obey a theory-based interpretation are moved according to the underlying theory), and then computing the stress of this modified configuration. This can be done using the modified configuration as the starting solution and setting the number of iterations equal to 0. By comparing the new stress with the original stress the effect of the slight point shifts on the goodness-of-fit can be examined. This allows one to determine whether a “microscopic” interpretation of the points is feasible or not.

Note that the approaches presented in this section do not tackle the global fit of an MDS solution directly. Rather, they add some additional flavor to the overall goodness-of-fit assessment by exploring the stability of a solution.

Degenerate Solutions, Interpretability, and Restricted Solutions

In some rare applications of ordinal MDS it can be the case that the stress is extremely small but the MDS distances do not explain the data properly. Such “degenerate” solutions can be identified by looking at the Shepard diagram for gross steps, by looking at MDS configuration for dense clusters, and by comparing an interval (or ratio) MDS fit with an ordinal MDS fit.

[INSERT FIGURE 2 HERE]

Let us illustrate a degenerate solution using a simple dataset from Lawler (1967). It

consists of a 6×6 matrix reflecting dissimilarities between trait-method combinations related to manager performance. First, we fit a two-dimensional, ordinal solution which leads to a stress-1 value of 0.004. From the configuration plot and the Shepard diagram (top panels of Figure 2) we see that the input dissimilarities have basically become dichotomized. Therefore, the monotonic regression is not a good choice as a transformation function if we want to assume that the data contain at least traces of valid quantitative information beyond ranks. The linear transformation, as used in interval MDS, does a much better job even though the stress is getting larger (0.137). This transformation leads to a more accurate structural representation of our data (bottom panel of Figure 2), assuming that the data are at least approximately on an interval scale.

This argument also implies that interpretability can matter in terms of evaluating MDS fit: one might not take the solution with the lowest stress value but rather pick a model with weaker dissimilarity-distance fit yet a more satisfactory fit of the MDS structure to substantive meaning.

Related to interpretability is the fact that sometimes there exists an underlying psychological theory on the configurations. The SMACOF framework allows one to fit unrestricted (exploratory) MDS models as well as restricted (confirmatory) MDS models. MDS models can be restricted in two basic ways: First, we can impose restrictions on the configurations by means of geometric shapes such as a circle in two dimensions (see Cox & Cox, 1991). Second, we can put external restrictions on the configurations based on an underlying theory that makes the MDS model confirmatory (De Leeuw & Heiser, 1980; Borg & Lingoes, 1980).

In either case, the stress values of the unrestricted MDS and the restricted MDS fit can be compared. If the unrestricted stress is not clearly lower than the restricted stress, one may pick the restricted MDS solution as the model of choice. We give an example of such a model comparison in the section on unfolding.

Permutation Tests for MDS

As mentioned above, using random data to obtain a stress distribution under a null hypothesis has severe limitations. We are drawing dissimilarities from a population over which we assume a distribution that does not necessarily reflect the data generating process under question².

A better way to create a stress sampling distribution under the H_0 of misfit is by using permutation tests. Permutation tests are widely used in modern statistics for many kinds of different testing scenarios where some distributional assumptions are not fulfilled, particularly in small samples. The basic idea of any permutation test is to find a resampling scenario that generates samples under the null hypothesis such that we get a sampling distribution that is not tied to strong distributional assumptions of the population (see Welch, 1990). Under some circumstances the permutation scenario is quite simple (e.g., in t -tests, ANOVA, etc.) whereas in other applications such as multivariate data settings the resampling strategy can be trickier.

MDS is such a multivariate technique where we have to carefully consider various permutation possibilities. As far as permutation strategies are concerned, we distinguish between directly observed proximities vs. derived proximities (more details follow below). Our null and alternative hypotheses can be formulated as follows:

- H_0 : Stress/configuration are obtained from a random permutation of dissimilarities.
- H_1 : Stress/configuration are obtained from a something other than a random permutation of dissimilarities.

Thus, we can think of H_0 as a hypothesis involving little departure from complete exchangeability of dissimilarities in the data generating process (or small signal), and H_1 implying at least some process of non-exchangeability (or some signal) in the data due to

²The same issues apply to some methods for choosing the number of clusters in cluster analysis such as the gap statistic proposed by Tibshirani, Walther, and Hastie (2001).

at least one dissimilarity. Special cases under H_0 are random dissimilarities and constant dissimilarities.

Directly Observed Dissimilarities

Let us first consider the case where the dissimilarities are directly observed and collected in a symmetric $n \times n$ dissimilarity matrix $\mathbf{\Delta}$. We can assume the rather uninformative H_0 from above (which still is more informative than the null under random dissimilarities) where there exists exchangeability across dissimilarities and, thus, we can randomly permute the lower-triangular portion of $\mathbf{\Delta}$ (and preserve its symmetry). Formally, we get a new $\mathbf{\Delta}^{(t)}$ for each of the $t = 1, \dots, T$ replications. Each $\mathbf{\Delta}^{(t)}$ is subject to an MDS fit of the same type as the original MDS on the observed $\mathbf{\Delta}$. Thus, for each replication we get an $n \times p$ configuration matrix $\mathbf{X}^{(t)}$, and, correspondingly, a stress value $\sigma(\mathbf{X}^{(t)})$. After T replications we can establish a sampling distribution for the stress under H_0 . The expected value of this distribution gives a stress norm $\bar{\sigma}$ which is typically smaller than the classical random stress norms and takes into account basic properties of our data (e.g. ties, missing values, etc.). Therefore, it provides a “sharper” and more realistic H_0 than the one under fully random dissimilarities.

In terms of inference we can apply a one-sided testing strategy by cutting off the left 5% of the sampling distribution which gives us the critical value c . If our observed stress $\sigma(\mathbf{X}) \leq c$ (or we just look at the p -value), we reject the H_0 . This provides some evidence that the MDS captures a structural signal.

As a simple example, let us compute a ratio MDS and an interval MDS with $p = 2$ using the Lawler dataset. We perform $T = 500$ replications. For the ratio MDS fit we obtain an observed stress value of 0.241, a random stress norm of 0.312, a permutation stress norm of 0.246, and a p -value of 0.334. We see that the observed stress value is much smaller than the random stress norm but is very close to the permutation stress norm. In fact, the test result suggests that the model does not fit substantially better than the one

expected under the H_0 of randomly permuted dissimilarities.

By fitting a more flexible interval MDS we get a stress of 0.137, a permutation stress norm of 0.215, and a p -value of 0. Therefore, we reject the H_0 that the stress and, subsequently, the configuration are obtained from a random permutation of dissimilarities. There is some structural signal in the resulting interval MDS configuration.

Derived Dissimilarities

As a second scenario we focus on the more common case where the dissimilarities are derived from an underlying $N \times n$ matrix \mathbf{Y} with $\mathbf{\Delta} = g(\mathbf{Y})$. \mathbf{Y} is now subject to permutations. If we want to scale the variables, the most natural way of permuting is within columns Y_1, Y_2, \dots, Y_n . This implies the assumption that the persons are stochastically independent from each other (i.e. interchangeability of the objects), but the variables are not (in fact, variables could be even on different scale levels).

At this point we can think of keeping Y_1 fixed and permute the remaining $n - 1$ variables according to $\boldsymbol{\pi}_j = (\pi_j^{(1)}, \pi_j^{(2)}, \dots, \pi_j^{(N)})$ where $\boldsymbol{\pi}_j$ represents a random permutation of the N row indices for each variable $j = 2, \dots, n$. Note that each variable is subject to a different permutation. This setup, as used in Buja and Eyuboglu (1992) within a PCA context, would lead to a permuted data matrix

$$\mathbf{Y}_\pi^{(t)} = \begin{pmatrix} Y_{1,1} & Y_{\pi_2(1),2} & Y_{\pi_3(1),3} & \cdots & Y_{\pi_n(1),n} \\ Y_{2,1} & Y_{\pi_2(2),2} & Y_{\pi_3(2),3} & \cdots & Y_{\pi_n(2),n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Y_{N,1} & Y_{\pi_2(N),2} & Y_{\pi_3(N),3} & \cdots & Y_{\pi_n(N),n} \end{pmatrix}$$

For each replication t we apply $\mathbf{\Delta}^{(t)} = g(\mathbf{Y}_\pi^{(t)})$ and fit the MDS model which gives us $\sigma(\mathbf{X}^{(t)})$. Let us denote the number of replications for this particular setting by $T^* := T$.

