Two types of single-peaked data: Correspondence analysis as an alternative to principal component analysis

Marike Polak*, Willem J. Heiser, Mark de Rooij

Leiden University Institute for Psychological Research, Methodology and Statistics Unit, Leiden, The Netherlands

A R T I C L E   I N F O

A B S T R A C T

It has been argued that principal component analysis (PCA) is not appropriate for analyzing data conforming to single-peaked response models, also referred to as unfolding models. An overview of these findings is given, which relates them to the distinction between two types of unfolding models; namely, models that are either a quadratic function of the person-to-item distances or an exponential function of these distances. This distinction is easy to recognize empirically because the inter-item correlation matrix for the two types of data typically shows different patterns. Furthermore, for both types of unfolding models, correspondence analysis (CA), which is a rival method for dimensionality reduction, outperforms PCA in terms of representation of both person and item locations, especially for the exponential unfolding model. Finally, it is shown that undoubled CA outperforms doubled CA for both types of unfolding models. It is argued that performing CA on the raw data matrix is an unconventional, but meaningful approach to scaling items and persons on an underlying unfolding scale. A real data example on personality assessment is given, which shows that for this type of data (undoubled) CA is to be preferred over PCA.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

In this paper we explore the surplus value of correspondence analysis (CA) over principal component analysis (PCA) for analyzing one-dimensional, single-peaked responses, i.e. data conforming to a one-dimensional unfolding model. We will discuss continuous, binary and graded responses.

Single-peaked (unimodal) data naturally arise in a variety of research settings, e.g. marketing research (for example, DeSarbo et al. (2002)), ecological research (for example, De'ath (1999)), and archeology (for example, Kendall (1971)). In psychology single-peaked response curves can be found, for instance, in attitude measurement (Roberts et al., 2000): people with moderate tolerance towards abortion are less likely to agree with items that are either very much in favor of abortion or very much against it.

The essence of an unfolding model is that the probability of agreement with a certain item is inversely related to the distance between the position of the item on the latent continuum and the position of the respondent; the closer an item is located near the respondent's position on the latent continuum, the more likely the respondent will agree with it. In this case, the latent continuum is called bipolar: ranging from a negative extreme (very much against abortion), via a neutral midpoint (neither against nor in favor of abortion), to a positive extreme (very much in favor of abortion). In the unfolding literature the positions of respondents or objects on this continuum are referred to as ideal points (Coombs, 1964).

There is a vast amount of literature on the inappropriateness of PCA for analyzing data conforming to an unfolding model (Coombs and Kao, 1960; Ross and Cliff, 1964; Davison, 1977; Van Schuur and Kiers, 1994; Van Schuur and Kruijtbosch, 1995; Andrich, 1996; Rost and Luo, 1997; Andrich and Styles, 1998; Roberts et al., 1999, 2000; Maraun and Rossi, 2001).

---

* Corresponding address: Psychometrics and Research Methodology Group, Leiden University Institute for Psychological Research, Leiden University, Wassenaarseweg 52, P.O. Box 9555, 2300 RB Leiden, The Netherlands. Tel.: +31 71 5273663.
E-mail address: polak@fsw.eur.nl (M. Polak).

The main conclusions from this literature are, first, that PCA of one-dimensional unfolding data results in a two-component solution, leading to erroneous conclusions about the dimensionality of the data. Second, the component scores of the persons with extreme positions on the latent scale underestimate the true positions. Third, component loadings of the items with extreme positions on the latent scale are underestimated, resulting in a non-optimal item selection. In the two-component PCA solution both persons and items either lie on a semi-circle (cf. Davison (1977)) or on a semi-circle with inwardly folded endpoints (cf. Roberts et al. (2000)). This inward bending of the endpoints is what is meant by underestimation of the true positions of the extreme persons and items. The current paper offers CA as an alternative to PCA and relates the different problems with PCA, described in the above, to the distinction between two types of unfolding models. To start with the latter, on the one hand, we have models that are a quadratic function of the person-to-item distances, and on the other hand, we have models that are an exponential function of these distances.

The often quoted paper by Davison (1977, see also Maraun and Rossi (2001)) discusses the quadratic unfolding model. For data conforming to this type of model, PCA only suffers the aforementioned "extra-component" problem, but not the problem of underestimation of the locations of extreme persons or items. That is, in the two-component solution, the person and item locations lie on a semi-circle. Furthermore, the inter-item correlation matrix of this type of data shows a "simplex-like" pattern, also referred to as Robinson pattern (Hubert et al., 1998). That is, when the items are ordered according to their location on the latent scale, the correlations along the diagonal of the matrix will be highly positive, moving downward and to the left, the correlations will decrease first to zero, and will decrease further to negative in the lower left-hand corner.

The papers from the field of unfolding item response theory (IRT) (for example, Andrich (1996), Rost and Luo (1997), Andrich and Styles (1998) and Roberts et al. (2000)) discuss exponential unfolding models. For data conforming to this type of models, PCA suffers both the "extra-component" problem, and the problem of underestimation of the locations of extreme persons and items. That is, in the two-component solution, the person and item locations lie on a semi-circle with inwardly bending extremes, a pattern which can be described as a "horseshoe" pattern (cf. Greenacre (1984, p. 226–232)). The problem of the inwardly bending extremes in the PCA solution has been discussed also in the field of ecology (Swan, 1970; Noy-Meir and Austin, 1970; Hill, 1973; De'ath, 1999).

In this paper CA is proposed as an alternative to PCA, since CA is known to represent single-peaked data correctly: Ter Braak (1985) showed that CA approximates the maximum likelihood solution of the Gaussian ordination model; Heiser (1981) showed that CA recovers the person and item order of error-free ratings conforming to an unfolding model. In Section 1.1, using CA as unfolding technique is explained further.

It is known that when unfolding data are strongly one-dimensional, a two-dimensional CA representation will show what is often referred to as the "arch-effect", where the items and persons are ordered along an arch (but also on the first dimension) according to their position on the scale (for example, Hill (1974) and Hill and Gauch (1980)). We prefer the term "arch" to "horseshoe", to stress the importance of the outward bending of the extremes of the arch. In that case, the first dimension reflects the correct order of items and persons, as opposed to solutions with inwardly bending extremes (which is the usual shape of a horseshoe), where the order of the items and persons gets mixed up at the endpoints of the first dimension.

When rating scale data are analyzed with CA, the variables are usually "doubled" to create pairs of variables, which form the positive and the negative poles of the rating scale (see for example, Greenacre (1993, Chapter 19), and Greenacre (2007, Chapter 23)). We will explain this type of data coding in CA in Section 1.2. In this paper we show that when unfolding data are doubled, CA, like PCA, is hampered by the undesirable inward bending of the extremes.

In this paper CA with and without doubling is compared to PCA with and without varimax rotation. For this purpose, we simulated continuous, binary, and graded responses using three different unfolding models, which are described in Section 1.3. The first is the quadratic unfolding model as discussed by Davison (1977), which results in continuous responses. Of the exponential unfolding model two variants are compared: the Gaussian ordination model (Ihm and van Groenewoud, 1984), which results in binary responses and the generalized graded unfolding model (Roberts et al., 2000), which results in graded responses. Furthermore, an empirical data set concerning the measurement of personality development was analyzed.

