

Symmetrical and Non-symmetrical Variants of Three-Way Correspondence Analysis for Ordered Variables

Rosaria Lombardo, Eric J. Beh and Pieter M. Kroonenberg

Abstract. In the framework of multi-way data analysis, this paper presents symmetrical and non-symmetrical variants of three-way correspondence analysis that are suitable when a three-way contingency table is constructed from ordinal variables. In particular, such variables may be modelled using general recurrence formulae to generate orthogonal polynomial vectors instead of singular vectors coming from one of the possible three-way extensions of the singular value decomposition. As we shall see, these polynomials, that until now have been used to decompose two-way contingency tables with ordered variables, also constitute an alternative orthogonal basis for modelling symmetrical, non-symmetrical associations and predictabilities in three-way contingency tables. Consequences with respect to modelling and graphing will be highlighted.

Key words and phrases: Symmetrical and non-symmetrical three-way correspondence analysis, ordinal categorical variables, orthogonal polynomials, trivariate moment decomposition.

1. INTRODUCTION

In the literature of multi-way data analysis, the basic multi-way techniques have been mainly applied to continuous data (Kroonenberg, 2008, 2014, Smilde, Bro and Geladi, 2004). Here, we aim to deal with multi-way categorical data. The study of multi-way contingency tables takes place via the analysis of lower and higher-order associations (interactions) among variables. For this purpose, association models (Clogg, 1982, van der Heijden, de Falguerolles and de Leeuw, 1989, Kroonenberg and Anderson, 2006, Agresti and Gottard, 2007, Kateri, 2014) and partitions of the three-way chi-squared statistic have been proposed in some detail (Lancaster, 1951,

Choulakian, 1988, Beh and Davy, 1998, 1999, Loisel and Takane, 2016, Lombardo, Takane and Beh, 2020).

Carlier and Kroonenberg (1996, 1998) and Kroonenberg (2008, Chapter 17) developed a comprehensive approach to deal with three-way categorical data, that is, three-way correspondence analysis, which includes the Lancaster partitioning of the chi-squared (and phi-squared) statistic (Lancaster, 1951), modelling of the global dependence in the three-way table and graphing of the global dependence via three-way correspondence analysis. Rather than using three-way correspondence analysis, it is more usual to examine sets of two-way tables and employ multiple correspondence analysis to an indicator matrix or a Burt matrix (Benzécri, 1973, 1977, Greenacre, 1984, Chapter 5, Greenacre, 1990, 2017, Greenacre and Blasius, 2006), thereby ignoring higher-order associations among variables (Beh and Lombardo, 2019).

Based on earlier versions of the papers by Carlier and Kroonenberg (1996, 1998), Lombardo (1994) discussed a variant of three-way correspondence analysis, called *non-symmetrical three-way correspondence analysis* for the analysis of a dependence relationship among the categorical variables, published as Lombardo, Carlier and D’Ambra (1996). Non-symmetrical three-way correspondence analysis is a generalisation of the two-way technique introduced by Lauro and D’Ambra (1984) and

Rosaria Lombardo is Associate Professor of Statistics, Economics Department, University of Campania L. Vanvitelli, via Gran Priorato di Malta, Capua (CE), Italy (e-mail: rosaria.lombardo@unicampania.it). Eric J. Beh is Professor of Statistics and Head of Statistics discipline, School of Mathematical and Physical Sciences, University of Newcastle, Callaghan, 2308, New South Wales, Australia (e-mail: eric.beh@newcastle.edu.au). Pieter M. Kroonenberg is Emeritus Professor of Multivariate Data Analysis, The Three-Mode Company, University of Leiden, Wasstraat 11, 2313 JG Leiden, The Netherlands (e-mail: kroonenb@fsw.leidenuniv.nl).

D'Ambra and Lauro (1989), in which a dependence structure is assumed such that the row variable is predicted by the column variable, or vice versa. A practical description of two-way non-symmetrical correspondence analysis is given by Kroonenberg and Lombardo (1999).

1.1 Three-Way Correspondence Analysis for Ordinal Data

In this paper, we propose variants of symmetrical and non-symmetrical three-way correspondence analysis to analyse ordinal three-way contingency tables, whereby an 'ordinal table' means that the three variables have ordinal categories. These variants use a novel decomposition, the *trivariate moment decomposition*, which has orthogonal polynomials as its basis. In contrast to much of the literature that deals with stochastically modelling ordered categorical data, but in line with the *algorithmic modeling culture* (Breiman, 2001), orthogonal polynomials are used for modelling the association among ordinal variables, including possible asymmetrical (or non-symmetrical) associations that exists among their categories (Beh and Lombardo, 2014, pages 91–100).

In order to analyse the association between the rows, columns and tubes¹ of a three-way contingency table, various extensions or generalisations of the singular value decomposition have been proposed; see, for example, Tucker (1966), Harshman (1970), Carroll and Chang (1970), Kiers (1989) and Kroonenberg (1992, 2008, 2020). In this paper, we focus on the Tucker3 decomposition (Tucker, 1966, Kroonenberg, 1983, 2008, Smilde, Bro and Geladi, 2004) which can be used to derive three orthogonal sets of singular vectors (components), one for each way of the data array. The Tucker3 decomposition also consists of a three-way core array which contains the strengths of the links between these components. However, the Tucker3 components are not the only type of basis vectors that can be defined for component spaces. Orthogonal polynomials can serve as an alternative set of orthogonal components which span each of the spaces of the three ways. For high-dimensional problems, polynomial decompositions can be comparable to non-parametric analysis of variance models (Lin and Zhang, 2006, Wahba, 1990, Wahba et al., 1995). As will be discussed in detail in Section 3, the decomposition based on orthogonal polynomials is generally called the *trivariate moment decomposition*. When among the three variables, at least one is nominal, we define a *hybrid decomposition* that uses polynomial components for the ordered categorical variables and Tucker3 components for the nominal variables.

The main aim of this paper is to therefore acquire insight into the structure of variable dependence taking into account the ordered information of categorical variables.

As an example, we will examine the study of Van Herk and Van de Velden (2007) where three-way correspondence analysis is applied to study ratings and rankings of items in five European countries, but replacing the singular vectors of two ways by orthogonal polynomials. The findings of Van Herk and Van de Velden (2007) show that ratings and rankings are assessed quite similarly by people from different countries. Here, we show that the orthogonal polynomial components can serve to evaluate differences that can exist between the ratings and the rankings across the countries.

1.2 Organisation of the Paper

This paper is organised as follows. In Section 2, we provide a summary of basic three-way correspondence analysis, both its symmetrical and non-symmetrical variants. In Section 3, we describe how the basic correspondence models can be adapted to three variables with ordinal categories. In this case, we will primarily concentrate on the symmetrical variant in which the multidimensional spaces have orthogonal polynomials as their support. The symmetrical variant for unordered variables has already been described in detail in Carlier and Kroonenberg (1996) and part of their basic results can also be used for handling ordered variables in the non-symmetrical case. In Section 4, we will illustrate the technique and its interpretational aspects using the rating-ranking data previously analysed by Van Herk and Van de Velden (2007).

2. SYMMETRICAL AND NON-SYMMETRICAL THREE-WAY CORRESPONDENCE ANALYSIS

Symmetrical three-way correspondence analysis is conceptually a straightforward generalisation of two-way symmetrical correspondence analysis. While the singular value decomposition is the core of the two-way analysis, a variant of the three-way generalisation of the singular value decomposition is the central element of three-way correspondence analysis. In practice, Tucker's (1966) three-mode principal component analysis is the variant that is most commonly used and it is the one used here; see Carlier and Kroonenberg (1996, Section 2.5) and Kroonenberg (2008, Chapter 17), for further details.

2.1 Non-symmetrical Three-Way Correspondence Analysis

Suppose we have three categorical variables observed on a sample of n subjects that are cross-classified to form a three-way contingency table. When one of the variables of a three-way contingency table, say the row variable, is a response variable predicted by the other two variables, then we speak of non-symmetrical three-way correspondence analysis (Lombardo, 1994, Lombardo, Carlier and D'Ambra, 1996, Beh and Lombardo, 2014, Chapter 11). In this case, it is not the deviation from independence of

¹A tube is a k -vector x_{ij} ; the third-way analogue of a row or column.

the variables which is the focus of the analysis, but the increase in predictability of the row categories given the column and tube categories. The increase can be measured by using the *Marcotorchino index* τ_M (Marcotorchino, 1985) which is a three-way generalisation of the two-way τ index of Goodman and Kruskal (1954), used in two-way non-symmetrical analysis (Kroonenberg and Lombardo, 1999).

2.2 The Marcotorchino Index and the Pearson Statistic

To discuss the Marcotorchino index as well as Pearson’s Φ^2 statistic, we will use the standardised notation and terminology for three-way analysis (Kiers, 2000). Let $\underline{P} = \{p_{ijk}\}$ be the $I \times J \times K$ array² of the observed joint relative frequencies, so that $\sum_{ijk} n_{ijk}/n = \sum_{ijk} p_{ijk} = 1$ with n_{ijk} being the absolute frequencies. For columns and tubes in symmetrical and non-symmetrical three-way correspondence analysis, $\underline{D}_J = \{d_{\bullet j \bullet} = p_{\bullet j \bullet}\}$, $\underline{D}_K = \{d_{\bullet \bullet k} = p_{\bullet \bullet k}\}$ are the diagonal weight matrices containing the relative marginal frequencies of the column and tube variables, respectively. For the rows in the symmetrical variant, the diagonal weight matrix is $\underline{D}_I = \{d_{i \bullet \bullet} = p_{i \bullet \bullet}\}$, but in the non-symmetrical case $\underline{D}_I = \{d_{i \bullet \bullet} = 1\}$, that is, the diagonal weights are one, rather than $p_{i \bullet \bullet}$ (Marcotorchino, 1985).