The problem with this permutation strategy is that it destroys the basic dissimilarity structure. Following arguments described in Linting (2007), we can permute more systematically. We pick one variable Y_i at a time subject to permutation. The remaining

set of variables remain unpermuted. We then apply the dissimilarity computations $g(\cdot)$ that must involve this particular variable. For instance, using the Euclidean distance as $g(\cdot)$ and focusing on the first variable (i.e., $i = 1$), the data representation above changes to

$$\mathbf{Y}_\pi^{(t,1)} = \begin{pmatrix} Y_{\pi_1(1),1} & Y_{1,2} & Y_{1,3} & \cdots & Y_{1,n} \\ Y_{\pi_1(2),1} & Y_{2,2} & Y_{2,3} & \cdots & Y_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Y_{\pi_1(N),1} & Y_{N,2} & Y_{N,3} & \cdots & Y_{N,n} \end{pmatrix}$$

The Euclidean distances related to the first variable are

$$\delta_{1j} = \sqrt{\sum_{s=1}^N (y_{1s} - y_{js}^{(t,1)})^2}, \quad (4)$$

with $j > 1$ (up to n). This gives a vector of dissimilarities $\boldsymbol{\delta}_1$ which fills up the first column of the dissimilarity matrix below the diagonal element $(1, 1)$. In the second run we permute variable Y_2 only and keep the remaining variables constant. Eq. (4) changes to δ_{2j} with $j > 2$ and gives the second column in $\boldsymbol{\Delta}^{(t)}$ below diagonal element $(2, 2)$. More generally, for fixed i we compute

$$\delta_{ij} = \sqrt{\sum_{s=1}^N (y_{1s} - y_{js}^{(t,i)})^2} \quad \forall j > i, \quad (5)$$

which gives $\boldsymbol{\delta}_i$ of length $n - i$ and fills up the i -th column in $\boldsymbol{\Delta}^{(t)}$ below the diagonal element (i, i) , accordingly. The advantage of this strategy is that it breaks up the proximity structure between the variable i and each of the remaining variables, while keeping the structure among the remaining variables intact.

Note that in order to establish a single permuted $\boldsymbol{\Delta}^{(t)}$, n permutations have to be carried out. Thus, the total number of permutations is $T = n \times T^*$. Analogous to the observed dissimilarity case, we get a stress sampling distribution under H_0 (based on $\sigma(\mathbf{X}^{(t)})$), can compute a permutation stress norm $\bar{\sigma}$, and perform inference. H_0 and H_1 can be formulated in an equivalent manner as above; the main difference is that this version of the test is generally sharper since the null is a bit more informative.

Odds and Ends

Within a PCA context, Buja and Eyuboglu (1992) and Skinner (1979) found little difference between random data and permuted data in terms of test performance. Linting (2007) attributes this to the permutation scheme they used and achieves some success using a permutation scheme where only one variable is permuted at a time. Similar issues apply to MDS: The systematic permutation scheme gives a more informative null distribution (in the sense of producing a sharper test) than the one based on full permutations. Still, the user should not overemphasize the p -value and judge the fit solely on the basis of a permutation test.

The permutation strategies proposed deal with rather uninformative null hypotheses. This may not be desirable. If there is more insight into possible dependency structures to derive more informative null hypotheses (e.g., based on the experimental design) the permutation scheme can be adapted correspondingly.

Overall, we strongly suggest that, if possible, one permutes the data matrix directly rather than the dissimilarity matrix. Apart from dependency issues—which are solved by permuting the raw data matrix—this strategy generally provides a sharper test. Still, some further methodological research remains to be done; a corresponding outlook can be found in the discussion section.

Example I: PTSD Wenchuan Earthquake

In this section we present a fully worked out example of (good practice) goodness-of-fit exploration in MDS. The dataset we use is from clinical psychology. McNally et al. (2015) collected data related to PTSD (posttraumatic stress disorder) symptoms reported by survivors of the Wenchuan earthquake in China. They used the PTSD checklist-civilian (PCL-C; Weathers, Litz, Herman, Huska, & Keane, 1993) which consists of 17 PTSD symptom items scaled on a five-point Likert scale (1 ... “not at all”; 5 ... “extremely”). The item labels and corresponding short variable names are given in

Table 1.

[INSERT TABLE 1 HERE]

We are interested in scaling the PTSD symptoms and exploring associations amongst them. First, we compute the dissimilarities using the Euclidean distance, which results in a 17×17 symmetric dissimilarity matrix. Subsequently, we fit a two-dimensional, ordinal MDS which produces a stress-1 value of 0.133. According to Kruskal's verbalization this fit would be "fair".

Next we produce a scree plot with varying numbers of dimensions from $p = 1$ to a maximum of $p = 16$. Figure 3 shows the results. Researchers could make an argument here for three, four, or even five dimensions (stress-1 values: 0.064, 0.042, 0.028). As opposed to EFA, where a large number of dimensions is no obstacle as long as the factors can be somehow interpreted, in MDS we tend to keep the number of dimensions as low as possible (typically $p = 2$ or $p = 3$) so that a configuration plot can be produced. The price we pay is a (slightly) worse fit. In this example we could pick a 3D solution; there is not that much decrease in stress for higher-dimensional solutions and, at the same time, the 3D stress is clearly smaller than the 2D stress. Still, we have to be careful with such purely descriptive arguments. It could be that the stress increase is caused by a few points (cf. SPP), or that the basic theoretical symptom structure is already nicely reflected by the 2D solution. We will explore these issues further below.

[INSERT FIGURE 3 HERE]

Let us just stick to the 2D solution for the moment and check whether the classical scaling initial configuration actually produced the lowest possible stress value, or whether a particular random start would give us a lower local minimum. At the same time we investigate whether structurally different solutions exist with almost the same stress. We

fit 100 additional ordinal MDS models based on 100 random starts with the result that for a few random starts we get slightly lower stress-1 values (with 0.13279 as the lowest one) than those from our initial fit (0.13281), with similar configuration patterns. Let us continue with the random start solution that gave the lowest stress value.

Next, we produce various Shepard plots. Remember that the original items use five-point Likert responses. An ordinal MDS is one attractive option. However, we can also claim that there is some metric information in the data and consider the data (and dissimilarities) on a metric scale level. Thus, in addition to the ordinal MDS, we fit a 2D interval MDS, and a 2D spline MDS (spline degree 2 and 2 knots). Eventually, we can make a data-driven decision on whether to treat the dissimilarities on an ordinal or metric level by picking a proper MDS model.

[INSERT FIGURE 4 HERE]

The MDS fits provide an interesting picture. We see that the interval MDS leads to a relatively high stress (0.23). However, the extreme point on the lower end of the abscissas of the Shepard diagrams represents an outlier that influences the transformation function. If we were to eliminate this point from the calculation, the stress for the interval MDS solution would reduce to 0.175. This is an admissible strategy if we aim for a simple transformation function.

The spline transformation captures the dissimilarity patterns nicely, suggests that there is a structural break in the dissimilarity transformation, and leads to a stress value of 0.155. This value is almost as low as the one based on ordinal MDS. We could now further tune the spline model (e.g., more knots, higher polynomial degree).

In the next step we explore influential points for the ordinal fit via the SPP. The symptoms with the largest SPP contributions are **lossint** (11.91%), followed by **future** (9.68%), and by **dreams** (9.54%). The bubble plot in Figure 5 incorporates the SPP into the configuration.

[INSERT FIGURE 5 HERE]

If we were to remove these points and fit a two-dimensional, ordinal MDS model we would get a stress of (0.085), which is almost as good as the 3D solution above. We could certainly do that. However, in this particular example we decide to keep the points in the setting since, as we will see below, there is some theory behind the configurations and we want to scale all the symptoms in the same space.

Now let us look at the ordinal random stress norms (500 replications, $n = 17$, $p = 2$). The simulation results in a stress norm of 0.277 and a lower $2 \times sd$ boundary of 0.255. Not too surprisingly, our solution is way below this boundary.

Next, we run a permutation test. This example uses derived proximities; therefore, we can perform our permutations on the data directly. For each replication (we use $T = 500$ here), the Euclidean distances are computed, and, subsequently, a 2D ordinal MDS is fitted and the stress value extracted. This gives us a null distribution for the stress values and we can map our observed stress onto this distribution.

[INSERT FIGURE 6 HERE]

The results are given in Figure 6. Since none of the permuted stress values are smaller than our observed stress value, we get a p -value of 0. We also see that the permutation stress values are, on average, smaller (mean: 0.227) than the ones from the random stress norms (mean: 0.277). This again shows that the stress norm obtained from the permutation approach is sharper than the one obtained from random dissimilarities. This is not surprising as the permutation null hypothesis is more informative than the nullest of nulls.

[INSERT FIGURE 7 HERE]

In order to examine the stability of the solution we explore jackknife configurations. The leave-one-out jackknife version applied here computes $n - 1$ additional MDS solutions. The jackknife plot is given in Figure 7. The labels are at the position of the points of the configuration based on all variables, the big dots denote the jackknife centroids, and the surrounding points connected to the centroid show the configurations based on the leave-one-out principle. All leave-one-out subconfigurations were subject to a Procrustes transformation. We see that the solution is extremely stable when single variables are taken out, apart some slight instabilities for `lossint` and `dreams` (to a minor extent also for `numb` and `future`), as well as one outlier for `avoidth`. Note that these “unstable” symptoms correspond exactly to the critical points as identified by SPP.

[INSERT FIGURE 8 HERE]

Finally, we look at the interpretability of the solution. The configuration plot is given in Figure 8. None of the two dimensions can be meaningfully interpreted; this is not uncommon in MDS. What we can do, however, is establish three regions (or facets) of symptoms. These regions correspond exactly to the three symptom clusters conceptualized in the DSM-IV (American Psychiatric Association, 1994): intrusive recollection (*B-cluster*), avoidance/numbing (*C-cluster*), and arousal (*D-cluster*).