1.1. CA as unfolding technique

In this paragraph we explain to what types of data CA can be applied, and we explain the rationale behind using CA as unfolding technique. CA is a multivariate technique primarily developed for the analysis of contingency table data (Benzécri, 1992; Greenacre, 1984). However, the technique can be applied to a broader range of data types, as long as the entries of the table contain measures of association strength between row entries and column entries. The association measure is assumed to be some non-negative quantity, where lack of association is indicated by a zero entry (Heiser, 2001).

In the current paper, we use CA as an unfolding technique. Typically an advantage of CA in this context is that it simultaneously scales both persons and items. Of the three most common normalizations in CA (i.e. row principal, column principal, and symmetrical normalization) we choose row principal normalization, so that a person is represented as the centroid (weighted average, with weights proportional to the ratings) of the items he has rated. This approach results in an interpretation of person scores as ideal points (Coombs, 1964). A higher rating of a given person on a given item results in a smaller person-to-item distance in the CA solution. Hence the expected responses are a single-peaked function of the person scores in the CA solution.

The distances between persons in a CA solution with row principal normalization approximate chi-square distances from below (Meulman, 1982, p. 33). The chi-square distance between two persons differs from the usual Euclidean distance, in that for each item, the squared difference between the persons' scores is weighted by the inverse marginal proportion (i.e. the mass) of each item. As a consequence, persons and items with low mass tend to lie more in the periphery of the CA solution (see for example Ter Braak and Prentice (1988), reprinted in Ter Braak and Prentice (2004, p. 262–263)). In the context of attitude items with ratings ranging from 0 (totally disagree) to 5 (totally agree), an item that few persons choose, i.e. an extreme item will have a low mass. Analogously, a person who agrees with only one item, is very likely to have an extreme opinion, and will have a low mass. We will show that for these extreme items and persons CA (without doubling) typically results in appropriate scale values.

1.2. Data coding in CA: Undoubled versus doubled data

In this paragraph we discuss two types of data coding in CA: undoubled and doubled data. These two approaches are also known as, respectively, asymmetric and symmetric treatment of response categories (see Gifi (1990, p. 294–295)).

Asymmetric treatment of response categories implies performing CA on the raw data table, where, for the simple case of binary responses, disagreement is denoted with 0 and agreement with 1. In effect, only agreement implies similarity between respondents, and not disagreement.

Symmetric treatment of the response categories demands a type of recoding of the data commonly known as "doubling" (see, for example, Benzécri (1992) or (Greenacre, 1984)). This is a type of data coding that complements a respondent's original ratings with the reverse of these ratings, which are obtained by subtracting the ratings from the maximum rating. For example, for a person with the ratings 0, 2, and 4 on three items with a six-point scale ranging from 0 to 5, the complete set of doubled scores would be 0, 2, 4 along with 5, 3, 1. In effect, both shared *agreement* with a certain item and shared *disagreement* imply similarity between respondents. An argument for this procedure is, that agreement with a statement is the same as disagreement with the opposite of this statement, so that all items need not be worded in the same direction.

However, in Heiser (1981, Chapter 4) as well as in Benzécri (1992, p. 391, where we assume that in the final paragraph on p. 391 the word "not" is missing by mistake after the word "is" in the sentence "But the presence or absence in a plant of a quality such as being a perennial is of the same nature") it is stressed that if response categories are thought to give an asymmetrical type of information, CA should be preformed on the undoubled (raw) data. Even when all items are not worded in the same direction, no reverse scoring is needed, as long as the *disagree*-category is coded with a zero score. In this case, the "attraction power" of items to persons, which is reflected in small person-to-item distances the CA solution, is determined by high ratings. As a consequence, the proximity between persons in the CA solution depends on (the level of) shared agreement, and not on shared disagreement.

The argument for asymmetric treatment of the response categories is that a respondent can have only one reason for agreeing with a certain statement, but either one of two different reasons for disagreeing. That is, a respondent disagrees with the statement when he is either too "positive" to agree with the statement or too "negative".

1.3. Three different single-peaked models

In the following we discuss the three different unfolding models that were used to generate single-peaked response data. We classify these models as either one of two different types of unfolding models, that is, quadratic or exponential. The first model is a quadratic function of the person-to-item distances, whereas the second and the third model are exponential functions of these distances.

1.3.1. Model 1: Metric unfolding model

To recognize single-peaked data empirically, Davison (1977) postulated predictions about the correlations and factor structure of responses z_{ij} to various items where the responses fit a metric, unidimensional unfolding model. Two models were compared. Firstly a model producing error free data:

$$z_{ij} = a_j(x_i - y_j)^2 + b_j, \tag{1}$$

where

z_{ij} is the response of person i on item j;

a_j is the discrimination parameter for item j;

x_i is the ideal point for person i on the underlying continuum;

y_j is the location of item j on the underlying continuum;

b_j is the maximum of the curve for item j.

The discrimination parameter for a given item j, a_j, indicates the steepness of the response curve. In ecology, the inverse of the discrimination parameter is called the *tolerance* of species j, which is a measure of ecological amplitude. That is, the steeper the response curve, the smaller a species tolerance. Note that $a_j < 0$, otherwise the response curve would have a minimum instead of a maximum.

Secondly, Davison discussed a model producing fallible data:

$$z_{ij} = a_j(x_i - y_j + \mathbf{E}_{ij})^2 = (x_i - y_j)^2 + \sigma^2_{\mathbf{E}_{ij}} + e_{ij},\tag{2}$$

where

$\mathbf{E}_{ij}$ is a random normal deviate;

$\sigma^2_{\mathbf{E}_{ij}}$ is the variance of $\mathbf{E}_{ij}$;

$e_{ij} = z_{ij} - (x_i - y_j)^2 - \sigma^2_{\mathbf{E}_{ij}}$.

Under the assumption of model (1) with $a_j = -1$, it follows from the results of Ross and Cliff (1964) that the matrix $\mathbf{Z}$ with elements z_{ij} has rank 3. One of the three components involves the quantities $y_j^2 + b_j$, which are constant across the rows of $\mathbf{Z}$, the other one involves the quantities x_i^2, which are constant across the columns of $\mathbf{Z}$, and the third one the x_i and y_j themselves. In addition, Ross and Cliff showed that centering the columns of $\mathbf{Z}$ reduces its rank to two. In addition to these results, Davison (1977) concluded that (a) the item by item correlation matrix displayed a simplex-like pattern, (b) the signs of first-order partial correlations can be specified in an empirically testable manner, and (c) the items will have a semi-circular, two-factor structure. Along the semi-circle, variables will be ordered by their positions on the latent dimension. This latter fact is influenced by the amount of error included in the model. The most extreme items become mixed up with the last but one extreme items. These conclusions were based on data sets with 100 persons and 10 items, where the items had fixed equally spaced true scale values y_j ranging from -3.00 to $+3.00$, and the 100 person scores x_i were randomly sampled from a normal distribution, $N(0, 1)$. It turned out that the correlations and factor structure were robust to non-normality of the person score distribution.