Based on these definitions the deviation from the three-way independence model is defined as

$$\underline{\Pi}_S = \left\{ \pi_{ijk}^S = \frac{p_{ijk}}{p_{i \bullet \bullet} p_{\bullet j \bullet} p_{\bullet \bullet k}} - 1 \right\}$$

while the expression for the column-tube centered profiles in a three-way table with a dependence structure becomes

$$\underline{\Pi}_M = \left\{ \pi_{ijk}^M = \frac{p_{ijk}}{p_{\bullet j \bullet} p_{\bullet \bullet k}} - p_{i \bullet \bullet} \right\}.$$

Therefore, π_{ijk}^M indicates the difference between the value of category i , predicted based on the knowledge of the combined (or interactively coded) predictor category (j, k) , $\frac{p_{ijk}}{p_{\bullet j \bullet} p_{\bullet \bullet k}}$, and the marginal row proportion of category i , $p_{i \bullet \bullet}$.

The inner product of the three-way array $\underline{\Pi}_M = (\pi_{ijk}^M)$ with itself in $\mathfrak{R}^{I \times J \times K}$ is equal to the squared norm or squared length of the array and, taking into account the diagonal weight matrices $\underline{D}_I = \{d_{i \bullet \bullet} = 1\}$, \underline{D}_J and \underline{D}_K , it can be defined as

$$\begin{aligned} \|\underline{\Pi}_M\|^2 &= \langle \underline{\Pi}_M, \underline{\Pi}_M \rangle \\ (1) \quad &= \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K p_{\bullet j \bullet} p_{\bullet \bullet k} (\pi_{ijk}^M)^2. \end{aligned}$$

Equation (1) is the numerator of the Marcotorchino statistic (1985), τ_M , which is a relative measure of predictability of the rows given the columns and tubes so that

$$(2) \quad \tau_M = \frac{\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K p_{\bullet j \bullet} p_{\bullet \bullet k} (\pi_{ijk}^M)^2}{1 - \sum_{i=1}^I p_{i \bullet \bullet}^2} = \frac{\tau_{M_{num}}}{\tau_{M_{den}}}.$$

Observe that the denominator is equivalent to the heterogeneity index (Gini, 1912) and measures the overall error in prediction and does not depend on the predictor categories. For this reason, the denominator can be neglected when analysing the predictability of a row given the column and tube categories (Beh and Lombardo, 2014, pages 459–461).

For symmetrical three-way correspondence analysis, the comparable measure is Pearson’s Φ^2 statistic in which, $\underline{D}_I = \{d_{i \bullet \bullet} = p_{i \bullet \bullet}\}$ rather than $\underline{D}_I = \{d_{i \bullet \bullet} = 1\}$. This measure of deviation from independence may be written as

$$(3) \quad \Phi^2 = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K p_{i \bullet \bullet} p_{\bullet j \bullet} p_{\bullet \bullet k} \left(\frac{p_{ijk}}{p_{i \bullet \bullet} p_{\bullet j \bullet} p_{\bullet \bullet k}} - 1 \right)^2.$$

2.3 Partitioning the Pearson and Marcotorchino Indices

The orthogonal projections of $\underline{\Pi}_S$ and $\underline{\Pi}_M$ onto the subspaces $\mathfrak{R}^0, \mathfrak{R}^I, \mathfrak{R}^J, \mathfrak{R}^K, \mathfrak{R}^{IJ}, \mathfrak{R}^{IK}, \mathfrak{R}^{JK}$ and \mathfrak{R}^{IJK} can be written in an ANOVA-like fashion (Lancaster, 1951, Choulakian, 1988, Carlier and Kroonenberg, 1996, Lombardo, Carlier and D’Ambra, 1996). For simplicity, a generic element of the space $\mathfrak{R}^{I \times J \times K}$ will be written in ANOVA-form without the S or M superscript as follows:

$$\begin{aligned} \pi_{ijk} &= (\pi_{\bullet \bullet \bullet}) + (\pi_{i \bullet \bullet} - \pi_{\bullet \bullet \bullet}) \\ &\quad + (\pi_{\bullet j \bullet} - \pi_{\bullet \bullet \bullet}) + (\pi_{\bullet \bullet k} - \pi_{\bullet \bullet \bullet}) \\ &\quad + (\pi_{ij \bullet} - \pi_{i \bullet \bullet} - \pi_{\bullet j \bullet} + \pi_{\bullet \bullet \bullet}) \\ (4) \quad &\quad + (\pi_{i \bullet k} - \pi_{i \bullet \bullet} - \pi_{\bullet \bullet k} + \pi_{\bullet \bullet \bullet}) \\ &\quad + (\pi_{\bullet j k} - \pi_{\bullet j \bullet} - \pi_{\bullet \bullet k} + \pi_{\bullet \bullet \bullet}) \\ &\quad + (\pi_{ijk} - \pi_{ij \bullet} - \pi_{i \bullet k} - \pi_{\bullet j k} + \pi_{i \bullet \bullet} \\ &\quad + \pi_{\bullet j \bullet} + \pi_{\bullet \bullet k} - \pi_{\bullet \bullet \bullet}). \end{aligned}$$

Here a dot (\bullet) indicates that the weighted mean has been taken over an index i, j, k , with respect to the associated weights $d_{i \bullet \bullet}, d_{\bullet j \bullet}$ and $d_{\bullet \bullet k}$, respectively. For example,

$$\begin{aligned} \pi_{\bullet \bullet \bullet} &= \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K d_{i \bullet \bullet} d_{\bullet j \bullet} d_{\bullet \bullet k} \pi_{ijk} = 0 \quad \text{and} \\ \pi_{\bullet j k} &= \sum_{i=1}^I d_{i \bullet \bullet} \pi_{ijk}. \end{aligned}$$

In the partition given by equation (4) the terms are pairwise orthogonal. However, as the elements of $\underline{\Pi}_M$

²Underlined bold capital letters indicate three-way arrays.

and $\underline{\Pi}_S$ generally consist of the deviation from independence based upon the observed frequencies, the fixed term, $\pi_{\bullet\bullet\bullet}$, together with the one-way terms $\pi_{i\bullet\bullet}$, $\pi_{\bullet j\bullet}$ and $\pi_{\bullet\bullet k}$ in equation (4) are all zero leaving only the bivariate and trivariate associations in the model (see [Loisel and Takane, 2016](#), [Lombardo, Takane and Beh, 2020](#)). Therefore, Pearson's Φ^2 statistic can be partitioned as

$$(5) \quad \begin{aligned} \Phi^2 &= \|\underline{\Pi}_S\|^2 = \langle \underline{\Pi}_S, \underline{\Pi}_S \rangle \\ &= \Phi_{1,2}^2 + \Phi_{1,3}^2 + \Phi_{2,3}^2 + \Phi_{1,2,3}^2, \end{aligned}$$

where the first three terms of equation (5) are the pairwise phi-squared statistics obtained by aggregating across the categories of each variable. That is, the row-column pairwise symmetric association is assessed through the term $\Phi_{1,2}^2$, and is (after multiplying it by n) asymptotically a $\chi_{(I-1)(J-1)}^2$ random variable with $(I-1)(J-1)$ degrees of freedom. Similarly, the pairwise row-tube symmetric term $n\Phi_{1,3}^2$ is asymptotically $\chi_{(I-1)(K-1)}^2$ and the column-tube association is assessed by the term $n\Phi_{2,3}^2$ which is asymptotically a χ^2 random variable with $(J-1)(K-1)$ degrees of freedom. The last term, $n\Phi_{1,2,3}^2$, measures the trivariate association and is asymptotically $\chi_{(I-1)(J-1)(K-1)}^2$.

Similarly, we may orthogonally partition the Marcotorchino statistic, $\tau_{M_{\text{num}}}$ ([Lombardo, Carlier and D'Ambra, 1996](#)) into four association terms consisting of three two-way associations and one three-way association. The partition of $\tau_{M_{\text{num}}} = \|\underline{\Pi}_M\|^2$ takes the form

$$(6) \quad \begin{aligned} \tau_{M_{\text{num}}} &= \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K p_{\bullet j\bullet} p_{\bullet\bullet k} \left(\frac{p_{ijk}}{p_{\bullet j\bullet} p_{\bullet\bullet k}} - p_{i\bullet\bullet} \right)^2 \\ &= \sum_{i=1}^I \sum_{j=1}^J p_{\bullet j\bullet} \left(\frac{p_{ij\bullet}}{p_{\bullet j\bullet}} - p_{i\bullet\bullet} \right)^2 \\ &\quad + \sum_{i=1}^I \sum_{k=1}^K p_{\bullet\bullet k} \left(\frac{p_{i\bullet k}}{p_{\bullet\bullet k}} - p_{i\bullet\bullet} \right)^2 \\ &\quad + \sum_{j=1}^J \sum_{k=1}^K \frac{1}{I} p_{\bullet j\bullet} p_{\bullet\bullet k} \left(\frac{p_{\bullet jk}}{p_{\bullet j\bullet} p_{\bullet\bullet k}} - 1 \right)^2 \\ &\quad + \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K p_{\bullet j\bullet} p_{\bullet\bullet k} \left(\frac{p_{ijk} - \alpha p_{ijk}}{p_{\bullet j\bullet} p_{\bullet\bullet k}} \right)^2, \end{aligned}$$

where $\alpha p_{ijk} = p_{ij\bullet} p_{\bullet\bullet k} + p_{i\bullet k} p_{\bullet j\bullet} + p_{\bullet jk} / I - p_{i\bullet\bullet} \times p_{\bullet j\bullet} p_{\bullet\bullet k} - p_{\bullet j\bullet} p_{\bullet\bullet k} / I$ ([Lombardo, Carlier and D'Ambra, 1996](#)). This can be expressed more succinctly as

$$(7) \quad \begin{aligned} \tau_{M_{\text{num}}} &= \|\underline{\Pi}_M\|^2 = \langle \underline{\Pi}_M, \underline{\Pi}_M \rangle \\ &= \tau_{1,2} + \tau_{1,3} + \tau_{2,3} + \tau_{1,2,3}. \end{aligned}$$

The first three terms of this partition consist of the contribution to the increase in predictability of the bivariate associations, and the last term consists of the contribution of

the trivariate association. The first two terms are derived by summing over the other index and they are equal to the numerator of the Goodman–Kruskal τ index between the response (row) variable and each of the two predictor variables. They represent the increase in predictability of the I row categories given the J column categories and the K tube categories, respectively. The third term is $1/I$ times the Pearson phi-squared statistic for the $J \times K$ contingency table formed by aggregating over the row categories. This term represents the design interaction and is a measure of the symmetric association between the two predictor variables. The last term is the contribution to the increase in predictability of the three-way asymmetric association among the three variables.