Looking at all the diagnostics in combination, leads us to the conclusion that a two-dimensional, ordinal MDS is our representation of choice. The stress value is low enough, there are no highly influential points, and the solution is stable and well interpretable.

Multidimensional Unfolding: Basics

In this second part of the article we focus on a variant of MDS called *unfolding*. The first unfolding models were proposed by Coombs (1964) and later extended to multidimensional unfolding. Traditionally, unfolding models are models of preferential

choice: n judges are asked to rank or rate m stimuli according to their preference.

Unfolding scales the judges as well as the stimuli in a joint space (*dual scaling*) so that the distances between a point that represents a judge and the points that represent the choice objects correspond to the observed preference scores.

The preference values can be converted into dissimilarities. This viewpoint is very important for two reasons: First, it allows one to consider unfolding models as a special variant of MDS. Second, it makes unfolding applicable to a wide range of data settings.

Formally, unfolding takes a rectangular dissimilarity matrix Δ of dimension $n \times m$ with elements δ_{ij} ($i = 1, \dots, n$ and $j = 1, \dots, m$) as input. The stress in Eq. (2) changes to

$$\sigma(\mathbf{X}_1, \mathbf{X}_2) = \sum_{i=1}^n \sum_{j=1}^m (\hat{d}_{ij} - d_{ij}(\mathbf{X}_1, \mathbf{X}_2))^2, \quad (6)$$

with the fitted Euclidean distances (p -dimensional space) expressed as

$$d_{ij}(\mathbf{X}_1, \mathbf{X}_2) = \sqrt{\sum_{s=1}^p (x_{1is} - x_{2js})^2}. \quad (7)$$

Analogous to standard MDS, the stress can be normalized to stress-1. Obviously, two configurations are involved in unfolding: The $n \times p$ matrix \mathbf{X}_1 of row configurations (e.g., representing individuals), and the $m \times p$ matrix \mathbf{X}_2 of column configurations (representing choice objects).

Both configurations can be represented in a joint plot. Note that unfolding operates directly on dissimilarities and fits distances. Persons are represented as “ideal points” so that the distances from each ideal point to the object points correspond to the preference scores (Borg & Groenen, 2005, p. 293). In the configuration plot all distances within persons, within objects, and between persons and objects are interpretable³.

The way the unfolding model was formulated above allows us to embed it into the SMACOF family and thus apply the majorization principle for optimization (see De Leeuw

³Note that the latter does not hold for correspondence analysis (CA), for instance, where no distances among the row and the column categories in a frequency table are defined (see Greenacre, 2007).

& Mair, 2009). There are several other ways to specify and estimate unfolding models and formulate corresponding extensions as described in Chapters 14 and 16 of Borg and Groenen (2005). As far as the estimation of unfolding models is concerned, avoiding trivial/degenerate solutions can be challenging. Details can be found in Chapter 15 of Borg and Groenen (2005), who categorize approaches that tackle this issue into three classes: adapting the unfolding data, adjusting the transformation function, and modifying the loss function.

Metric unfolding variants use the same transformation function as MDS. Nonmetric (ordinal) unfolding uses again $\hat{d}_{ij} = f(\delta_{ij})$ (monotone regression) but is technically quite complicated because a penalization on nonmetric transformations of dissimilarities with small variation is needed (Busing, Groenen, & Heiser, 2005). Another important extension is row-conditional unfolding where the choices are not directly comparable across the rows. Such row-conditionality reduces the number of constraints for the unfolding solution. So, before allowing for ever softer models that make fewer and fewer constraints, one should consider the consequences on the interpretability of the resulting models.

Unfolding Goodness-of-Fit

Most of the steps for standard MDS goodness-of-fit assessment can be applied to unfolding as they were explained above. Some of them, however, need to be adapted due to the dual scaling nature of unfolding. Let us run through them step-by-step. As in MDS, the unfolding solution is sensitive to starting solutions. Again, we can fit the model with multiple random starts and pick the one with the lowest stress-1 value or one that is theoretically much more convincing. In unfolding we can have a theoretical starting configuration for either the columns (rows random), the rows (columns random), or both in order to support convergence of the algorithm to a meaningful solution. The default setting in the **smacof** package uses a singular value decomposition of the input dissimilarity matrix.

For unfolding, neither stress rules-of-thumb nor random stress norms have been

proposed in the literature. In an environment like R random stress norms can be created by means of a simple Monte Carlo simulation. We do not provide any of them here since we will use a permutation test that provides sharper norms anyway.

A scree plot can be established as in standard MDS. Shepard diagrams can be produced for the row and column dissimilarities separately. We also get two sets of SPP values: one for the rows and one for the columns. A bubble plot can be created by incorporating the SPP information into the configuration plot (again, the larger the bubbles, the higher the SPP). In order to test for stability, jackknife and bootstrap approaches can be carried out in a similar fashion as in MDS (including Procrustes transformation).

Similar to MDS, unfolding techniques also provide us with the opportunity to fit models with circular restrictions on the configurations (either rows only, columns only, or both). We are not aware of any other confirmatory unfolding variant proposed in the literature.

Permutation tests are easy to achieve. For unconditional unfolding we can permute the input dissimilarity matrix (across rows and columns) which gives a stress null distribution. If the data reflect rank orders of a judge across the variables, one may permute within rows since the rankings are not interchangeable across rows. If a row-conditional unfolding model is fitted, the permutation has to be performed within rows since the model abandons full interchangeability. The null and alternative hypotheses can be formulated as above.

Example II: Unfolding on Personal Values

The following example uses data from Borg, Bardi, and Schwartz (2016) on people's attitudes towards various personal values. Personal values are broad trans-situational goals that serve as guiding principles in a person's life. According to Schwartz (1992), there are 10 basic values: tradition, conformity, security, power, achievement, hedonism, stimulation,

self-direction, universalism and benevolence. The inter-correlations among items measuring a person's attitudes towards these values exhibit certain gradients that can be visualized as a circle in a two-dimensional MDS space. The circle is read such that adjacent values (e.g., tradition and conformity) tend to be positively related, and those on opposing ends (e.g., security and self-direction) are negatively related.

The research question in Borg et al. (2016) was whether the Schwartz value circle also exists within persons and not only across persons (as claimed in the literature so far). The instrument they used to measure personal values is the Schwartz Value Survey (SVS; Schwartz, Sagiv, & Boehnke, 2000), which asks persons to rate various personal values "as guiding principles in their life" on a 7-point rating scale. The data collected on 327 persons and 10 variables were centered (row-wise) and converted from preferences into dissimilarities.

We start our analysis by fitting a two-dimensional metric (ratio) unfolding solution. This leads to a stress-1 value of 0.174. None of the random starts led to a lower stress, nor did a theory-based starting configuration for the column points and random coordinates for the row points. A scree plot is given in Figure 9. No clear elbow is visible but we see that for two dimensions the stress is already considerably low. Two dimensions also correspond to the dimensionality proposed for the original Schwartz value space.

[INSERT FIGURE 9 HERE]

Thus, let us proceed with $p = 2$. The resulting configuration plot is given in Figure 10. We see that the configurations reflect the approximately circular structure as hypothesized by Schwartz (1992) which is a psychologically interesting result.

[INSERT FIGURE 10 HERE]

Next we look at the SPP and check for influential observations (see Figure 11). There

are no influential outliers on the person side (four out of 327 individuals have an SPP of slightly larger than 1%). On the variable side the highest SPP values are the following: stimulation (15.53%), hedonism (15.06%), tradition (13.31%), and power (10.18%); the remaining SPP values are below 10%. We conclude that no variable is an obvious outlier, and that further replications must show whether stimulation and hedonism are always the personal values with the highest SPP values.

[INSERT FIGURE 11 HERE]

Now we perform a permutation test on the full dissimilarity matrix with 500 replications. In this data setting we assume that the cells are fully interchangeable. The stress norm (mean stress) for the permutation samples is 0.28. Our observed stress is clearly smaller than the average stress norm; the p -value is 0.

What is interesting in this application is that we actually have a well-established underlying theory even though this has not been explored at an individual level (apart from Borg et al., 2016). It says that values can be arranged on a circle with the points ordered in a particular way, and with certain points being opposed to each other. The first configuration plot in Figure 10 already suggests that we might not lose too much goodness-of-fit if we force the value configurations to be on a circle. Hence, we fit a corresponding two-dimensional unfolding model with circular restrictions on the columns. The resulting stress is 0.181 which is only slightly larger than the unrestricted stress (0.174). The restricted configuration plot is given in Figure 12.

[INSERT FIGURE 12 HERE]

The picture given by the permutation test is similar to that of the unrestricted solution. Thus, we use the two-dimensional, spherical unfolding model as our final model. We conclude that the value circle also exists within persons which, in conjunction with the

Schwartz theory, provides a good interpretability of the solution while keeping the stress value low.