It should be noted that in CA not only the columns are centered, but the rows as well (double centering; Gifi (1990, Chapter 8)). For this case, Schoenemann (1970) showed that double centering of $\mathbf{Z}$ further reduces the rank to one, and that the x- and y-scores are recovered up to a scale factor. Therefore, when we generate data under the Davison model, we will obtain exactly one component with non-zero inertia in CA, due to the double centering. However, the joint scale of the scores depends on the chosen normalization, and may not be equal to the original one.

1.3.2. Model 2: Gaussian ordination model

Here it will be shown that CA approximates the Gaussian ordination model. This is a well-known model in the field of ecology for the single-peaked relationships between the abundance of a species and some environmental variable. However, it could also model the single-peaked relationships between the attitude of a person and some attitude item. Results follow from Ter Braak (1987) and Ihm and van Groenewoud (1984). We will start with the Gaussian ordination model as proposed by Ihm and van Groenewoud. This model is somewhat more general than the standard model since it has an extra parameter (α_i^*) to account for different masses of the persons. The response z_{ij} of person i on item j is approximated by a model using maximum likelihood given a binomial (or multinomial) distribution. The Gaussian ordination model is

$$\pi_{ij} = \alpha_i^* \beta_j^* \exp\left(-\frac{1}{2}(x_i - y_j)^2 / t_j^2\right),\tag{3}$$

where

π_{ij} is the probability that person i agrees with item j;

x_i is the ideal point for person i on the underlying continuum;

y_j is the location of item j on the underlying continuum;

β_j^* is the maximum of the curve for item j;

t_j^2 is the discrimination parameter for item j.

Assuming $t_j = t$ (equal discrimination parameters) we can rewrite (3) into

$$\pi_{ij} = \alpha_i \beta_j \exp\left(x_i y_j\right),\tag{4}$$

with $\alpha_i = \alpha_i^* / \exp(\frac{-x_i^2}{2t^2})$ and $\beta_j = \beta_j^* / \exp(\frac{-y_j^2}{2t^2})$.

Using the Taylor expansion of first order we obtain

$$\pi_{ij} \approx \alpha_i \beta_j \left(1 + x_i y_j\right).\tag{5}$$

The least-squares estimate of $\alpha_i \beta_j$ is $\widehat{\alpha_i \beta_j} = \frac{z_{i+} z_{+j}}{z_{++}}$.

Inserting this expression in (5) we obtain

$$\pi_{ij} \approx \frac{z_{i+} z_{+j}}{z_{++}} \left(1 + x_i y_j\right),\tag{6}$$

which is the CA model with one component. Note that the first-order Taylor expansion works well for small values of the interaction term $x_i y_j$. But the relation of CA with Gaussian ordination model holds true as well for large values (Ter Braak, 1985, 1987). See also Ter Braak (1988) and Zhu et al. (2005) for this link in constrained CA.

1.3.3. Model 3: Probabilistic IRT unfolding model GGUM

The generalized graded unfolding model (GGUM) is a parametric item response model that has been well developed and incorporates features such as variable item discrimination and variable threshold parameters for the response categories (Roberts et al., 2000). The GGUM allows for binary or graded responses, but will be used in the current paper to generate responses on a six-point rating scale. One premise of the GGUM is that for each person there are two *subjective responses* associated with each observable response, except for the totally agree response. These subjective responses can be seen as two distinct reasons for a person's response. For instance, when a person strongly disagrees with a certain items this could be for either of two reasons. If on the underlying continuum the item is located more to the right extreme than the person, the person disagrees *from below* the item. However, if the item is located more to the left extreme than the person, the person disagrees *from above* the item. The probability that a person will respond using a particular observable answer category is defined as the sum of the probabilities associated with the two corresponding subjective responses. Specifically, the model has the form:

$$P(Z_j = z | x_i) = \frac{\exp\left\{t_j\left[z(x_i - y_j) - \sum_{k=0}^{z}\tau_{jk}\right]\right\} + \exp\left\{t_j\left[(M - z)(x_i - y_j) - \sum_{k=0}^{z}\tau_{jk}\right]\right\}}{\sum_{\omega=0}^{C}\left(\exp\left\{t_j\left[\omega(x_i - y_j) - \sum_{k=0}^{\omega}\tau_{jk}\right]\right\} + \exp\left\{t_j\left[(M - \omega)(x_i - y_j) - \sum_{k=0}^{\omega}\tau_{jk}\right]\right\}\right)},$$

(7)

where

Z_j is an observable response to attitude item j;

$z = 0$ ($z = 0, 1, 2, \ldots, C$) corresponds to the strongest level of disagreement;

$z = C$ corresponds to the strongest level of agreement;

M is the number of subjective response categories minus 1;

C is the number of response categories minus 1 ($M = 2C + 1$);

x_i is the location of person i on the attitude continuum;

y_j is the location of item j on the attitude continuum;

t_j is the discrimination of attitude statement j; and

τ_{jk} is the location of the kth subjective response category threshold on the attitude continuum relative to the location of item j.

2. Method

The aim of the present research is to compare the performance of CA (with and without doubled items) and PCA (with and without varimax rotation) in terms of the recovery of the "true" scale values. Three types of scale values were of interest: person scale values, item scale values, and scale values of persons and items taken together, referred to as the joint scale.

We chose to include CA with doubled items, with the aim of testing the presumption that, in the case of unfolding data, asymmetric treatment of response categories (implied by CA of undoubled data) gives a better recovery of the true scale values, than symmetric treatment of response categories (implied by CA of doubled data); even when ratings are analyzed. This is important, since doubling is the standard procedure for analyzing ratings with CA. In this paper it is shown in which situations this procedure is not suited.

Furthermore, we performed a standardized PCA, that is, PCA of the inter-item correlation matrix, because it is the general approach in the literature we referred to in this paper (for example, Davison (1977)). We included PCA with varimax rotation, since this is the most commonly used rotation in the context of item scaling and item selection (for example, Tabachnick and Fidell (2001, Chapter 13)). It is relevant to see that rotation aggravates the problems that hamper the PCA solution for unfolding data. Note that CA with doubled items and PCA are closely related (but not the same). Several authors have discussed the similarities between doubled CA and PCA (for example, Leclerc (1980, p. 56), Greenacre (1984, p. 182–183) and Van de Velden (2004, p. 103–104)).

2.1. Three unfolding benchmark datasets

To be able to compare our results with the results from the different fields of research discussed in the introduction of this paper, we simulated data using the three models explained in Section 1.3. First, we generated a deterministic data set for each of the three unfolding models, used as benchmarks. Data generation was done as follows: we simulated responses for 300 persons on 20 items, which is an average sample size and an average test length.