2.4 Testing the Strength of the Association

An important property of the components from the partition of Pearson's phi-squared statistic (in the symmetrical case) and the numerator of Marcotorchino statistic (in the non-symmetrical case) is that they are orthogonal and approximately independent under proper conditions. Therefore, the relative contribution of the associations to the overall measures can be assessed with respect to each other both in absolute terms and with respect to their degrees of freedom. In the case of a symmetrical design the Φ^2 contributions can be referred to using the χ^2 -distribution, as can the Marcotorchino index in the non-symmetrical case (equation (2)). In the latter case the associated test statistic is the generalisation of [Light and Margolin's \(1971\) C-statistic](#), referred to as the C_M -statistic, and defined by

$$(8) \quad C_M = (n-1)(I-1)\tau_M \sim \chi_{\alpha, df}^2.$$

Here $\chi_{\alpha, df}^2$ is the $1-\alpha$ percentile of the chi-squared distribution with degrees of freedom $df = IJK - I - J - K + 2$. Note that the left-hand side of equation (8) is an asymptotically chi-squared random variable under the sufficient, but not necessary condition, that $1 - \sum_i p_{i\bullet\bullet}^2 \approx 1/I$. Furthermore, the C_M -statistic can be partitioned as follows:

$$\begin{aligned} C_M &= \frac{(I-1)(n-1)\tau_{1,2}}{1 - \sum_{i=1}^I p_{i\bullet\bullet}^2} + \frac{(I-1)(n-1)\tau_{1,3}}{1 - \sum_{i=1}^I p_{i\bullet\bullet}^2} \\ &\quad + \frac{(I-1)(n-1)\tau_{2,3}}{I(1 - \sum_{i=1}^I p_{i\bullet\bullet}^2)} + \frac{(I-1)(n-1)\tau_{1,2,3}}{1 - \sum_{i=1}^I p_{i\bullet\bullet}^2} \\ &= C_{IJ} + C_{IK} + C_{JK} + C_{IJK}. \end{aligned}$$

The first term, C_{IJ} , is identical to the C -statistic of [Light and Margolin \(1971\)](#) for the response (row) and predictor (column) variables after aggregating across the tube categories. As discussed by [Light and Margolin \(1971\)](#), this measure is asymptotically a χ^2 random variable with $(I-1)(J-1)$ degrees of freedom. Therefore, C_{IJ} can be used to determine whether there exists a significant

asymmetric association between the row and column categories. Similarly, the statistic C_{IK} can be used to formally test for the association between the row and tube categories by comparing it with the $1 - \alpha$ percentile of the χ^2 -distribution with $(I - 1)(K - 1)$ degrees of freedom. The C -statistic, C_{JK} , for the association between the two predictors is akin to the Pearson Φ^2 statistic for the two-way contingency table formed by aggregating across each of the row categories. For further details on the C_M -statistic, see Beh, Simonetti and D'Ambra (2007) and Beh and Lombardo (2014, Section 11.5.2).

2.5 Modelling Association

The most common three-way generalisation of the singular value decomposition is the three-mode principal component model (Tucker, 1966) or the Tucker3 model, as it is now commonly referred to (Kroonenberg and de Leeuw, 1980). This three-way decomposition is used for both the symmetrical and non-symmetrical variants of three-way correspondence analysis in an analogous way to the use of the singular value decomposition in two-way correspondence analysis (see Carlier and Kroonenberg, 1996, Lombardo, Carlier and D'Ambra, 1996). The general form of the Tucker3 decomposition for the global association is

$$\begin{aligned} \pi_{ijk}^S &= \hat{\pi}_{ijk}^S + e_{ijk} \\ (9) \quad &= \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} a_{ip} b_{jq} c_{kr} + e_{ijk}, \end{aligned}$$

where P , Q and R (with $P \leq I$, $Q \leq J$ and $R \leq K$) represent the number of columns in the component matrices \mathbf{A} , \mathbf{B} and \mathbf{C} for the first (rows), second (columns) and third (tubes) way, respectively. The orthogonality of the vectors $\mathbf{a}_p = \{a_{ip}\}$, $\mathbf{b}_q = \{b_{jq}\}$ and $\mathbf{c}_r = \{c_{kr}\}$ is defined with respect to the weight matrices \mathbf{D}_I , \mathbf{D}_J and \mathbf{D}_K , respectively. The orthogonality of the vectors ensures the orthogonality of core elements g_{pqr} so that the Pearson X^2 statistic can also be expressed in terms of the sum-of-squares of the elements g_{pqr} from the core array \mathbf{G} , since $X^2 = n \sum_{pqr} g_{pqr}^2$.

The elements $g_{pqr} = \langle \hat{\mathbf{\Pi}}^S, \mathbf{a}_p \otimes \mathbf{b}_q \otimes \mathbf{c}_r \rangle$, where \otimes is the Kronecker product, can be interpreted as the generalised, or three-way, analog of the singular values. The magnitude of e_{ijk} indicates the error of the approximation between the observed π_{ijk}^S and its predicted value, $\hat{\pi}_{ijk}^S$ (see Kroonenberg, 2008, Chapter 4).

2.6 Orthogonal Polynomials

Before discussing the modelling of variables with ordered categories using orthogonal polynomials (Section 3), we will briefly introduce these polynomials. Orthogonal polynomials arise in the literature in several

forms. Their basic form is such that the i th-order polynomial includes as its highest-order term x^i .

In the present proposal the orthogonal polynomials are computed from the ordered categories and their related marginal proportions. These are defined in such a way that the zero'th-order polynomial is a constant, the first-order polynomial is based on the means, the second-order polynomial is based on the variances, the third-order term is based on the skewness, etc. Polynomials with these characteristics can be obtained via recurrence formulae (Favard, 1935, Chihara, 1978, 1990) which were considered for categorical variables by Emerson (1968) and will be referred to as *orthogonal polynomials*; for further details see, for example, Beh (1997, 1998), Beh and Lombardo (2014, pages 94–96), Lombardo, Beh and Kroonenberg (2016) and Beh and Lombardo (2021).

Consider the ordered scores of an ordered variable. Scores are defined to reflect this ordered structure and are typically chosen to be consecutive integers, starting from unity (Beh, 1998). These sets of integers for the three variables are indicated by $i = 1, 2, \dots, I$ for the rows, $j = 1, 2, \dots, J$ for the column and $k = 1, 2, \dots, K$ for the tube categories and are referred to as *natural scores*. By substituting the ordered scores and the marginal proportions into Emerson's (1968) recurrence formulae, the calculation of the orthogonal polynomials is similar for all ordered variables. Their meaning will be briefly described for the column categories; the vector of such column orthogonal polynomials will be denoted using the Greek letter β .

- Zero'th-order polynomial: The zero'th-order polynomial, β_0 , is equal to 1.
- First-order polynomial: The first-order polynomial, β_1 , indicates the variation in the mean of the categories.
- Second-order polynomial: The second-order orthogonal polynomial, β_2 , is a dispersion term as it is a function of the squared scores.
- Higher-order polynomials: Higher-order polynomials represent higher-order moments of the ordered categories.

The normalisations of the orthogonal polynomials are then carried out using the norm with respect to the metric of the marginal proportions. For example, for the columns, the norm is $\|\beta_v\|_{\mathbf{D}_J}$ where the subscript v indicates the order of the polynomial $v = 0, \dots, J - 1$. Thus, the column polynomials are orthonormal with respect to the marginal proportions $p_{\bullet j}$ such that

$$\sum_{j=1}^J p_{\bullet j} \beta_{jv} \beta_{jv'} = \begin{cases} 1, & v = v', \\ 0, & v \neq v'. \end{cases}$$

Similarly, the row (α_i) and tube (γ_k) polynomials are orthonormal with respect to the marginal proportions $p_{i\bullet\bullet}$ and $p_{\bullet\bullet k}$, respectively.

3. ORDINAL THREE-WAY CORRESPONDENCE ANALYSIS

Both the standard symmetrical and non-symmetrical variants of three-way correspondence analysis use singular vectors for the orthogonal decomposition of the deviations. However, specific properties of the variables such as ordinality of their categories are not catered for, even though ordinality often features heavily in the interpretation. There are several contributions in the literature that provide ways of performing classical two-way correspondence analysis with order restrictions imposed on a set of categories. See, for example, Nishisato (1980, Chapter 2), Böckenholt and Böckenholt (1990), Takane, Yanai and Mayekawa (1991) and Takane and Jung (2009).