Note that a single example can not take account of the breadth and all the challenges of unfolding models as described in Chapters 14 and 16 of Borg and Groenen (2005). In this particular example it occurred that the person points form a somewhat dense cloud, and the stimulus (i.e. value) points arrayed in a circle around it. Such a configuration picture is not the general case, however. A recent application which gives a different picture is presented in Mair, De Leeuw, and Wurzer (2015). In their example, in contrast to the value example here, the interpretation of the row points of the dissimilarity matrix is substantively relevant as well as the row-to-column distances.

Discussion

This article described strategies for goodness-of-fit assessment in MDS and unfolding. The main message is that researchers should not rely on mechanically using conventional stress rules and norms but should rather explore the goodness-of-fit in more depth. The diagnostic tools, measures, tests, and visualizations presented here give researchers a detailed insight into the structure of an MDS fit.

The permutation strategies elaborated in this article provide benchmarks for stress values under weakly informative null hypotheses. This makes the old random stress norms—that only provide norms for ordinal MDS anyway—obsolete. Again, one should not just look at the resulting p -value but rather integrate the results of permutation tests into the global goodness-of-fit picture, especially in relation to the permutation stress norm. Further research in the area of permutation can involve the adaption of the Mantel-test principle (Mantel, 1967) in order to test various structural hypotheses on the configuration.

As far as MDS extensions mentioned at the beginning of this article are concerned, nowadays it is somewhat rare that researchers in fields like Marketing use straight nonmetric MDS models. Rather, it may be necessary to account for characteristics such as

individual or group differences, preference vectors, or other sophisticated unfolding variants as presented in Chapter 16 of Borg and Groenen (2005). The full bundle of diagnostic tools presented here is applicable to these variants as well, except that the permutation testing principle may have to be adapted to the particular data setting. For some of these extended models one should also check the significance of additional parameters included in the model. In INDSCAL, for instance, one often obtains dimension weights that have substantial variance even though these weights explain very little variance when compared to unit weight (see Borg & Lingoes, 1978). Hence, what one should always do is check how a model with additional parameters compares to a simpler model without those parameters.

One particular MDS development which deviates conceptually from the MDS models presented in this article is Bayesian MDS. Bayesian methods for MDS and unfolding provide possibilities for estimating the number of dimensions directly (Oh & Raftery, 2001; Matlosz, 2013). These models have not been widely applied so far. Even though the number of dimensions p is part of the optimization, the user should not exclusively rely on this criterion when it comes to goodness-of-fit assessment but rather investigate various solutions using the diagnostics presented above, if applicable.

To conclude, this article equips researchers with a broad set of tools to explore goodness-of-fit in MDS and unfolding in great detail. Researchers should examine different starting solutions, use graphical fit displays (scree plot, Shepard diagram, bubble plots), descriptive measures such as SPP for influential outlier detection, permutation tests and norms, jackknife/bootstrap resampling for stability examination, and, last but not least, take into account the interpretability of a solution.

References

- American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)* (4th ed.). Washington, DC: American Psychiatric Association.
- Borg, I., Bardi, A., & Schwartz, S. H. (2016). Does the value circle exist within persons or only across persons? *Journal of Personality*. (Forthcoming)
- Borg, I., & Groenen, P. J. F. (2005). *Modern multidimensional scaling: Theory and applications* (2nd ed.). New York: Springer.
- Borg, I., & Lingoes, J. C. (1978). What weights should weights have in individual differences scaling? *Quality and Quantity*, *12*, 223–237.
- Borg, I., & Lingoes, J. C. (1980). A model and algorithm for multidimensional scaling with external constraints on the distances. *Psychometrika*, *45*, 25–38.
- Borg, I., & Mair, P. (2016). The choice of initial configurations in multidimensional scaling: local minima, fit, and interpretability. *Austrian Journal of Statistics*. (Forthcoming)
- Borg, I., & Shye, S. (1995). *Facet theory: Form and content*. Newbury Park, CA: Sage.
- Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, *27*, 509–540.
- Busing, F. M. T. A., Groenen, P. J. F., & Heiser, W. J. (2005). Avoiding degeneracy in multidimensional unfolding by penalizing on the coefficient of variation. *Psychometrika*, *70*, 71–98.
- Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an n -way generalization of Eckart-Young decomposition. *Psychometrika*, *35*, 283–320.
- Cliff, N. (1973). Scaling. *Annual Review of Psychology*, *24*, 473–506.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.
- Cox, T. F., & Cox, M. A. A. (1991). Multidimensional scaling on a sphere. *Communications in Statistics: Theory and Methods*, *20*, 2943–2953.
- Cox, T. F., & Cox, M. A. A. (2001). *Multidimensional scaling* (2nd ed.). Boca Raton, FL:

Chapman & Hall/CRC.

- De Leeuw, J. (1977). Applications of convex analysis to multidimensional scaling. In J. Barra, F. Brodeau, G. Romier, & B. van Cutsem (Eds.), *Recent developments in statistics* (p. 133-145). Amsterdam, The Netherlands: North Holland Publishing Company.
- De Leeuw, J. (1984). Differentiability of Kruskal's stress at a local minimum. *Psychometrika*, *49*, 111-113.
- De Leeuw, J. (1988). Convergence of the majorization method for multidimensional scaling. *Journal of Classification*, *5*, 163-180.
- De Leeuw, J., & Heiser, W. J. (1980). Multidimensional scaling with restrictions on the configuration. In P. Krishnaiah (Ed.), *Multivariate analysis, volume V* (pp. 501-522). Amsterdam, The Netherlands: North Holland Publishing Company.
- De Leeuw, J., & Mair, P. (2009). Multidimensional scaling using majorization: SMACOF in R. *Journal of Statistical Software*, *31*(3), 1-30.
- De Leeuw, J., & Mair, P. (2015). Shepard diagram. In *Wiley StatsRef: Statistics Reference Online*. New York: Wiley.
- De Leeuw, J., & Meulman, J. (1986). A special jackknife for multidimensional scaling. *Journal of Classification*, *3*, 97-112.
- De Leeuw, J., & Stoop, I. (1984). Upper bounds for Kruskal's stress. *Psychometrika*, *49*, 391-402.
- Gower, J. C., & Legendre, P. (1986). Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification*, *3*, 5-48.
- Greenacre, M. (2007). *Correspondence analysis in practice* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.
- Groenen, P. J. F. (1993). *The majorization approach to multidimensional scaling: Some problems and extensions* (Unpublished doctoral dissertation). Leiden University, Leiden, The Netherlands.

- Heiser, W. J., & Meulman, J. (1983). Constrained multidimensional scaling, including confirmation. *Applied Psychological Measurement*, *7*, 381–404.
- Jacoby, W. G., & Armstrong, D. A. (2014). Bootstrap confidence regions for multidimensional scaling solutions. *American Journal of Political Science*, *58*, 264–278.
- 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.
- Lawler, E. E. (1967). The multitrait-multirater approach to measuring job managerial performance. *Journal of Applied Psychology*, *51*, 369–381.
- Linting, M. (2007). *Nonparametric inference in nonlinear principal components analysis: Exploration and beyond* (Unpublished doctoral dissertation). Leiden University, Leiden, The Netherlands.
- Mair, P., De Leeuw, J., & Wurzer, M. (2015). Multidimensional unfolding. In *Wiley StatsRef: Statistics Reference Online*. New York: Wiley.
- Mair, P., Rusch, T., & Hornik, K. (2014). The Grand Old Party: A party of values? *SpringerPlus*, *3*(697).
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, *27*, 209–220.
- Matlosz, K. (2013). *Bayesian multidimensional scaling model for ordinal preference data* (Unpublished doctoral dissertation). Columbia University, New York.
- McNally, R. J., Robinaugh, D. J., Wu, G. W. Y., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clinical Psychological Science*, *3*, 836–849.
- Oh, M. S., & Raftery, A. E. (2001). Bayesian multidimensional scaling and choice of dimension. *Journal of the American Statistical Association*, *96*, 1031–1044.

- R Development Core Team. (2016). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <http://www.R-project.org> (ISBN 3-900051-07-0)
- Ramsay, J. O. (1977). Maximum likelihood estimation in MDS. *Psychometrika*, *42*, 241–266.
- Ramsay, J. O. (1988). Monotone regression splines in action. *Statistical Science*, *3*, 425–461.
- Rusch, T., Mair, P., & Hornik, K. (2015). *Cops: Cluster optimized proximity scaling* (Tech. Rep. No. 2015/1). Vienna, Austria: WU Vienna University of Economics and Business.
- Schwartz, S. H. (1992). Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, *25*, 1–62.
- Schwartz, S. H., Sagiv, L., & Boehnke, K. (2000). Worries and values. *Journal of Personality*, *68*, 309–346.
- Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function II. *Psychometrika*, *27*, 219–246.
- Skinner, H. A. (1979). Dimensions and clusters: A hybrid approach to classification. *Applied Psychological Measurement*, *3*, 327–341.
- Spence, I., & Ogilvie, J. C. (1973). A table of expected stress values for random rankings in nonmetric multidimensional scaling. *Multivariate Behavioral Research*, *8*, 511–517.
- Spence, I., & Young, F. W. (1978). Monte Carlo studies in nonmetric scaling. *Psychometrika*, *43*, 115–117.
- Stenson, H. H., & Knoll, R. L. (1969). Goodness of fit for random rankings in Kruskal's nonmetric scaling procedures. *Psychological Bulletin*, *72*, 122–126.
- Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B*

(Statistical Methodology), 63, 411–423.