We generated data following model 1 according to (1), with $a_j = -1$, and b_j equal to the maximum of $(x_i - y_j)^2$, so that the minimum z_{ij} equaled 0. The model 2 data were generated according to (3), with $t_j = 1$. Finally, the model 3 data were generated according to (7), with $t_j = 1$, and a constant interthreshold distance of .4 (these values are based on previous studies by Roberts et al. (2000)). To create data comparable to the data analyzed by Davison (1977), we sampled

true person scores x_i from a normal $N(0, 1)$ distribution and we chose item locations y_j equally spaced ranging from -3 to $+3$ on the latent continuum for the three benchmark datasets. For the probabilistic models (models 2 and 3) 500 data sets were generated, which were averaged and rounded.

Note that model 1 (the metric unfolding model) produces continuous data, model 2 (the Gaussian ordination model) produces dichotomous (i.e. 0/1) data, and model 3 (the IRT-unfolding model; GGUM) produces graded response data on a six-point rating scale.

In the current paper we estimate item location, and not category location, for all types of responses, even for graded responses. We regard the item categories as measures of association strength between an item and a person. In other words, we make the assumption (like one does in PCA) that on a scale from 0 to 5 a rating of 5 means this person prefers the item 5 times as much as a person with a rating of 1.

To compare the performance of the two types of CA and of PCA, of each solution only the first dimension was selected. Since the model-generated data are one-dimensional, the major target is that the first dimension of each solution shows the true order of items and persons. The quality of the recovery of person locations, item locations and the joint scale is expressed in terms of correlations.

Second, we performed a Monte Carlo simulation to include sampling error. The procedure is described in the next section.

2.2. Monte Carlo simulation

The aim of the simulation study is to compare the performance of CA (with and without doubled items) and PCA (with and without varimax rotation) in terms of variation of the parameter estimates.

True person scores were again sampled from a normal $N(0, 1)$ distribution and item locations were equally spaced, ranging from -3 to $+3$ on the latent continuum. Responses were simulated 300 times according to the three different models discussed previously. To simulate responses according to model 1 we used the variant defined by (2). We followed the same procedure as Davison (1977) by adding $\mathbf{E}_{ij}$, a random normal deviate, $N(0, .25)$ to the responses. Davison (1977) considered the error level of $\sigma^2_{\mathbf{E}_{ij}} = .25$ yielding a realistic reliability of responses on psychological variables. Models 2 and 3 result in probabilities for each response category. These probabilities were used to sample responses.

The quality of the recovery of person locations, item locations and the joint scale in terms of correlations for the 200 simulations will be summarized with boxplots.

2.3. Real data: The Developmental Profile

We also analyzed data on personality development collected with the Developmental Profile (DP) (Abraham et al., 2001). The DP is an instrument for personality assessment consisting of 10 subscales, referred to as developmental levels, each consisting of 9 items, referred to as developmental lines.

Each developmental level describes a central or specific aspect of behavior, characteristic of a specific phase in the development of psychosocial capacities. The developmental levels in the DP are organized in a hierarchy, according to the degree to which they are associated with the severity of maladaptive psychosocial functioning. The lower six levels refer to maladaptive behavior; the upper four levels refer to adaptive behavior. The developmental lines describe various categories of behavior as they are manifested on each of the developmental levels.

It is assumed that the 10 developmental levels may be seen as separate (but not independent) subscales consisting of 9 items (behavior patterns defined on 9 developmental lines). For each level these 9 items are manifestations of level specific functioning in 9 different domains. All items are scored by a trained professional based on a semi-structured interview. A four-point scale is used to indicate the degree to which each personality characteristic is present ($0 =$ not present; $1 =$ present to a limited degree; $2 =$ clearly present; $3 =$ very clearly present). The developmental profile of an individual is defined as his (sum-)score on each of the ten developmental levels.

In the current paper the aim is to investigate the hypothesis that the 10 developmental levels (i.e. the subscales of the DP) are ordered on one underlying bipolar dimension ranging from maladaptive to adaptive psychosocial functioning. The sample consisted of 736 patients who were classified as either forensic inpatients ($N = 24$), inpatients ($N = 450$), outpatients of ambulant ($N = 163$), and normal controls ($N = 99$). Level scores were computed as the sum of the scores on the 9 items corresponding to that level. Because there were only 0-scores on the items of the highest level, this level was discarded from all analysis, resulting in a 736 by 9 data matrix.

3. Results

3.1. The three benchmark datasets

This section of results consists of two parts. First the matrices of inter-item correlations for the three benchmark datasets are compared. Second, the results of the two types of CA are compared to the results of the two types of PCA.

The inter-item correlations for the benchmark data conforming to model 1 are displayed in Table 1. The correlation matrix shows a strong Robinson pattern. The inter-item correlations for the benchmark datasets conforming to model 2 and 3 are

Table 1
Correlations between items for the benchmark dataset conforming to model 1

Item	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14	v15	v16	v17	v18	v19	v20
v1	1.00	1.00	1.00	1.00	.99	.98	.96	.92	.80	.51	.00	−.45	−.67	−.78	−.84	−.88	−.90	−.91	−.92	−.93
v2	1.00	1.00	1.00	1.00	.99	.99	.97	.93	.81	.53	.02	−.43	−.66	−.77	−.83	−.87	−.89	−.90	−.92	−.92
v3	1.00	1.00	1.00	1.00	1.00	.99	.98	.94	.83	.55	.04	−.40	−.64	−.75	−.81	−.85	−.88	−.89	−.90	−.91
v4	1.00	1.00	1.00	1.00	1.00	.99	.98	.95	.85	.58	.08	−.37	−.61	−.73	−.79	−.83	−.86	−.88	−.89	−.90
v5	.99	.99	1.00	1.00	1.00	1.00	.99	.96	.87	.61	.12	−.33	−.58	−.70	−.77	−.81	−.84	−.85	−.87	−.88
v6	.98	.99	.99	.99	1.00	1.00	1.00	.97	.90	.66	.18	−.27	−.53	−.65	−.73	−.77	−.80	−.82	−.84	−.85
v7	.96	.97	.98	.98	.99	1.00	1.00	.99	.93	.72	.26	−.19	−.45	−.59	−.67	−.71	−.75	−.77	−.79	−.80
v8	.92	.93	.94	.95	.96	.97	.99	1.00	.97	.81	.39	−.05	−.32	−.47	−.56	−.61	−.65	−.67	−.70	−.71
v9	.80	.81	.83	.85	.87	.90	.93	.97	1.00	.92	.59	.18	−.10	−.26	.35	−.42	−.46	−.49	−.51	−.53
v10	.51	.53	.55	.58	.61	.66	.72	.81	.92	1.00	.86	.54	.29	.14	.04	−.03	−.08	−.11	−.14	−.16
v11	.00	.02	.04	.08	.12	.18	.26	.39	.59	.86	1.00	.90	.74	.63	.54	.49	.44	.41	.39	.37
v12	−.45	−.43	−.40	−.37	−.33	−.27	−.19	−.05	.18	.54	.90	1.00	.96	.91	.86	.82	.79	.77	.75	.74
v13	−.67	−.66	−.64	−.61	−.58	−.53	−.45	−.32	−.10	.29	.74	.96	1.00	.99	.97	.95	.93	.92	.90	.89
v14	−.78	−.77	−.75	−.73	−.70	−.65	−.59	−.47	−.26	.14	.63	.91	.99	1.00	.99	.99	.98	.97	.96	.95
v15	−.84	−.83	−.81	−.79	−.77	−.73	−.67	−.56	−.35	.04	.54	.86	.97	.99	1.00	1.00	.99	.99	.98	.98
v16	−.88	−.87	−.85	−.83	−.81	−.77	−.71	−.61	−.42	−.03	.49	.82	.95	.99	1.00	1.00	1.00	1.00	.99	.99
v17	−.90	−.89	−.88	−.86	−.84	−.80	−.75	−.65	−.46	−.08	.44	.79	.93	.98	.99	1.00	1.00	1.00	1.00	1.00
v18	−.91	−.90	−.89	−.88	−.85	−.82	−.77	−.67	−.49	−.11	.41	.77	.92	.97	.99	1.00	1.00	1.00	1.00	1.00
v19	−.92	−.92	−.90	−.89	−.87	−.84	−.79	−.70	−.51	−.14	.39	.75	.90	.96	.98	.99	1.00	1.00	1.00	1.00
v20	−.93	−.92	−.91	−.90	−.88	−.85	−.80	−.71	−.53	−.16	.37	.74	.89	.95	.98	.99	1.00	1.00	1.00	1.00