In this paper, we go one step further and incorporate the ordinality of the variables into the three-way modelling itself by replacing the singular vectors with the orthogonal polynomials (Emerson, 1968) described in the previous section. These are suitable for this purpose due to their orthogonality, their ease of computation and their interpretational qualities (see below for details). The decomposition technique will be referred to as the *trivariate moment decomposition* as it is an analogue of its two-way variant, *bivariate moment decomposition* (see, e.g., Beh, 1997, Beh and Lombardo, 2004, Sections 3.10, 3.11, Lombardo, Carlier and D’Ambra, 1996) and it assumes that all three variables are ordered. However, in a general framework for nonparametric multivariate function estimation, polynomial decomposition can be considered akin to the smoothing spline analysis of variance (SS-ANOVA) model (Lin and Zhang, 2006, Wahba, 1990) for categorical variables.

Since the essence of the symmetrical and non-symmetrical variants is the same with respect to the inclusion of orthogonal polynomials, we concentrate on the symmetrical case as it is the best known version.

When a row variable consists of I ordered categories, we can compute I orthogonal polynomials, using Emerson’s (1968) recurrence formulae, the ordered row scores and the row marginal proportions. Each polynomial has a specific interpretation and measures of associations derived from them can be formally tested. As indicated above, we generate as many orthogonal polynomials as there are categories: the zero’th order orthogonal polynomial is trivial, the first-order polynomial is linear (respecting the ordinality of categories), the second-order polynomial is quadratic (describing the dispersion of categories), and so on, and all are centered and orthogonal to each other. Thus, in a visual display of the first and second orthogonal polynomials, the categories lie ordered on a parabola centred at the origin.

3.1 Modelling Ordinal Three-Way Polynomial Correspondence Analysis

In the trivariate moment decomposition for symmetrical and non-symmetrical correspondence analysis, the decompositions of the Pearson Φ^2 statistic and of the Marcotorchino $\tau_{M_{num}}$ statistic are defined by replacing \mathbf{a}_p , \mathbf{b}_q and \mathbf{c}_w (see equation (9)) with their orthogonal polynomial equivalents. In particular, for the symmetrical case, π_{ijk}^S is given by

$$(10) \quad \pi_{ijk}^S = \sum_{u=0}^U \sum_{v=0}^V \sum_{w=0}^W z_{uvw} \alpha_{iu} \beta_{jv} \gamma_{kw}.$$

For the decomposition of equation (10), the row, column and tube orthogonal polynomials are weighted such that $\sum_{i=1}^I p_{i\bullet\bullet} \alpha_{iu}^2 = 1$, $\sum_{j=1}^J p_{\bullet j\bullet} \beta_{jv}^2 = 1$ and $\sum_{k=1}^K p_{\bullet\bullet k} \gamma_{kw}^2 = 1$, respectively. Note that the indices u, v, w run from 0 to U, V and W , respectively, (where $U = I - 1, V = J - 1, W = K - 1$) and correspond to the orders of the polynomials. The values z_{uvw} in equation (10) are analogous to the core elements g_{pqr} in the nominal case, hence they are also referred to as *polynomial core elements* and are defined by

$$z_{uvw} = \langle \underline{\Pi}_S, \alpha_u \otimes \beta_v \otimes \gamma_w \rangle.$$

For example, using the results from Section 2.6, the polynomial core element, z_{111} , is given by

$$\begin{aligned} z_{111} &= \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \pi_{ijk}^S p_{i\bullet\bullet} p_{\bullet j\bullet} p_{\bullet\bullet k} \alpha_{i1} \beta_{j1} \gamma_{k1} \\ &= \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K \pi_{ijk}^S p_{i\bullet\bullet} p_{\bullet j\bullet} p_{\bullet\bullet k} \\ &\quad \times \frac{(i - \mu_I)}{\sigma_I} \frac{(j - \mu_J)}{\sigma_J} \frac{(k - \mu_K)}{\sigma_K}, \end{aligned}$$

where i, j and k are the natural scores for the i th row, j th column and k th tube category, respectively. Also, μ_I, μ_J and μ_K are the weighted means of these scores, while σ_I, σ_J and σ_K are the corresponding standard deviations, respectively.

The term z_{111} is akin to the magnitude of the generalised correlation between the linear polynomials of the ordered row, column and tube categories, respectively. A comprehensive discussion of generalised correlations for two-way contingency tables can be found in Rayner and Best (1996) and for three-way tables in Rayner and Beh (2009), and their interpretation extends naturally to the non-symmetrical case.

It is the decomposition consisting of orthogonal polynomials and polynomial core array elements as defined by equation (10) that is generally referred to as the *trivariate moment decomposition* of $\underline{\Pi}_S$. From Section 2.3 and

equation (5), it follows that the total inertia of the ordered three-way contingency table is equal to the Pearson Φ^2 statistic, to the squared norm of the array $\underline{\Pi}_S$, and to the sum of the squared polynomial core elements, such that

$$(11) \quad \Phi^2 = \|\underline{\Pi}_S\|^2 = \sum_{u=0}^U \sum_{v=0}^V \sum_{w=0}^W z_{uvw}^2.$$

When analysing the association between the three variables using the orthogonal polynomials to yield each term of the Φ^2 partition (equation (5), see also Tables 2, 3, 4, 5), we can also re-express the decomposition given by equation (10) as an additive polynomial model for the global symmetrical dependence measure as a function of the association terms such that

$$(12) \quad \begin{aligned} \pi_{ijk}^S &= \sum_{u=1}^U \sum_{v=1}^V z_{uv0} \alpha_{iu} \beta_{jv} + \sum_{u=1}^U \sum_{w=1}^W z_{u0w} \alpha_{iu} \gamma_{kw} \\ &+ \sum_{v=1}^V \sum_{w=1}^W z_{0vw} \beta_{jv} \gamma_{kw} \\ &+ \sum_{u=1}^U \sum_{v=1}^V \sum_{w=1}^W z_{uvw} \alpha_{iu} \beta_{jv} \gamma_{kw}. \end{aligned}$$

Here the z_{uv0} , z_{u0w} and z_{0vw} -terms represent the marginal bivariate generalised correlations of the partition. For example, the first term is defined as $z_{uv0} = \langle \underline{\Pi}_S, \alpha_u \otimes \beta_v \otimes \gamma_0 \rangle$.

It should be noted that the computation of the terms z_{uv0} , z_{u0w} and z_{0vw} requires special attention. For example, z_{uv0} is computed using all the row and column polynomials while only the tube polynomial remains constant, at $w = 0$. The terms z_{u0w} and z_{0vw} are computed similarly; see Section 4. The definitions of the above terms can be used to show that the Pearson Φ^2 statistic of equations (3) and (11) can also be expressed as

$$(13) \quad \begin{aligned} \Phi^2 &= \sum_{u=0}^U \sum_{v=0}^V \sum_{w=0}^W z_{uvw}^2 \\ &= \sum_{u=1}^U \sum_{v=1}^V z_{uv0}^2 + \sum_{u=1}^U \sum_{w=1}^W z_{u0w}^2 \\ &+ \sum_{v=1}^V \sum_{w=1}^W z_{0vw}^2 + \sum_{u=1}^U \sum_{v=1}^V \sum_{w=1}^W z_{uvw}^2. \end{aligned}$$

See Beh and Davy (1998) for more details on this partition.

3.2 Hybrid Decomposition for Nominal and Ordinal Variables

When the three-way contingency table consists of nominal and ordinal variables, the approximation of the global dependence, or total inertia, involves computing Tucker3

components for the nominal variables and orthogonal polynomials for the ordered variables (Lombardo and Beh, 2017). Two cases involving such a structure can arise: (1) two ordered variables and one nominal variable, or (2) two nominal variables and only one ordered variable. As an example of case 1, after computing the polynomials for the two ordered variables, say the row and column variables, and the Tucker3 components for the tube variable, the *hybrid decomposition* takes on the form

$$(14) \quad \begin{aligned} \pi_{ijk}^S &= \hat{\pi}_{ijk}^S + e_{ijk} \\ &= \sum_{u=0}^U \sum_{v=0}^V \sum_{r=1}^R z_{uvr} \alpha_{iu} \beta_{jv} c_{kr} + e_{ijk}, \end{aligned}$$

where the α_u and β_v are the row and column polynomials, respectively, and the c_r are the tube singular vectors (Tucker3 components). The values z_{uvr} in equation (14) are defined by $z_{uvr} = \langle \hat{\underline{\Pi}}^S, \alpha_u \otimes \beta_v \otimes c_r \rangle$.

While the number of the orthogonal polynomials for the rows and columns should always be equal to the number of categories in a variable (see Section 2.6), the number of Tucker3 components for the tube variable can be smaller ($R \leq K$). Note that in the hybrid decomposition the solutions are not nested, in contrast to the trivariate moment decomposition (equation (10)) which allows for the definition of nested solutions. A complete orthogonal decomposition is always used when all the three variables are ordered, as it is for model (10), but is seldom used in practice when the variables are not all ordered.

3.3 Number of Components in Trivariate Moment Decompositions and in Hybrid Decompositions

To determine which polynomial components are important for approximating the data with the model, it is necessary to inspect the inertia that is accounted for. Unlike the singular vector based decomposition, one needs to evaluate all terms in the orthogonal decomposition because there is no decreasing order in the inertia that is accounted for.

The size of a polynomial triplet (u, v, w) depends on how well its orthogonal polynomials approximate the association structures in a given set of data, and this may be a third-order cubic polynomial rather than the first-order linear one. The size of the squared core values indicate how well a certain triplet succeeds in accounting for the variability. Clearly, the higher the order of a polynomial, the more difficult its interpretation becomes. Therefore, a preference is given to those triplets of polynomials with lower-orders combined with large squared core values.