Torgerson, W. S. (1952). Multidimensional scaling: I. Theory and method. *Psychometrika*, 17, 401–419.

Vera, J. F., Heiser, W. J., & Murillo, A. (2007). Global optimization in any Minkowski metric: a permutation-translation simulated annealing algorithm for multidimensional scaling. *Journal of Classification*, 24, 277–301.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993). *The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility*. Paper presented at the meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Welch, W. J. (1990). Construction of permutation tests. *Journal of the American Statistical Association*, 85, 693–698.

Table 1

PTSD checklist-civilian items

Variable	Item Wording
intrusion	Repeated, disturbing memories, thoughts, or images of a stressful experience from the past?
dreams	Repeated, disturbing dreams of a stressful experience from the past?
flash	Suddenly acting or feeling as if a stressful experience were happening again?
upset	Feeling very upset when something reminded you of a stressful experience from the past?
physior	Having physical reactions (e.g., heart pounding, trouble breathing, sweating) when something reminded you of a stressful experience from the past?
avoidth	Avoiding thinking about or talking about a stressful experience from the past or avoiding having feelings related to it?
avoidact	Avoiding activities or situations because they reminded you of a stressful experience from the past?
amnesia	Trouble remembering important parts of a stressful experience from the past?
lossint	Loss of interest in activities that you used to enjoy?
distant	Feeling distant or cut off from other people?
numb	Feeling emotionally numb or being unable to have loving feelings for those close to you?
future	Feeling as if your future will somehow be cut short?
sleep	Trouble falling or staying asleep?
anger	Feeling irritable or having angry outbursts?
concen	Having difficulty concentrating?
hyper	Being “super-alert” or watchful or on guard?
startle	Feeling jumpy or easily startled?

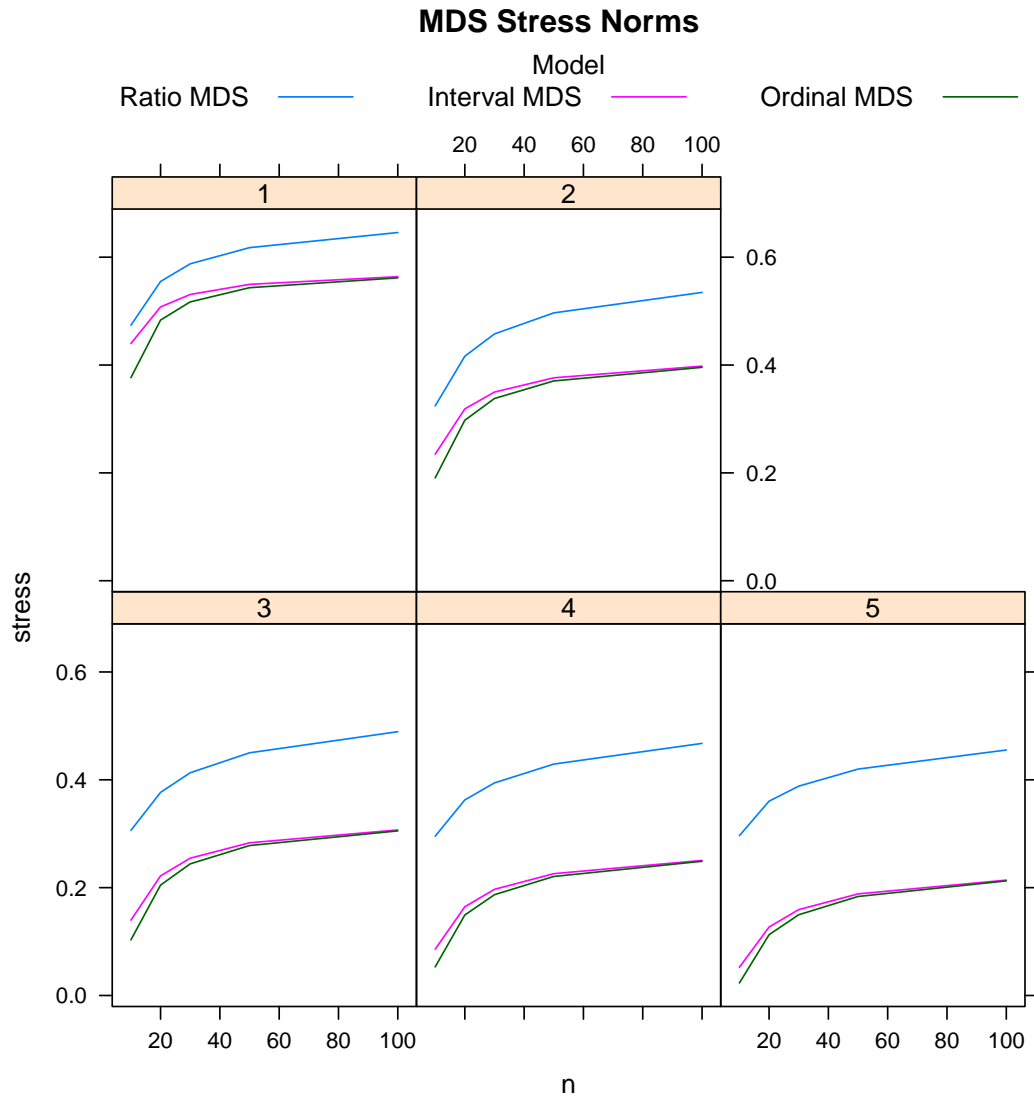


Figure 1. Random stress norms for ratio, interval, and ordinal MDS across different dimensions p (panels) and different numbers of objects n (x-axis).

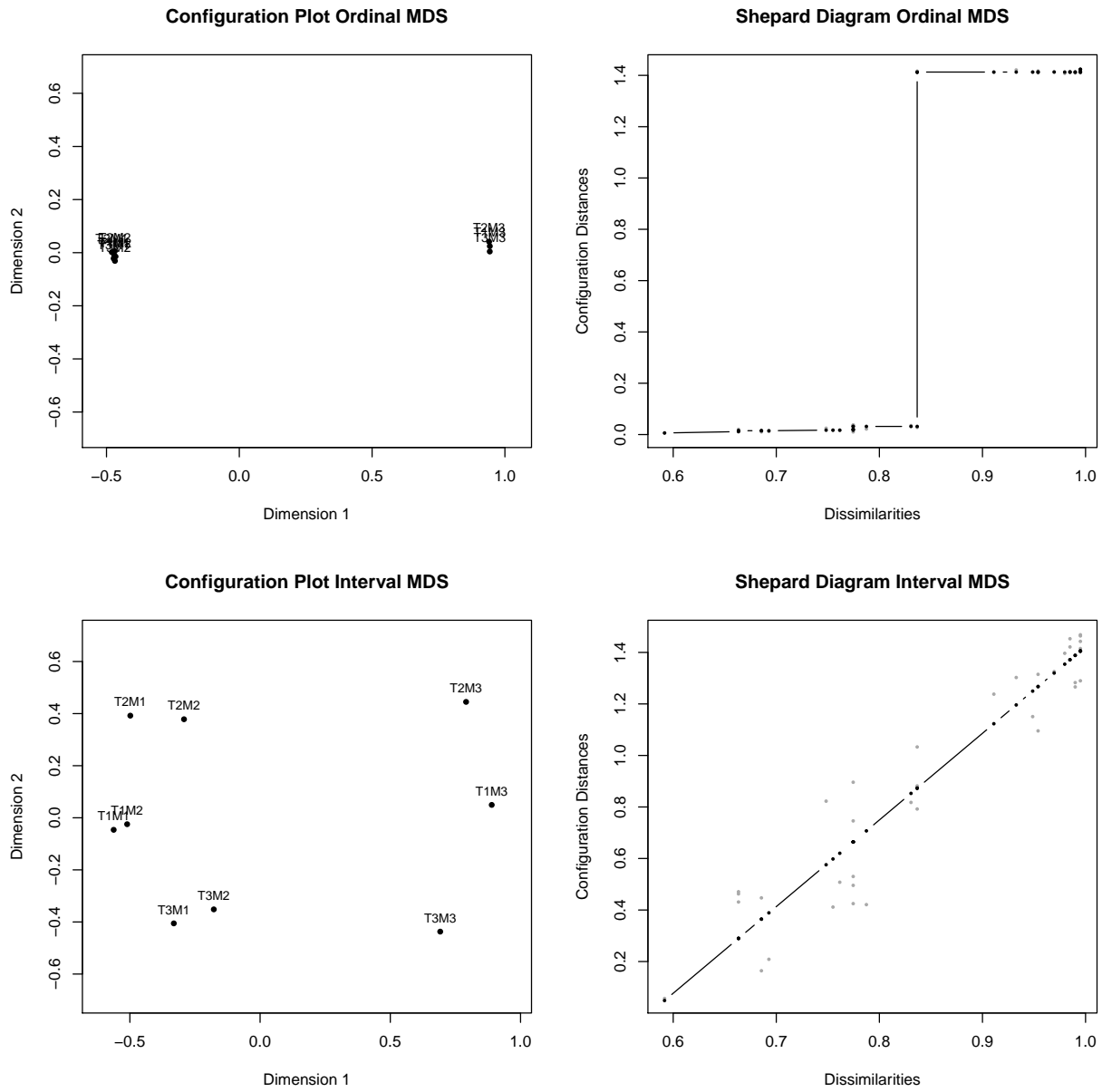


Figure 2. Top panel: configuration plot and Shepard diagram for ordinal solution. Bottom panel: configuration plot and Shepard diagram for interval solution