Table 2
Correlations between items for the benchmark dataset conforming to model 3

Item	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14	v15	v16	v17	v18	v19	v20
v1	1.00	.87	.82	.79	.72	.64	.49	.29	.00	−.26	−.59	−.74	−.81	−.73	−.66	−.63	−.57	−.48	−.40	−.33
v2	.87	1.00	.89	.85	.82	.72	.62	.42	.08	−.21	−.57	−.79	−.83	−.84	−.76	−.68	−.65	−.58	−.48	−.40
v3	.82	.89	1.00	.88	.84	.80	.68	.54	.22	−.12	−.47	−.72	−.87	−.87	−.87	−.75	−.69	−.68	−.59	−.49
v4	.79	.85	.88	1.00	.90	.86	.76	.62	.35	.03	−.42	−.65	−.81	−.89	−.88	−.86	−.77	−.69	−.65	−.60
v5	.72	.82	.84	.90	1.00	.89	.83	.73	.47	.18	−.28	−.60	−.77	−.88	−.91	−.88	−.86	−.77	−.70	−.67
v6	.64	.72	.80	.86	.89	1.00	.89	.82	.61	.31	−.14	−.47	−.76	−.84	−.89	−.91	−.88	−.87	−.77	−.69
v7	.49	.62	.68	.76	.83	.89	1.00	.87	.73	.48	.03	−.32	−.61	−.82	−.83	−.86	−.88	−.87	−.85	−.75
v8	.29	.42	.54	.62	.73	.82	.87	1.00	.83	.64	.22	−.13	−.46	−.67	−.79	−.80	−.83	−.86	−.83	−.82
v9	.00	.08	.22	.35	.47	.61	.73	.83	1.00	.82	.55	.22	−.14	−.40	−.54	−.66	−.68	−.73	−.78	−.77
v10	−.26	−.21	−.12	.03	.18	.31	.48	.64	.82	1.00	.74	.51	.16	−.09	−.25	−.39	−.49	−.52	−.58	−.63
v11	−.59	−.57	−.47	−.42	−.28	−.14	.03	.22	.55	.74	1.00	.79	.61	.36	.19	.05	−.06	−.19	−.24	−.33
v12	−.74	−.79	−.72	−.65	−.60	−.47	−.32	−.13	.22	.51	.79	1.00	.81	.70	.54	.38	.29	.20	.06	−.01
v13	−.81	−.83	−.87	−.81	−.77	−.76	−.61	−.46	−.14	.16	.61	.81	1.00	.85	.78	.67	.58	.52	.41	.29
v14	−.73	−.84	−.87	−.89	−.88	−.84	−.82	−.67	−.40	−.09	.36	.70	.85	1.00	.90	.84	.78	.72	.65	.57
v15	−.66	−.76	−.87	−.88	−.91	−.89	−.83	−.79	−.54	−.25	.19	.54	.78	.90	1.00	.89	.85	.82	.76	.71
v16	−.63	−.68	−.75	−.86	−.88	−.91	−.86	−.80	−.66	−.39	.05	.38	.67	.84	.89	1.00	.90	.86	.83	.79
v17	−.57	−.65	−.69	−.77	−.86	−.88	−.88	−.83	−.68	−.49	−.06	.29	.58	.78	.85	.90	1.00	.89	.85	.82
v18	−.48	−.58	−.68	−.69	−.77	−.87	−.87	−.86	−.73	−.52	−.19	.20	.52	.72	.82	.86	.89	1.00	.88	.83
v19	−.40	−.48	−.59	−.65	−.70	−.77	−.85	−.83	−.78	−.58	−.24	.06	.41	.65	.76	.83	.85	.88	1.00	.86
v20	−.33	−.40	−.49	−.60	−.67	−.69	−.75	−.82	−.77	−.63	−.33	−.01	.29	.57	.71	.79	.82	.83	.86	1.00

similar with respect to their pattern, so only the correlation matrix from model 3 is displayed in Table 2. The correlation matrix shows a distinctively different pattern. Namely, the correlations show a pattern similar to the Robinson pattern, but with "inwardly bending extremes". That is, the correlations between the most extreme items are not the highest negative correlations, like in Table 1, but are only moderately negative.

The performance of the two types of CA and PCA for the three unfolding benchmark datasets in terms of recovery of the true parameters for persons, items, and the joint scale is quantified by the correlations between true and estimated parameter values, which are reported in Table 3.

Table 3 shows that, for all models, CA recovers all scale values well ($r > .99$). The only exception is the relatively poor recovery of the joint scale by CA for the model 1 data. It appeared that for this type of data a symmetrical normalization (i.e. the standard coordinates of both items and persons multiplied by the square root of the singular values) of the CA scale values results in a better recovery of the joint scale values ($r = .98$), whereas the recovery of the joint scale of the symmetrically scaled PCA did not improve as much ($r = .94$). However, the CA (with row principal normalization) on the doubled data gives the best recovery of the joint scale for the model 1 data. We elaborate this matter in the Discussion section.

Furthermore, it can be seen that the pattern of results for the model 2 data and the model 3 data is very similar. For the model 2 and model 3 data, undoubled CA gives the best recovery of all scale values, except that for the model 2 data CA of the doubled items gives a slightly better recovery.