Note that if higher-order polynomials have large squared core values but none of the lower-order ones do, then the higher-order polynomial can generally be considered as an unexplainable source of variation due to the high frequency fluctuations between the categories. The higher-order polynomials, in most cases, are associated

with the unstructured information, showing the noise in the data and often cannot be interpreted sensibly. On the other hand, lower-order polynomials take care of smoothing the profiles of the categories over the other categories. They provide the analyst with insightful, parsimonious, and relatively smooth descriptions of the associations or increases in predictability in three-way tables. Therefore, the association structures can be modelled using constant, linear, quadratic and cubic orthogonal polynomials.

However when performing the hybrid decomposition we also need to compute generalised singular vectors. As a result, when assessing the dimensionality in a more formalised manner, it may be appropriate to consider various methods based on numerical and visual inspections (such as the convex hull or Pareto frontier); see, for example, Kroonenberg and Oort (2003), Ceulemans and Kiers (2006, 2009), Kroonenberg and Anderson (2006), Ceulemans, Timmerman and Kiers (2011) and Lorenzo-Seva, Timmerman and Kiers (2011). For more of a discussion on the interpretation of lower and higher order polynomials, refer to Rayner and Best (1996) and Beh and Lombardo (2014, Chapter 6).

Since our primary interest is in describing the utility of orthogonal polynomials as components, the number of dimensions to retain will be based upon the varying degrees of explained inertia by the different combinations of polynomials and generalised singular vectors (see Section 4).

3.4 Graphing the Trivariate Moment Decomposition: Polynomial Biplots

To visually portray the association structure between categorical variables in a three-way contingency table, Carlier and Kroonenberg (1996) and Kroonenberg (2008) discussed various types of biplots. Here, we focus on interactive biplots (also referred to as nested-mode biplots by Kroonenberg, 2008, Section 11.5.4), since they are the only ones that allow for a simultaneous representation of all row, column and tube categories in a single display (see also Gower, Le Roux and Gardner-Lubbe, 2016).

Therefore to depict graphically the symmetric structure of the ordered categorical variables, we consider *three-way polynomial interactive* biplots. Such biplots represent a generalisation to the three-way case of the polynomial biplots presented with respect to the two-way ordered variants of correspondence analysis (Lombardo, Beh and Kroonenberg, 2016). These three-way biplots essentially follow the construction of those for nominal variables proposed by Kroonenberg (2008, pages 273–280), but with the singular vectors replaced by the orthogonal polynomials. However for the non-symmetrical case, their interpretation is quite different since they portray the predictability of the response given the predictors rather than the mutual association. The non-symmetrical association between the three variables can be visualised using three-way polynomial interactive biplots by reflecting how the

column-tube category interactions impact upon the row categories.

As described in the following sections, three-way polynomial biplots can take two forms depending on whether the row coordinates or the column-tube coordinates are isometric. The isometry indicates which components are in principal coordinates and thus which distances between which categories can be determined and properly assessed.

3.4.1 *Interactively-coded polynomial biplot.* [rows: standard polynomial coordinates; column \times tubes principal polynomial coordinates].

Following Kroonenberg’s (2008, page 273) terminology for the interactively-coded isometric biplot, the *reference* mode is the variable (in our case, the row) visually depicted using standard coordinates and the interactive-mode (column-tube category combination) is in principal coordinates. We consider the case where the reference mode is the row variable, while the pair consisting of the j th column and the k th tube categories will be coded interactively (j, k) and will be represented by a single point. The coordinates of the two modes follow from

$$\begin{aligned} \pi_{ijk}^S &= \sum_{u=0}^U \alpha_{iu} \left[\sum_{v=0}^V \sum_{w=0}^W z_{uvw} \beta_{jv} \gamma_{kw} \right] \\ &= \sum_{u=0}^U \alpha_{iu} h_{jk,u}. \end{aligned}$$

Here $h_{jk,u}$ represents the interactively coded (column-tube category) *principal polynomial coordinates*, that can be expressed as

$$h_{jk,u} = \sum_{i=1}^I p_{i \bullet \bullet} \alpha_{iu} \pi_{ijk}^S.$$

Each of the I slices of the array $\underline{\Pi}_S$ is approximated by the linear combination of the matrices $\mathbf{H}_u = (h_{jk,u})$ (for $u = 0, 1, 2, \dots, U$) with the coefficients α_{iu} . Therefore, the main aim of this plot is to explicitly code the column-tube category (j, k) and to join the points of the ordered categories.

For this biplot, we are dealing with a simple U -dimensional display space ($U = I - 1$ excluding the zero’th polynomial) where the coordinates for both the row and column-tube categories are displayed in the space defined by the U row orthogonal polynomial axes, that is, the row *standard polynomial coordinates*. Along these axes we can investigate the association between the row categories and the column-tube categories by observing the trends of the rows and those of the interactively coded categories in a joint graphical display. For example, in a linear-by-quadratic biplot, we display the row standard coordinates α_{i1} and α_{i2} , and the column-tube principal coordinates $h_{jk,1}$ and $h_{jk,2}$ where the scales for the two

axes are determined by the polynomial core values $z_{1\bullet\bullet}$ and $z_{2\bullet\bullet}$, respectively. In the symmetric case, the percent of variability explained by the u 'th axis is given by $z_{u\bullet\bullet}^2 / X^2 \times 100$.

3.4.2 *Single-variable polynomial biplots.* [rows: principal polynomial coordinates; columns \times tubes: standard polynomial coordinates].

Such a biplot can be constructed when the coordinates for both the single-variable and the interactively-coded variables are displayed in the space defined by the interactively-coded orthogonal polynomials (instead of in the space of the single-variable orthogonal polynomials). Therefore, the aim of this plot is to visually depict the association between the three variables such that the structure of the ordinal variable is preserved while reflecting the possibility that there exists a column-tube association. We define the *single-variable isometric polynomial biplot*. Here we are dealing with a space whose dimensionality is $VW = (J - 1 \times K - 1)$ (when excluding the zero'th polynomials). To obtain the single-variable biplot, the deviations from independence, given in π_{ijk}^S , are decomposed such that

$$\begin{aligned} \pi_{ijk}^S &= \sum_{u=0}^U \sum_{v=0}^V \sum_{w=0}^W [z_{uvw} \alpha_{iu}] [\beta_{jv} \gamma_{kw}] \\ &= \sum_{v=0}^V \sum_{w=0}^W f_{i,vw} h_{jk,vw}. \end{aligned}$$

Here $f_{i,vw}$ is the row principal polynomial coordinates, given by the product of the i th row polynomial for the polynomial core values z_{uvw} , while $h_{jk,vw}$ represents the interactively-coded standard polynomial coordinate, using the interactively-coded element of the j th column polynomial and the k th tube polynomial.

Therefore, the single-variable principal polynomial coordinates are defined by

$$\mathbf{F}_{(I \times VW)} = \mathcal{AZ}_{(U \times VW)} \quad \left(f_{i,vw} = \sum_{u=0}^U \alpha_{iu} z_{uvw} \right),$$

and the interactively-coded standard polynomial coordinates are

$$\mathbf{H}_{(JK \times VW)} = (\mathcal{B} \otimes \mathcal{C}) \quad (h_{jk,vw} = \beta_{jv} \gamma_{kw}).$$

We use the generalised correlations of the polynomial core array to select the relevant combination of orthogonal polynomials to display. For the single-variable biplots, the points for the standard coordinates of the column-tube pairs will be sequentially connected to show the column-tube trend, while the rows will be graphically depicted as points. For example, in a linear-by-quadratic biplot we display the row principal coordinates $f_{i,11}$ and $f_{i,22}$ (where the scales for the two axes are determined by the polynomial core values $z_{\bullet 11}$ and $z_{\bullet 22}$, respectively) and

the column-tube standard coordinates $h_{jk,11}$ (using the linear, column and tube polynomials) and $h_{jk,22}$ (using the quadratic, column and tube polynomials). In the symmetric case, the percent of variability reported by the interactive linear polynomials axis ($\beta_1 \gamma_1$) is $z_{\bullet 11}^2 / X^2 \times 100$ and for the interactive quadratic polynomial axis ($\beta_2 \gamma_2$) is given by $z_{\bullet 22}^2 / X^2 \times 100$.

However, when using the hybrid decomposition, as we do for the example in Section 4, we consider the generalised singular vectors for the nominal variable. As a result, for the single-variable biplot, the column-tube standard coordinates are hybrid orthogonal polynomials resulting from the combination of polynomials with generalised singular vectors.

3.4.3 *Interpretation of polynomial biplots.* Some general rules for the interpretation of the polynomial biplots are listed here for reference, but they will become clearer when discussing the example. For all types of polynomial biplots, the standard polynomial coordinates play the most important role in determining the strength of the association as they represent the polynomial axes on which all the other categories are projected.

– *Interactively-coded biplots.* The origin of this plot represents the mean of the ordered categories of the reference mode (rows in our application; see Section 4).

Linear (or first-order) polynomial. If the standard polynomial coordinate of a reference category is positive and large then its value is higher than the mean profile. It influences all those column-tube categories that take large positive coordinates on the first axis.

Quadratic (or second-order) polynomial. If the standard coordinate of a reference category is positive and large on the quadratic polynomial, it means that it has a large dispersion (positive difference with respect to the mean of the ordered reference scores). All those column-tube categories with a large coordinate on this second axis will be affected by that reference category.

In general when plotting the linear-by-quadratic (or the quadratic-by-linear) polynomial, the plot always shows a parabolic shape of different narrowness and concavity. However, when using two linear polynomial axes (which is possible only for interactively-coded biplots), the plot always results in a configuration akin to the spokes of a cartwheel, where all lines go through the origin due to the centring of the polynomials.