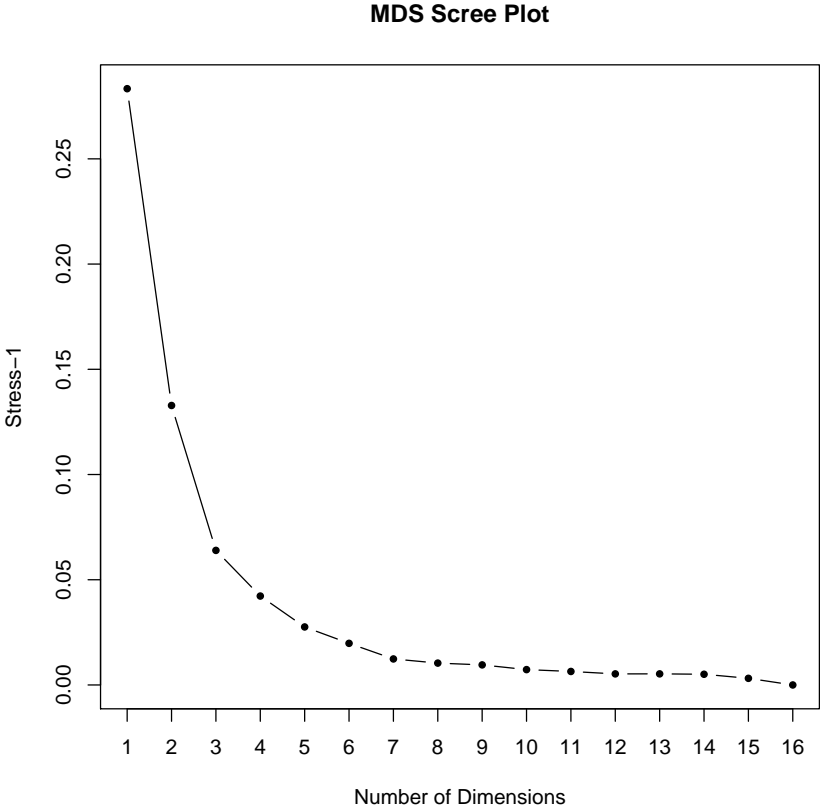


Figure 3. Scree plot for ordinal MDS solutions ($p = 1, \dots, 16$) on Wenchuan PTSD data.

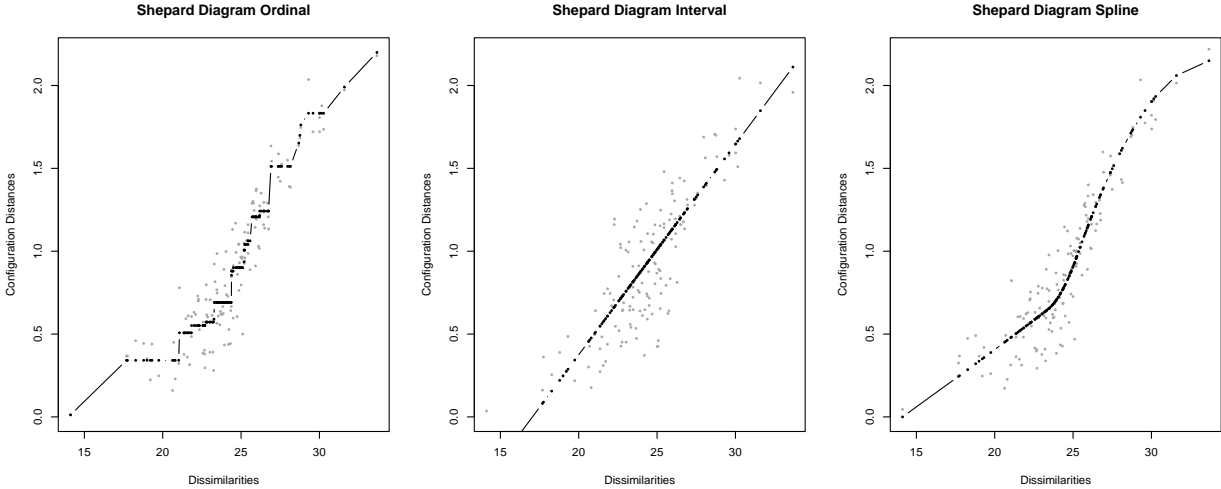


Figure 4. Shepard diagram for ordinal, interval, and spline MDS fit

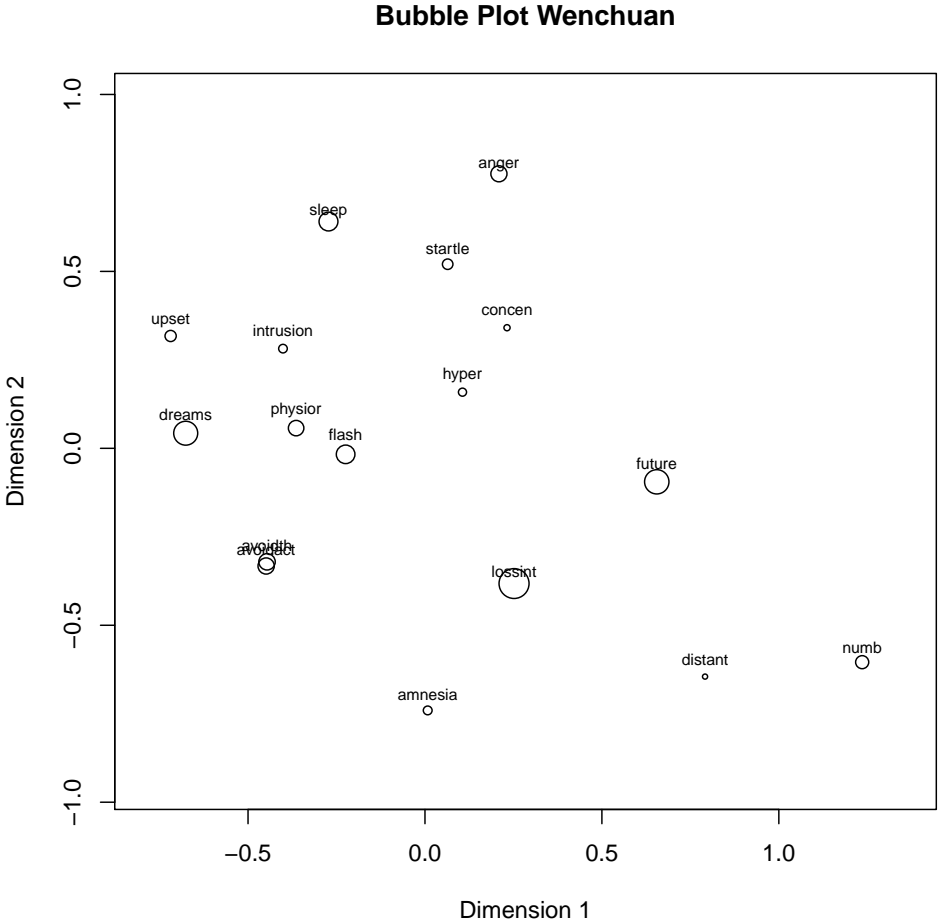


Figure 5. Bubble plot (the larger the bubble, the higher the SPP).

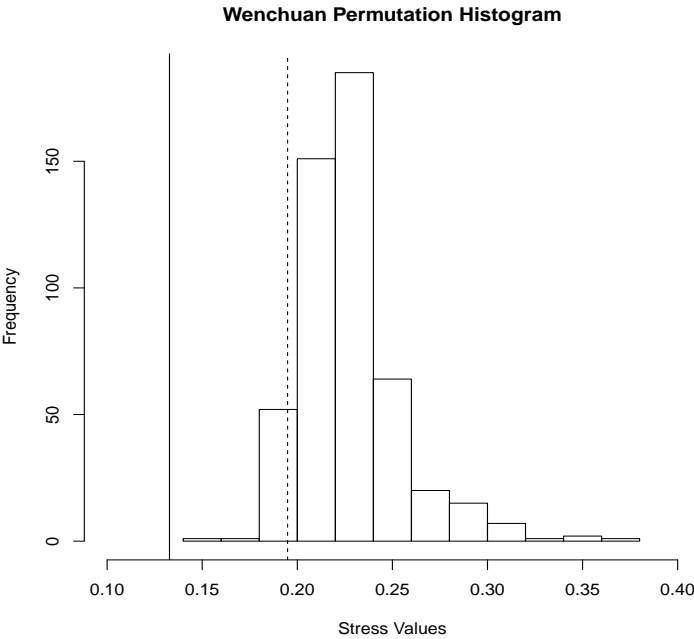


Figure 6. Permutation histogram: The solid vertical line denotes the observed stress value, the dotted line the 5% cutoff for the rejection region (stress of 0.195).