Table 3
Parameter recovery for various analysis techniques; correlations between true and estimated parameter values for items, persons, and the joint scale of items and persons

		CA	CAd	PCA	PCAr
Model 1	Items	.9936	.9938	1.0000	.9998
	Persons	.9958	.9995	.9989	.9432
	Joint	.8169	.9975	.9033	.8529
Model 2	Items	.9966	.9976	.7200	.5072
	Persons	.9950	.9764	.9548	.8253
	joint	.9946	.9764	.8642	.7463
Model 3	Items	.9996	.9991	.8708	.6406
	Persons	.9975	.9774	.9751	.8751
	Joint	.9953	.9733	.8852	.7944

(a) True person scale values. (b) True person scale values.

(c) True person scale values. (d) True person scale values.
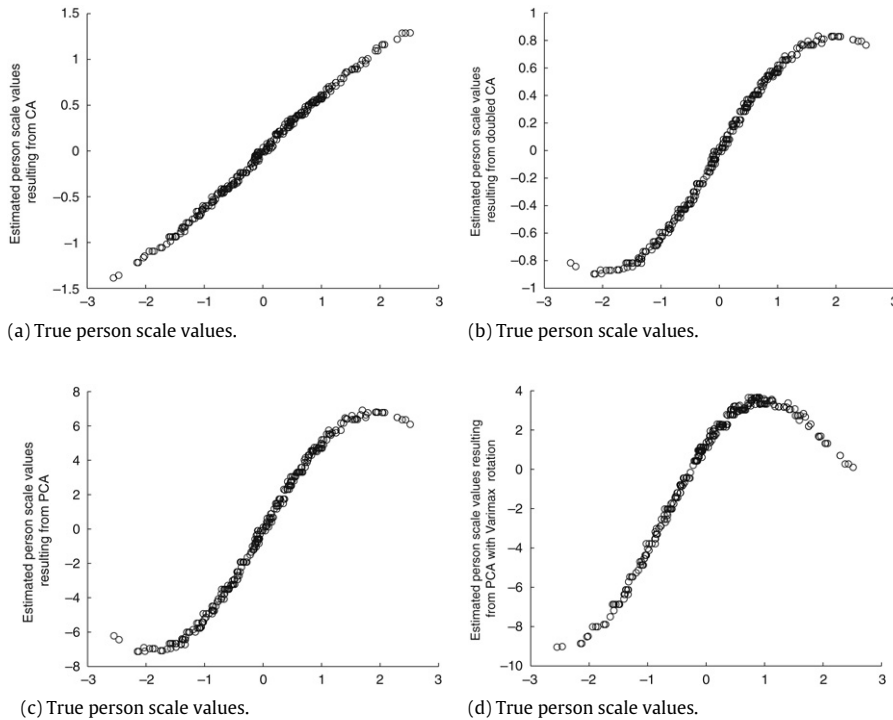
Fig. 1. True persons scale values plotted against the estimated scale values of CA (a), CA with doubled items (b), PCA (c), and PCA with varimax rotation (d) performed on the model 3 benchmark data.

For the model 1 data, the component loadings resulting from the PCA perfectly match the true item scale values. For the person scale values and the joint scale, CA of the doubled items gives the best recovery.

Although in some cases, for example for the person scores of the CA with doubled items or PCA on the model 3 data in Table 3, the correlation suggests a good fit ($r = .977$ and $r = .975$, respectively), there is a serious misfit in the extreme ends of the scale. This effect is due to the inward bending of the scores of the extreme persons illustrated in Fig. 1, which gives the scatter plots with the true persons scale values plotted against the estimated scales values of the four types of analysis. It is clear that only the person scale values estimated by CA are linearly related to the true scale values and are ordered correctly ($r = .998$).

3.2. Monte Carlo simulation

Fig. 2 shows the boxplots for the correlations between "true" and estimated parameter values for items, persons, and the joint scale of items and persons, for CA, CAd, PCA, and PCAr on 200 simulated data sets from 3 unfolding models. The overall pattern of results is very similar to the results presented in Table 3. Again the results for the model 2 data and the model 3 data are very similar. The boxplots for the model 2 and model 3 data show that CA gives the best recovery of all scale values.

Both the doubled CA and PCA estimations of the person and item scale values show the pattern with the inwardly bending endpoints, thereby underestimating the scale values of the extreme persons and items. The undoubled CA estimations the
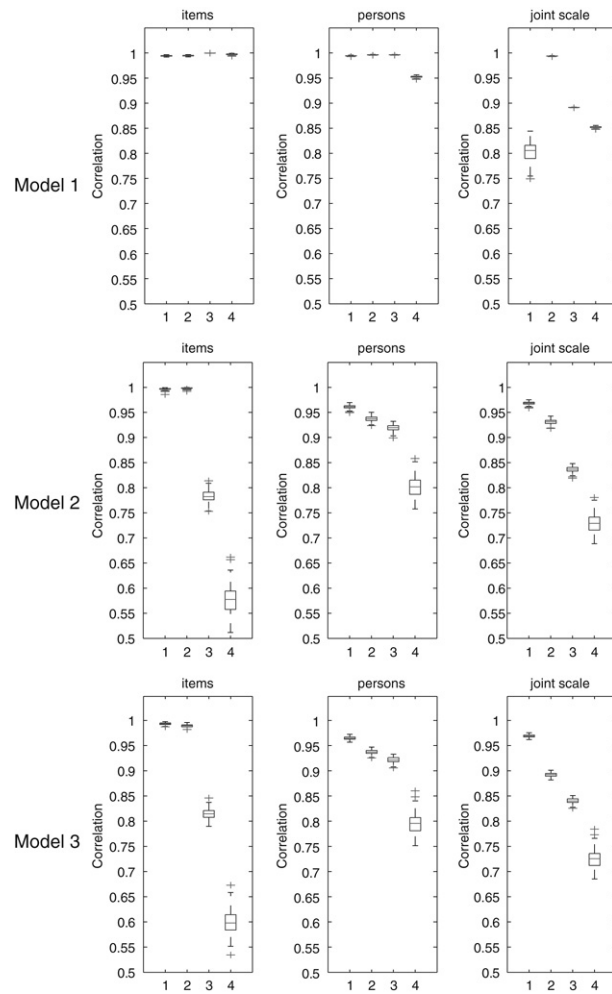
Fig. 2. Parameter recovery: boxplots of correlations between true and estimated parameter values for items, persons, and the joint scale based on CA (1), CA with doubled items (2), PCA (3), and PCA with varimax rotation (4) on 200 simulated data sets resulting from the 3 unfolding models.

person and item scale values show the arch-pattern, with a correct order along the first dimension. However, the difference between the recovery of item scale values in CA and CA with doubled items for the model 2 data are negligible.

For the model 1 data, PCA always gives a perfect recovery of the item scale values, although the differences with the CA techniques are minimal. Except for PCA with varimax rotation, all techniques recover the person scale values well for the model 1 data. For the joint scale CA with doubled items gives the best recovery. CA with row principal normalization gives a variable and poor recovery of the joint scale. This phenomenon will be further explained in the Discussion.

3.3. Real data: The Developmental Profile

Based on the results of the simulation study we chose to analyze the data with only CA and PCA without rotation. The inter-item correlations for the DP data are displayed in Table 4. The correlation matrix shows a pattern similar to the pattern in Table 2, although there are some reversals, for example in the column of v3 the order of v4, v5, and v6 is reversed. However, the pattern suggests that the PCA solution of for this data will possibly suffer from inward bending component loadings and component scores, besides the extra-component problem.