– *Single-variable biplots.* The origin of this plot represents the mean of the ordered, interactively-coded (column-tube) categories.

Linear (or first-order) polynomial. If the standard polynomial coordinate of the interactively coded column-tube category is positive and large then it

TABLE 1
Rating-ranking data: Partitioning of the Pearson index; $I = \text{Ratings}$; $J = \text{Rankings}$; $K = \text{Countries}$

Association	Index	% Inertia	df	X^2/df	Type
$n\Phi_{1,2}^2 = X_{IJ}^2$	18,359	86%	64	287	Ratings–Rankings
$n\Phi_{2,3}^2 = X_{JK}^2$	255	1%	32	8	Rankings–Countries
$n\Phi_{1,3}^2 = X_{IK}^2$	590	3%	32	18	Ratings–Countries
$n\Phi_{1,2,3}^2 = X_{IJK}^2$	2062	10%	256	8	Ratings \times Rankings \times Countries
Total	21,266	100%	384	55	Overall

means that category is above its mean. All row categories that are in the positive direction will be affected by that column-tube category.

Quadratic (or second-order) polynomial. If the standard coordinate of a column-tube category is positive and large on the quadratic polynomial, then that category has a large dispersion, that is, a positive difference with respect to its mean, and one or more row categories can be affected by that large column-tube coordinate.

Finally, we cannot say that one kind of polynomial biplot is better than the other because they portray similar information about the variable structure of the association between the variables, but does so differently. However, they can highlight different aspects of the association particularly at the origin of these polynomial biplots since they do not have the same meaning.

4. APPLICATION

To illustrate the advantages of the ordinal three-way correspondence analysis method outlined in this paper, when compared to the classical analysis, we will reanalyse Van Herk and Van de Velden’s (2007) three-way contingency table. This table consists of two ordered variables and one nominal variable which are *symmetrically* associated, that is, there is not a dependence design. This table can be analysed with a hybrid decomposition where the orthogonal polynomials are computed for the ordered row and column variables, and the Tucker3 components for the nominal tube variable.

The data, which are very useful in sensory analysis and marketing research, were part of a large European survey dataset (Van Herk and Van de Velden, 2007). In particular, in an empirical study based on the List of Values (LOV; Kahle, 1983), participants were asked to study the LOV list, and then rated each item on a nine-point scale. Subsequently, the participants ranked the same nine statements, assessing the most important item with a score “9”, and the least important item with “1”. The LOV was described to participants as a list of things that some people look for or want out of life. The items on the LOV list were *Sense*

of belonging, Excitement, Warm relationships with others, Self-fulfillment, Being well respected, Fun and enjoyment in life, Security, Self-respect and a Sense of accomplishment; for more details see Van Herk and Van de Velden (2007).

The three-way contingency table consists of nine rating values against nine ranking values given by the same participants across five European countries. In the literature, there has been an ongoing debate about the pros and cons of rating and ranking procedures (see Alwin and Krosnick, 1985, Rodrigue et al., 2000, Van Herk and Van de Velden, 2007) that is beyond the scope of this paper.

Van Herk and Van de Velden’s analysis was aimed at understanding whether the information from the ratings and rankings were consistent (or concordant) across five European countries. Our aim is to evaluate the relative effectiveness of the polynomial approach versus the traditional multi-way approach. Furthermore, while Table 1 shows that the partition of Pearson’s chi-squared statistic is identical to that of the standard analysis (Carlier and Kroonenberg, 1996, 1998), Tables 2, 3, 4 and 5 provide the partition of Pearson’s X^2 statistic for the ordered categorical variables when using the orthogonal polynomials (see equation (13)).

TABLE 2
Partitioning of the Pearson term X_{IJ}^2 : $I = \text{Ratings}$; $J = \text{Rankings}$

Ratings polynomials (α_u)	X_{IJ}^2	% Inertia	df	X^2/df
1st order: α_1	12,701	70%	8	1588
2nd order: α_2	3883	21%	8	485
3rd order: α_3	834	5%	8	104
<i>all-higher-order</i>	941	5%	40	24
Total	18,359	100%	64	287

Rankings polynomials (β_v)	X_{IJ}^2	% Inertia	df	X^2/df
1st order: β_1	13,820	75%	8	1727
2nd order: β_2	3178	17%	8	397
3rd order: β_3	605	3%	8	76
<i>all-higher-order</i>	756	5%	40	19
Total	18,359	100%	64	287

TABLE 3

Partitioning of the Pearson term X^2_{JK} : $I = \text{Ratings}$; $K = \text{Countries}$

Ratings polynomials (α_u)	X^2_{JK}	% Inertia	df	X^2/df
1st order: α_1	207	35%	4	52
2nd order: α_2	133	23%	4	33
3rd order: α_3	90	15%	4	23
<i>all-other-order</i>	159	27%	20	8
<i>Total</i>	589	100%	32	18

4.1 Rating-Ranking Data

The rating-ranking data were obtained from a sample of about 5000 participants from five European countries; France, Italy, Germany, UK and Spain. Since Van Herk and Van de Velden (2007) were interested in the comparative use of the different scales, they cross-classified the ratings and rankings. That is, they counted the number of times respondents gave a rating i and a ranking j to the same item. This provided them with a contingency table of size $9 \times 9 \times 5$ ($I \times J \times K$), in which an entry n_{ijk} denotes the number of times respondents of a country k assigned an item both a rating i and a rank j ; one could refer to this number as *rating-ranking concordance* or *consistency*.

Van Herk and Van de Velden (2007) analysed the association in the three-way table of *Ratings* \times *Rankings* \times *Countries* with standard three-way correspondence analysis.

However, in contrast to Van Herk and Van de Velden (2007), we modelled the association in the table and the ordinality of the two measurement variables by employing orthogonal polynomials for rating and ranking points and generalised singular vectors for country points. The similarity between ratings and rankings of the same set of items would be expected to be higher when inconsistency is low. However, as stated by Van Herk and Van de Velden (2007), inconsistency between ratings and rankings may happen due to the cognitive complexity of the task, because of potential misunderstandings of the procedure, or from the presence of random errors.

4.2 Aim of the Analysis

The main questions we shall address here are:

TABLE 4

Partitioning of the Pearson term X^2_{JK} : $J = \text{Rankings}$; $K = \text{Countries}$

Rankings polynomials (β_v)	X^2_{JK}	% Inertia	df	X^2/df
1st order: β_1	139	55%	4	35
2nd order: β_2	74	29%	4	18
3rd order: β_3	16	6%	4	4
<i>all-other-order</i>	26	6%	20	1
<i>Total</i>	255	100%	32	8

TABLE 5

Partitioning of the Pearson term X^2_{JK} : $I = \text{Ratings}$; $J = \text{Rankings}$; $K = \text{Countries}$

Ratings polynomials (α_u)	X^2_{JK}	% Inertia	df	X^2/df
1st order: α_1	638	30%	32	20
2nd order: α_2	702	34%	32	22
3rd order: α_3	217	10%	32	7
<i>all-higher-order</i>	505	26%	160	3
<i>Total</i>	2062	100%	256	8

Rankings polynomials (β_v)	X^2_{JK}	% Inertia	df	X^2/df
1st order: β_1	809	40%	32	25
2nd order: β_2	545	26%	32	17
3rd order: β_3	281	14%	32	9
<i>all-higher-order</i>	427	20%	160	3
<i>Total</i>	2062	100%	256	8

1. To what extent is there an association among *Ratings*, *Rankings* and *Countries*?

2. What are the trends of *Ratings* and *Rankings* across *Countries*?

3. Can these trends be more parsimoniously modelled and easier to interpret using orthogonal polynomials rather than singular vectors?

From Table 1, we can assess how much the variables *Ratings* and *Rankings*, as well as their mutual association, change across *Countries*. Fortunately, from Tables 2, 3, 4 and 5, we can also evaluate the importance of the orthogonal polynomials when identifying the most important association terms from the partition of Φ^2 . This information is also contained in the polynomial core elements of Table 8.

4.3 Size and Nature of the Dependence

Our analysis starts with investigating the importance of the two-way and three-way association terms with the aim of determining if and how *Ratings* were associated with *Rankings* across *Countries*. To this end, we partitioned the Pearson Φ^2 statistic into three two-way terms and one three-way term; see equation (5). Given the large number of observations the two-way terms and the three-way term are all statistically significant at the 0.01 level. Therefore, we will in general not refer to significance, but only compare the relative sizes of the inertia (X^2/df). Indeed, the mean of a χ^2 variable is equal to df , thus X^2/df allows for comparisons between X^2 values from different (asymptotical) chi-squared distributions. The overall association is $X^2 = 21,266$ ($df = 384$) which shows that there is evidence of a strong association among *Ratings*, *Rankings* and *Countries*.

The percentage contributions to the inertia of the four terms are given in Table 1, but can also be derived from

squaring the polynomial core elements using equation (13); this is akin to squaring the core elements obtained from the generalised singular value (Tucker3) decomposition (see Kroonenberg, 2008, Section 9.5.1). The major source of association in these data is the two-way association between *Ratings* and *Rankings*, which contributes to 86% of the total inertia.

Table 2 shows that, for both the ordered variables *Ratings* (*I*) and *Rankings* (*J*), the linear and quadratic polynomials play the most important role in the contribution to the total inertia. Using all row and column orthogonal polynomials lead to statistically significant scores of association for the *Ratings* and *Rankings* variables.