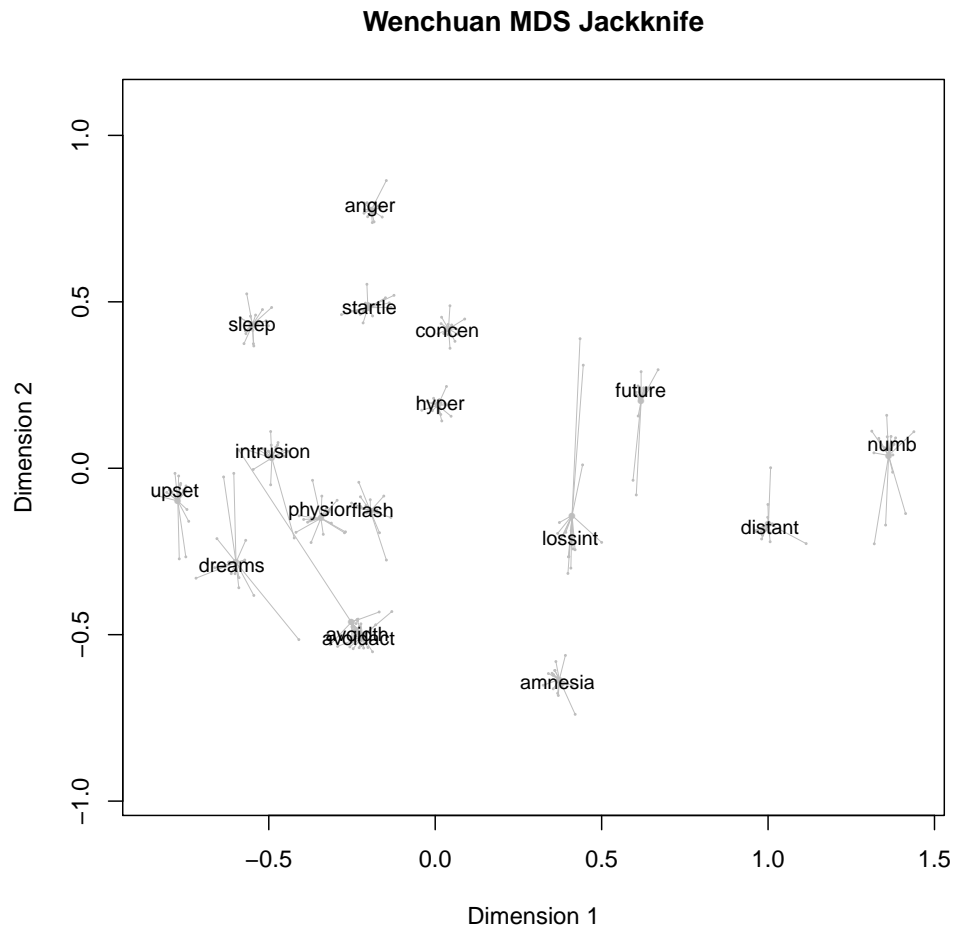


Figure 7. Jackknife plot. The points of the original configuration are shown by the symptom labels; the end points of the “stars” show the respective point position of the $n - 1$ jackknife configurations, subsequently connected to their centroid.

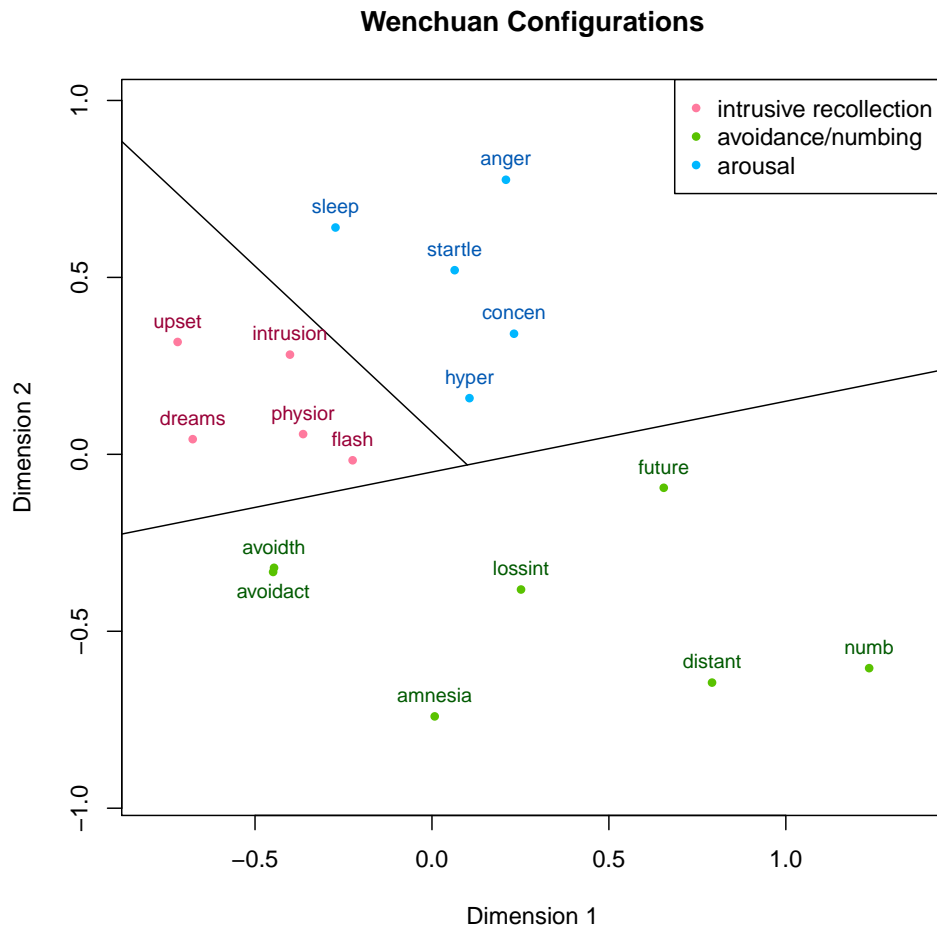


Figure 8. Final configuration plot for Wenchuan MDS. Regions represent the three DSM-IV clusters: intrusive recollection, avoidance/numbing, and arousal.

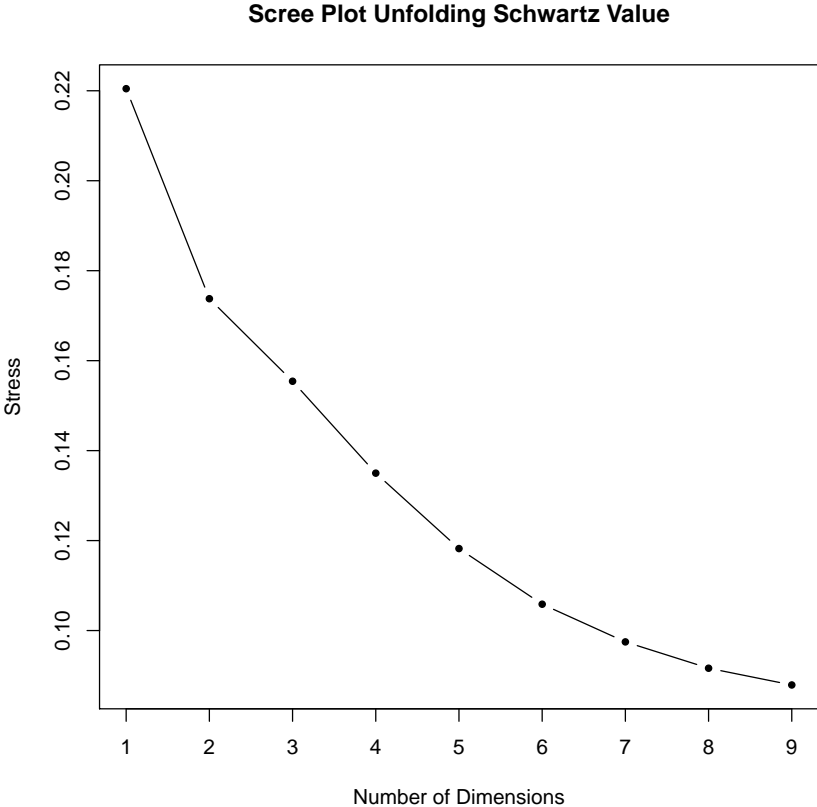


Figure 9. Scree plot for unrestricted unfolding solution on the Schwartz values.