Fig. 3 shows the unrotated two-component solution of PCA on the DP data. It can be seen that there is indeed an inward bending of the component loadings of the extreme items. The PCA solution accounts for 47.4% of the total variance (32.4 and 14.9 by axis 1 and 2, respectively). The CA solution in 2 dimensions is displayed in Fig. 4. This solution shows the well-known arch-effect (see, for example, Hill and Gauch (1980)), with the theorized order of the items reflected by the first axis. There is a reversal of items v2 and v3 along the arch (this was also the case on the first component of the PCA solution), but not on the first axis. The CA solution accounts for 55.0% of the total inertia (37.3% and 17.7% for axes 1 and 2, respectively).

There are two important differences between the two solutions, that could lead to fundamentally different interpretations. First, the inward bending of the factor loadings of the extreme items (v1 and v9) could lead to the conclusion

Table 4
Correlations between level scores of the Developmental Profile

Item	v1	v2	v3	v4	v5	v6	v7	v8	v9
v1	1.00	.48	.27	.11	.00	.00	−.37	−.28	−.19
v2	.48	1.00	.35	.31	.13	.08	−.37	−.27	−.24
v3	.27	.35	1.00	.05	.11	.21	−.21	−.13	−.15
v4	.11	.31	.05	1.00	.32	.11	−.24	−.26	−.28
v5	.00	.13	.11	.32	1.00	.26	−.11	−.15	−.17
v6	.00	.08	.21	.11	.26	1.00	.01	−.01	−.10
v7	−.37	−.37	−.21	−.24	−.11	.01	1.00	.59	.51
v8	−.28	−.27	−.13	−.26	−.15	−.01	.59	1.00	.48
v9	−.19	−.24	−.15	−.28	−.17	−.10	.51	.48	1.00
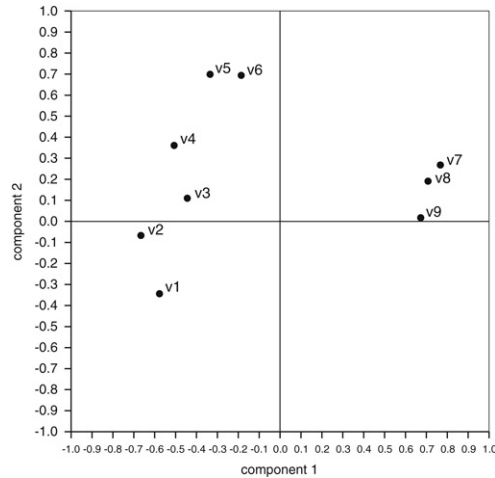
Fig. 3. Component loadings on the first and second component resulting from PCA on the DP data.
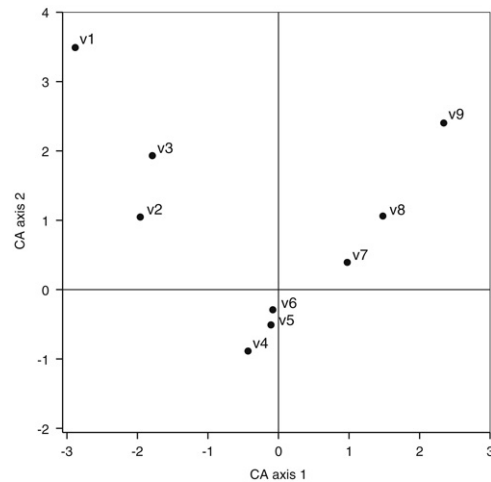

Fig. 4. Item scores on the first and second dimension resulting from CA on the DP data.

that these items do not "fit" into the scale as well as the moderately extreme items (v2 and v8). Second, v5 and v6 would clearly be discarded from the first subscale and would be seen as a separate (independent) subscale. In the CA literature it is known that the arch-effect is a result of a strong first dimension. In practice the CA solution depicted in Fig. 4 would be interpreted as a confirmation of the theory of an underlying bipolar axis on which the 9 items are ordered.

4. Discussion

Across all analyses, CA without doubling performs best for unfolding data generated with three different single-peaked models. We have to make one reservation however.

Both the analysis results for the three unfolding benchmark datasets and the results of the simulation study showed that in the case of the model 1 data CA recovered the joint scale poorly, whereas CA of the doubled data recovered the joint scale well. This is an exception in the current and existing results referred to in this paper. It was suggested that for this model a symmetrical normalization is to be preferred over a row principal normalization. This matter deserves further attention, but in the following we try to explain the difference in results of row principal and symmetrical CA.

The reason for the poor correlation between the true joint scale and the joint scale estimated by CA with row principal normalization is that the ranges of the person scale values and the item scale values differ substantially. CA on the simulated data based on model 1 resulted in person scale values ranging from −.73 to .57, while the item scale values ranged from −3.27 to 2.45. By choosing a symmetrical normalization these scale differences became smaller (in the case of symmetrical normalization the person scale values range from −2.02 to 1.79, while the item scale values range from −0.94 to 0.85).

Note that for both model 2 and model 3 CA recovered the joint scale very well. For instance, for the model 3 data the correlations between the true joint scale and the estimated joint scale are substantially higher for CA than for the other methods. CA on the simulated data based on model 3 resulted in person scale values ranging from −1.81 to 1.77, while the item scale values range from −2.52 to 2.43. Clearly, in this case both sets of scale values have a more comparable scale than in the case of the model 1 data.

An explanation for this difference can be found in the row masses of the datasets resulting from the different models. In the case of the model 1 data, the row masses show relatively little variation (on a scale with an average mass of one they range from .6 to 1.05, whereas in the case of the model 3 data the row masses range from .15 to 1.53). This difference is explained by the difference in scores of the extreme persons. Persons with a relative extreme position have their maximum score on the most extreme item on the corresponding pole of the scale. For these persons the scores on the other items decrease as the items lie further away from the extreme of the scale. In model 3 this decrease is steeper than in model 1, resulting in lower masses for these extreme persons. Since in CA the inverse masses serve as weights in determining the inter-point distances, the extreme persons in the model 3 data lie further away from the origin than for the model 1 data.

Note that the "symmetrical normalization" referred to in the above, (which is the SPSS variant) is not the same as the "symmetric map", which is often used by French researchers (see also Greenacre (1993; 2007, Chapter 9)). In "symmetrical normalization", the square root of the singular values is used as a scaling factor for the standard coordinates of both items and persons. In contrast, the "symmetric map" uses the singular values (and not their square root) as a scaling factor. For a discussion about the difference between these two scaling options, see Greenacre (2006). In this paper we are interested in the correlation between the true scale values and the CA scale values of the first axis, so that the scaling differences do not influence the current results.

It was shown that two types of unfolding models, i.e. models that are either a quadratic function of the person-to-item distances or an exponential function of these distances can easily be recognized empirically, because the inter-item correlation matrices for the two types of data typically show different patterns. Data conforming to the quadratic model show a Robinson pattern, whereas data conforming to the exponential model show a pattern similar to the Robinson pattern, but with "inwardly bending extremes".