The two remaining two-way association terms do not make a substantial contribution to the inertia of these data, as their contribution is very low; 3% and 1% of the inertia for X^2_{IK} and X^2_{JK} , respectively. However, the relative size (X^2/df) of the two-way association term X^2_{IK} (18) is considerably greater than that of the three-way term X^2_{IJK} (8). The association between *Ratings* and *Countries* in Table 3 indicates differences in rating patterns in the five countries, as well as the association between *Rankings* and *Countries* in Table 4 shows some differences in ranking patterns across the five European countries. As Van de Herk and Van de Velden pointed out in their paper, these association terms should be close to zero as the respondents were asked to use a precise rating and ranking system in all countries. However, in practice, respondents did not always perform a consistent rating and ranking evaluation.

In terms of the orthogonal polynomials, Table 3 shows that, when studying the association between *Ratings* and *Countries*, the contribution of the first four polynomials lead to statistically significant sources of association, while the polynomials of order greater than five do not. In Table 4, we illustrate the partition of the bivariate term X^2_{JK} using the orthogonal polynomials. Here the first and second order polynomials provide the greatest contribution to the association between *Rankings* and *Countries*. The first three orthogonal polynomials result statistically significant, while the polynomials of order greater than three are not significant.

Finally, the three-way association is quite dominant and contributes to 10% of the total inertia; see Table 1. The orthogonal polynomials which contribute to this trivariate association are all statistically significant and are given in Table 5. We note that the linear polynomial of *Ratings* is no longer the most important source of association, as the contribution of the quadratic polynomials to the inertia is higher (34%) in this trivariate association term. Also, its relative size ($X^2/df = 22$) is higher than the linear polynomial's size ($X^2/df = 20$).

TABLE 6
Rating–ranking data: Un-normalised polynomial coefficients for Ratings and Rankings

Term	Polynomials			Marg. Prop.
	1st	2nd	3rd	
<i>Ratings</i>				
	α_1	α_2	α_3	
<i>rat1</i>	-1.46	1.44	-1.17	0.39
<i>rat2</i>	-1.08	0.28	0.75	0.19
<i>rat3</i>	-0.70	-0.54	1.26	0.14
<i>rat4</i>	-0.32	-1.02	0.83	0.08
<i>rat5</i>	0.06	-1.16	-0.05	0.08
<i>rat6</i>	0.44	-0.97	-0.91	0.03
<i>rat7</i>	0.82	-0.45	-1.28	0.03
<i>rat8</i>	1.20	0.42	-0.69	0.02
<i>rat9</i>	1.58	1.62	1.35	0.04
<i>Rankings</i>				
	β_1	β_2	β_2	
<i>rank1</i>	-0.83	0.71	-0.51	0.13
<i>rank2</i>	-0.38	-0.33	1.00	0.12
<i>rank3</i>	0.07	-1.00	1.03	0.11
<i>rank4</i>	0.52	-1.30	0.16	0.11
<i>rank5</i>	0.97	-1.22	-1.05	0.11
<i>rank6</i>	1.42	-0.76	-2.03	0.11
<i>rank7</i>	1.87	0.07	-2.22	0.10
<i>rank8</i>	2.33	1.27	-1.04	0.11
<i>rank9</i>	2.78	2.85	2.08	0.11

4.4 Polynomial Axis

To compute the orthogonal polynomials, we assign natural scores to the ordered categories of the *Ratings* and *Rankings* variables and consider their associated marginal proportions when using Emerson's (1968) recurrence formulae. Table 6 summarises the first, second and third order polynomial coefficients for each of the two ordered ways.

Apart from the normalisations, the coefficients of the orthogonal polynomials in Table 6 show their linear, quadratic and cubic shapes. The normalisations are necessary to respect the orthonormality constraints, and therefore the polynomial components are multiplied by the square root of the marginal proportions. Thus, the zero'th polynomials are not equal to 1.00 but vary based on the square root of their marginal proportions (see the last column of Table 6); for example, after the normalisation, the first rating category labelled as *rat1* has a zero'th order polynomial coefficient of $\sqrt{0.39} = 0.62$.

4.4.1 Model dimensionality and Tucker3 components. When using the orthogonal polynomials, the number of dimensions is equal to the number of ordered categories, but when using the hybrid decomposition we have to also compute the generalised singular vectors. However for the sake of simplicity, in our application the dimensionality has been selected by inspecting the varying degrees of

TABLE 7
Tucker3 component for Countries

Countries	r_1	r_2
Italy (I)	-0.37	0.03
United Kingdom (UK)	-0.40	0.47
France (F)	-0.43	-0.50
Spain (S)	-0.47	0.57
Germany (G)	-0.54	-0.46

explained variance from the different interpretable plots. Therefore, for *Ratings* and *Rankings*, the dimensionality is equal to the number of the orthogonal polynomials, that is, nine, while for the nominal variable *Countries* only two dimensions are considered. Indeed, the improvement to the explained inertia of the second Tucker3 component was not very important, but this second tube component allowed for a linear-by-linear interactively-coded biplot to be constructed (Figure 4).

Table 7 reports the two generalised singular vectors of the *Countries* variable. We observe that the scores of Germany and France on the second Tucker3 component of *Countries* are negative while the scores of Spain, UK and Italy are positive. All countries have large scores for the first and second Tucker3 axis except Italy, whose score on the second axis is close to zero. This will affect the interactively-coded polynomial biplots of Figures 3 and 4.

The dimensionality choice was based on the amount of explained variance when adding/deleting dimensions to the tube variable. Doing so, the resulting explained inertia is 19,971, that is, 94% of the total chi-squared statistic. In contrast, the model dimension of the classical analysis performed by Van Herk and Van de Velden (2007) was chosen to be $2 \times 2 \times 1$ and its explained inertia was 17,726, that is, 83% of the total chi-squared statistic.

4.5 Polynomial Core Array and Generalised Correlations

The next step is to investigate which combinations of polynomials and singular vectors, that is, which triplets $(\alpha_u, \beta_v, \text{ and } c_r)$, contribute substantially to the inertia. This information is to be found in the (squared) polynomial core elements. Observe that the sum of the squared polynomial core elements is equal to the total explained inertia, that is, 19,971. As already indicated, the orthogonal polynomials are ordered by their degree and not by their inertia accounted for so that we first have to calculate all polynomials for the row and column ways, as well as the Tucker3 components for the third way, and the polynomial core array. Subsequently, we select those core elements with the highest z_{uvr}^2 . At present, this is a subjective process without stringent significance tests.

Table 8 shows the elements z_{uvr} of the polynomial core array for all combinations of u, v , and r . Each z_{uvr}^2 represents the inertia accounted for by the triplet $(\alpha_u, \beta_v, \text{ and } c_r)$

or the strength of the link between these components (see Section 3.3). An inspection of these polynomial core elements in Table 8 shows a strong link ($z_{111}^2 = 12,027$) between the linear polynomial of *Ratings* ($u = 1$) and the linear polynomials of *Rankings* ($v = 1$) when considering the first country component $r = 1$. This triplet shows that the association among *Ratings*, *Rankings* and *Countries* is for the largest part linear in both *Ratings* and *Rankings* variables, and accounts for 57% of the total explained inertia. The next largest polynomial core element is due to the triplet of the quadratic polynomial of *Ratings* with the linear polynomial of *Rankings* and the first component of *Countries* ($z_{211}^2 = 1867$). While it is much smaller, it still represents an important contribution to fitting the model to the data (9% of total). Another important contributor to the total explained inertia of polynomial core elements is z_{221}^2 and reflects 8% of this inertia. It refers to the quadratic components of *Ratings* and of *Rankings* when $r = 1$.

4.5.1 *Contributions of the orthogonal polynomials in an interactively-coded polynomial biplot.* From Table 8, we can also compute the percentage contribution of each of the orthogonal polynomials for the two ordered variables to the total inertia ($X^2 = 21,266$). Here, we focus our attention on the rating polynomials since *Ratings* is the reference variable in the interactively-coded biplot of Figure 2. Table 9 reports the percentage contribution for the *Ratings* variable. The contribution of the linear polynomial, α_1 , is $z_{1\bullet\bullet}^2 / X^2 \times 100 = 0.63 \times 100 = 63\%$, while the quadratic polynomial is $z_{2\bullet\bullet}^2 / X^2 \times 100 = 0.21 \times 100 = 21\%$. Therefore, Figure 2 accounts for 84% of the total inertia when using the linear-by-quadratic polynomials; we shall describe this figure in more detail in Section 5.2.

4.5.2 *Contributions of the orthogonal polynomials in a single-variable polynomial biplot.* For biplots constructed under a single-variable isometry the coordinates are determined from the axes of the interactively-coded *Rankings* (column) and *Countries* (tube) variables; from them we get $V = 9$ column polynomials and $R = 2$ generalised singular vectors for the *Countries* variable. This implies that, when constructing such biplots, one has to select which of the 18 hybrid polynomial axes to portray the association. For the sake of simplicity here, the choice of these axes is determined by the size of the inertia along each axis; see Table 10.