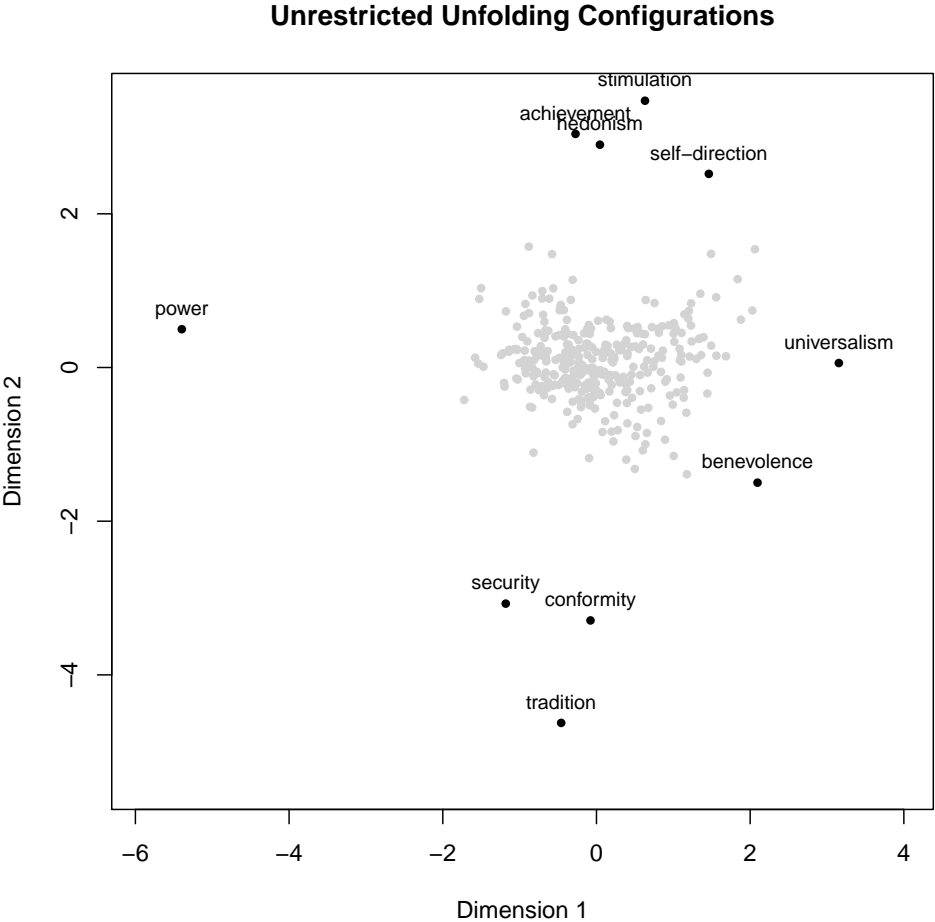


Figure 10. Unrestricted unfolding configurations on Schwartz values: the persons are scaled (gray points) as well as the variables (value items).

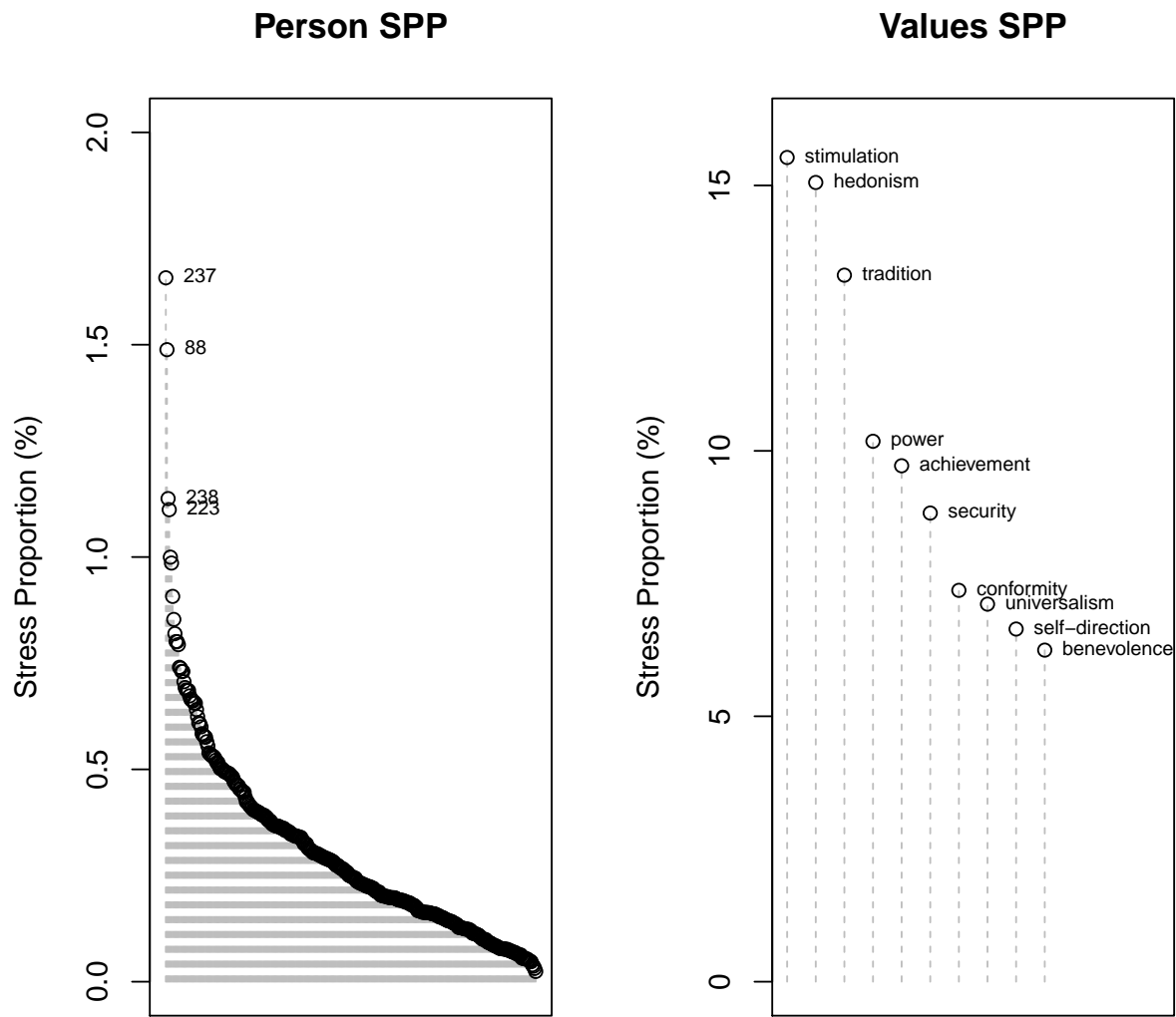


Figure 11. SPP plot for unfolding solution. Left panel: persons. Right panel: personal values.

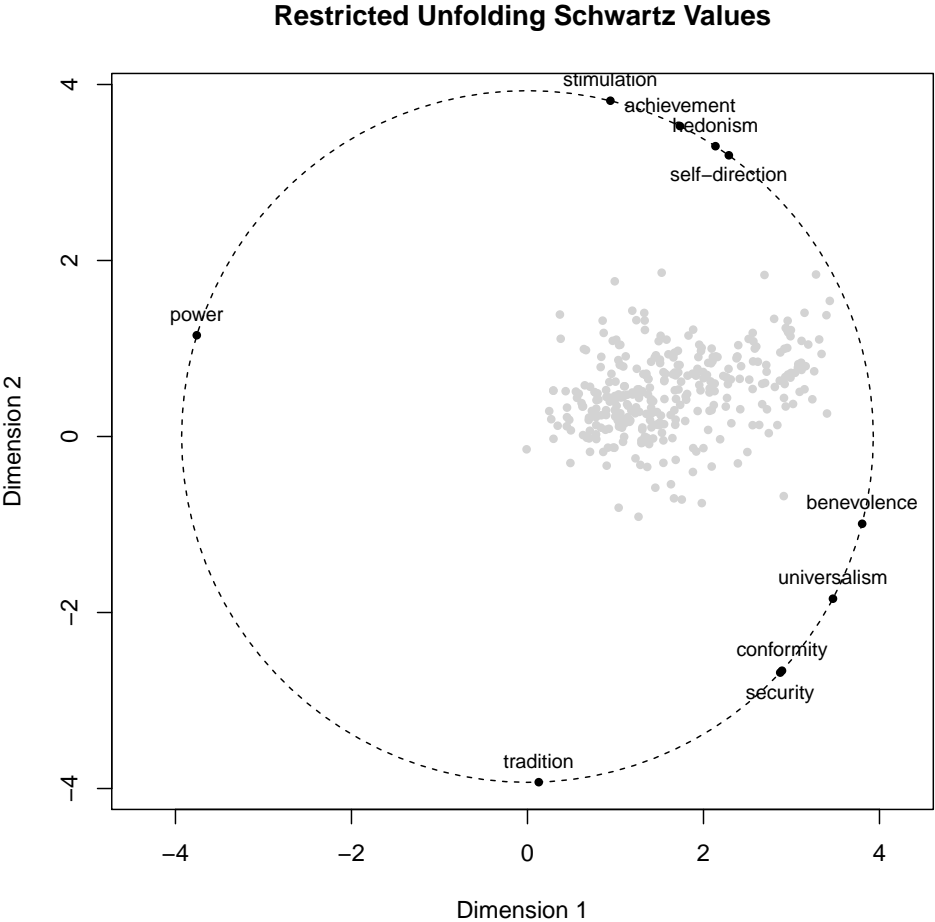


Figure 12. Circular restricted unfolding on Schwartz values with persons (gray points) and value items.