Particularly for this latter type of responses, we showed that CA of the raw data matrix outperforms both PCA (with or without rotation) and CA with doubling, with respect to scaling items and persons on the underlying unfolding scale. It turned out, that for this type of responses, the two-dimensional CA solution was not hampered by the inward bending of the extremes, whereas both the CA solution for the doubled data, and the PCA solution were hampered by this phenomenon.

Acknowledgement

References

Abraham, R.E., Van, H.L., Van Foeken, I., Ingenhoven, T.J.M., Tremonti, W., Pieper de Vries, I., De Groot, J., Van't Spijker, A., Spinhoven, P., 2001. The developmental profile. Journal of Personality Disorders 15, 457–473.

Andrich, D., 1996. A hyperbolic cosine latent trait model for unfolding polytomous responses: Reconciling Thurstone and Likert methodologies. British Journal of Mathematical and Statistical Psychology 49, 347–365.

Andrich, D., Styles, I.M., 1998. The structural relationship between attitude and behavior statements from the unfolding perspective. Psychological Methods 3, 454–469.

Benzécri, J.-P., 1992. Correspondence Analysis Handbook. Marcel Dekker, New York.

Coombs, C.H., 1964. A Theory of Data. Wiley, New York.

Coombs, C.H., Kao, R.C., 1960. On a connection between factor analysis and multidimensional unfolding. Psychometrika 25, 219–231.

Davison, M.L., 1977. On a metric unidimensional unfolding model for attitudinal and developmental data. Psychometrika 42, 523–548.

De'ath, G., 1999. Principal curves: A technique for indirect and direct gradient analysis. Ecology 80, 2237–2253.

DeSarbo, W.S., Kim, J., Chan Choi, S., Spaulding, M., 2002. A gravity based multidimensional scaling model for deriving spatial structures underlying consumer preference/choice judgements. Journal of Consumer Research 29, 91–100.

Gifi, A., 1990. Nonlinear Multivariate Analysis. Wiley, Chicester.

Greenacre, M.J., 1984. Theory and Applications of Correspondence Analysis. Academic Press, New York.

Greenacre, M.J., 1993. Correspondence Analysis in Practice. Academic Press, London.

Greenacre, M.J., 2006. Tying up the loose end in simple correspondence analysis. Working paper.

Greenacre, M.J., 2007. Correspondence Analysis in Practice, second edition. Chapman & Hall/CRC, London.

Heiser, W.J., 1981. Unfolding analysis of proximity data. Ph.D. Thesis, Leiden, The Netherlands, Leiden University.

Heiser, W.J., 2001. Correspondence analysis. In: Smelser, N.J., Baltes, P.B. (Eds.), International Encyclopedia of the Social and Behavioral Sciences. Pergamon, Oxford, pp. 2820–2824.

Hill, M.O., 1973. Reciprocal averaging: An eigenvector method of ordination. Journal of Ecology 61, 237–249.

Hill, M.O., 1974. Correspondence analysis: A neglected multivariate method. Applied Statistics 23, 340–354.

Hill, M.O., Gauch, H.G., 1980. Detrended correspondence analysis: An improved ordination technique. Vegetatio 42, 47–58.

Hubert, L., Arabie, P., Meulman, J., 1998. The representation of symmetric proximity data: Dimensions and classifications. The Computer Journal 41, 566–577.

Ihm, P., van Groenewoud, H., 1984. Correspondence analysis and Gaussian ordination. COMPSTAT Lectures 3, 5–60.

Kendall, M.G., 1971. Seriation from abundance matrices. In: Hodson, F.R. (Ed.), Mathematics in the Archeological and Historical Sciences. Edinburgh University Press, Edinburgh, pp. 215–252.

Leclerc, A., 1980. Quelques propriétés optimales en analyse des données, en termes de corrélation entre variables. Mathématiques et Sciences Humaines 70, 51–67.

Maraun, M.D., Rossi, N.T., 2001. The extra-factor phenomenon revisited: Unidimensional unfolding as quadratic factor analysis. Applied Psychological Measurement 25, 77–87.

Meulman, J.J., 1982. Homogeneity Analysis of Incomplete Data. DSWO Press, Leiden, The Netherlands.

Noy-Meir, I., Austin, M.P., 1970. Principal component ordination and simulated vegetational data. Ecology 51, 551–552.

Roberts, J.S., Donoghue, J.R., Laughlin, J.E., 2000. A general item response theory model for unfolding unidimensional polytomous responses. Applied Psychological Measurement 24, 3–32.

Roberts, J.S., Laughlin, J.E., Wedell, D.H., 1999. Validity issues in the Likert and Thurstone approaches to attitude measurement. Educational and Psychological Measurement 59, 211–233.

Ross, J., Cliff, N., 1964. A generalization of the interpoint distance model. Psychometrika 29, 167–176.

Rost, J., Luo, G., 1997. An application of a Rasch-based unfolding model to a questionnaire on adolescent centrism. In: Rost, J., Rolf, L. (Eds.), Applications of Latent Trait and Latent Class Models in the Social Sciences. Waxmann Munster, New York.

Schoenemann, P.H., 1970. On metric multidimensional unfolding. Psychometrika 35, 349–366.

Swan, J.M.A., 1970. An examination of some ordination problems by use of simulated vegetational data. Ecology 51, 89–102.

Tabachnick, B.G., Fidell, L.S., 2001. Using Multivariate Statistics. Allyn & Bacon, Boston, MA.

Ter Braak, C.J.F., 1985. Correspondence analysis of incidence and abundance data: Properties in terms of a unimodal response model. Biometrics 41, 859–873.

Ter Braak, C.J.F., 1987. Unimodal models to relate species to environment. Ph.D. Thesis, Wageningen University, The Netherlands.

Ter Braak, C.J.F., 1988. Partial canonical correspondence analysis. In: Bock, H.H. (Ed.), Classification and Related Methods of Data Analysis. North-Holland, Amsterdam, pp. 551–558.

Ter Braak, C.J.F., Prentice, I.C., 1988. A theory of gradient analysis. Advances in Ecological Research 18, 271–317.

Ter Braak, C.J.F., Prentice, I.C., 2004. A theory of gradient analysis. Advances in Ecological Research 34, 235–282.

Van de Velden, M., 2004. Optimal scaling of paired comparison data. Journal of Classification 21, 89–109.

Van Schuur, W.H., Kiers, H.A.L., 1994. Why factor analysis often is the wrong model for analysing bipolar concepts and what model to use instead. Applied Psychological Measurement 18, 97–110.

Van Schuur, W.H., Kruijtbosch, M., 1995. Measuring subjective well-being: Unfolding the Bradburn affect balance scale. Social Indicators Research 36, 49–74.

Zhu, M., Hastie, T.J., Walther, G., 2005. Constrained ordination analysis with flexible response functions. Ecological Modelling 187, 524–536.