The most dominant inertia lies along the second axis (β_1, c_1) and is calculated from the sum of the associated squared polynomial core elements: $z_{\bullet 11}^2 = \sum_{u=0}^U z_{u11}^2 = 13,929$. Therefore, its contribution to the total inertia is $13,929 / 21,266 \times 100 = 66\%$. Note this second axis is the hybrid polynomial computed using the linear polynomial of *Rankings* combined with the first generalised

TABLE 8
Polynomial core array: z_{uvr}

Countries: Tucker3 component r_1									
Rankings									
Ratings	v_0	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
0th order: u_0	0.00	1.33	-1.02	0.18	-0.29	-0.09	0.11	-0.06	-0.11
1st order: u_1	-0.52	-109.67	-25.92	-12.07	-3.21	-1.90	1.59	-0.30	-1.61
2nd order: u_2	0.95	43.21	-42.81	-13.76	-8.69	0.13	-1.98	-0.16	0.89
3th order: u_3	-0.38	-2.78	25.32	-14.34	-0.27	-5.09	-0.89	1.01	2.34
4th order: u_4	0.39	-4.22	-12.72	6.39	-11.00	3.36	-2.27	-0.99	-2.03
5th order: u_5	0.03	1.48	1.86	-6.16	-1.52	-10.87	5.18	-0.20	-1.64
6th order: u_6	-0.12	0.07	2.72	4.54	-1.55	-2.85	-11.64	3.29	1.17
7th order: u_7	0.12	-1.81	-1.67	-0.78	1.32	1.63	-2.31	-10.87	1.46
8th order: u_8	-0.26	-1.21	-0.41	-0.77	-1.60	-0.24	0.19	-2.41	-11.28

Countries: Tucker3 component r_2									
Rankings									
Ratings	v_0	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8
0th order: u_0	0.00	-5.24	2.53	-3.42	2.44	-2.53	0.98	-0.63	0.41
1st order: u_1	-13.85	-16.65	-8.93	2.03	-1.01	0.42	-0.35	0.43	0.91
2nd order: u_2	6.37	-12.20	-10.32	-9.75	-0.40	-0.28	-0.38	0.18	-1.06
3th order: u_3	-5.47	3.08	0.72	-2.45	-4.45	-1.07	0.29	-0.63	1.28
4th order: u_4	4.09	-4.49	1.96	3.74	0.79	-2.83	0.87	0.05	-0.54
5th order: u_5	-4.87	2.55	0.74	-1.20	1.68	3.00	-1.95	-0.58	-0.42
6th order: u_6	1.86	-2.19	0.46	0.15	-2.09	-1.77	2.56	0.67	0.62
7th order: u_7	-0.63	1.66	-0.11	-1.32	0.36	0.09	-0.99	0.39	-1.28
8th order: u_8	0.73	0.48	0.73	0.01	0.36	0.33	-0.09	1.43	2.48

TABLE 9
Percentage contributions of the rating polynomials to the total explained inertia

Row polynomials									
α_0	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8	% Inertia
axis1	axis2	axis3	axis4	axis5	axis6	axis7	axis8	axis9	
0	63	21	4	2	1	1	1	1	94

TABLE 10
Percentage contributions of the hybrid polynomials in the single-variable polynomial biplot

Hybrid column-tube polynomials									
β_0, c_1	β_1, c_1	β_2, c_1	β_3, c_1	β_4, c_1	β_5, c_1	β_6, c_1	β_7, c_1	β_8, c_1	
axis1	axis2	axis3	axis4	axis5	axis6	axis7	axis8	axis9	
0	66	16	3	1	1	1	1	1	

β_0, c_2	β_1, c_2	β_2, c_2	β_3, c_2	β_4, c_2	β_5, c_2	β_6, c_2	β_7, c_2	β_8, c_2	% Inertia
axis10	axis11	axis12	axis13	axis14	axis15	axis16	axis17	axis18	
1	2	1	0	0	0	0	0	0	94

vector of *Countries*. Similarly, along the third axis, where (β_2, c_1) , the inertia is calculated from the sum of the associated squared polynomial core elements: $z_{\bullet 21}^2 = \sum_{u=0}^U z_{u21}^2 = 3322$ and its contribution to the total inertia is $3322/21,266 \times 100 = 16\%$ (see Tables 8 and 10). Therefore, Figure 3 accounts for 82% of the total inertia when using the linear-by-quadratic hybrid polynomials. Finally, note that the 11th hybrid axis (β_1, c_2) is given by the linear polynomial of *Rankings* combined with the second generalised vector of *Countries* and its contribution to the total inertia is $309/21,266 \times 100 = 2\%$. Figure 4 portrays the association when using the linear-by-linear hybrid polynomials and accounts for 68% of the total inertia. Figures 3 and 4 will be described in more detail in the next section.

5. GRAPHING ASSOCIATION

As indicated above, numerical summaries are seldom enough for understanding complex relationships and this is certainly the case with three-way data. We argued above that polynomial biplots are suitable graphical tools to visualise the symmetrical or non-symmetrical association for ordered three-way contingency tables. In this section, we present examples of both the interactively-coded and the single-variable polynomial biplots to understand this dependence structure. To put the contribution of the polynomial biplots into perspective, in Section 5.1, we first visualise the association between the rating-ranking data after carrying out a classical symmetrical three-way correspondence analysis based on the Tucker3 decomposition. Then in Sections 5.2 and 5.3, we illustrate the two kinds of polynomial biplots (described in Section 3.4) based on the hybrid decomposition resulting from using the orthogonal polynomials for the ordered row and column variables (*Ratings* and *Rankings*) and the Tucker3 components for the nominal tube variable (*Countries*). It is also worth noting that in both cases these graphs display the complete information on the total association and not just one of the two-way association term (or even just the three-way association term). This implies that one may also focus on details in those association terms, even though we do not do so in the present example, since our primary interest is in describing the proposed methodology.

5.1 Biplots for the Classical Symmetrical Three-Way Correspondence Analysis

Figure 1 shows the association of *Ratings* and *Rankings* across *Countries* categories when using the Tucker3 components for the row, column and tube variables; see also Van Herk and Van de Velden (2007). The dimensionality of the space for their analysis was chosen to be $2 \times 2 \times 1$ and was based on the amount of explained variance when adding/deleting dimensions. Notwithstanding the dimension reduction, the resulting explained inertia

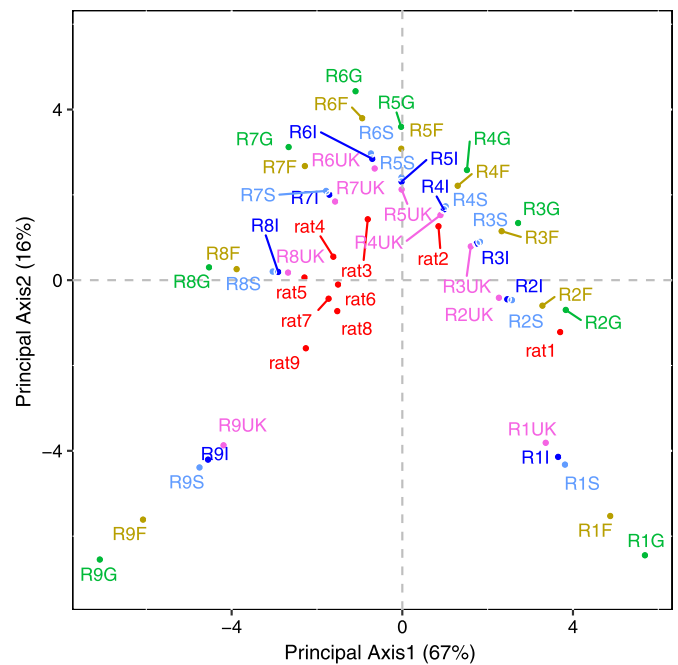


FIG. 1. Row isometric biplot according to classical symmetrical three-way correspondence analysis.

is 17,726, that is, 83% of the total chi-squared statistic. Having only two row components, the first two axes of the interactively-coded biplot represent 83% of the total inertia. Here, the *Rankings* and *Countries* categories are interactively-coded and are visually represented by standard coordinates while the *Ratings* categories are depicted using principal coordinates. In classical three-way correspondence analysis of Figure 1, the origin of the plot represents the independence among the variables.

In Figure 1, the point labelled *rat9* represents the rating score 9. The points labelled *R9G*, *R9F*, *R9S*, *R9I*, *R9UK* represent the rank of 9 given in Germany, France, Italy, Spain and UK, respectively. It can be seen that the items considered least important (i.e., having ratings/rankings of 1 and 2) are located on the right-hand side of the plot, while the items considered most important (i.e., those with ratings/rankings of 8 and 9) are located on the left-hand side of the plot. Indeed, a hierarchy from least to most important can be clearly distinguished when looking at the ranking scores across the five countries, but it is less evident when observing the rating values. On the right side of the plot, we can see only *rat1* and *rat2*, while the remaining rating categories are on the left side of Figure 1. Therefore, there is a clear inconsistency between the rankings of 6 and 7, and the rating of 3. As noted by Van Herk and Van de Velden (2007), the distances between scale points are not equal for either rating or ranking points. For example, the distance between rating points 8 and 9 is smaller than that between rating points 1 and 2. Furthermore, it can be observed that countries have been placed in the same order for each ranking score; Germany is located furthest from the origin, the UK is located closest

to the origin and the other countries lie in between the two. Figure 1 also shows more consistency between the high ranking score of Germany, France, Italy, Spain and UK ($R9G, R9F, R9S, R9I, R9UK$) and the high rating category, $rat9$. The ranking-country association suggests possible differences in response patterns between countries. Indeed, there are also differences in the use of ranking between countries. Participants in Germany (points $R1G$ through to $R9G$) show more discrimination in their rankings than participants in, for example, Italy or UK, whose points are closer to the origin of the plot.

5.2 Interactively-Coded Polynomial Biplot for the Symmetrical Hybrid Decomposition

As explained above, the polynomial interactively-coded biplot has the same structure as its classic version, but now the axes are defined in terms of the orthogonal polynomials. Thus a comparison between the two rests on the interpretation of the axes and on the meaning of the origin of the plot.

In the polynomial biplot of Figure 2, the *Ratings* variable is depicted using standardised coordinates, while the interactively coded categories of *Rankings* and *Countries* are depicted using principal coordinates; the polynomial axes of Figure 2 account for about 84% of the total explained inertia (see Table 8 and the inertia of row polynomials in Section 4.5.1). The horizontal polynomial axis of Figure 2 reflects the linear polynomial of the *Ratings* variable and is weighted with respect to the row proportions. Similarly, the vertical polynomial axis reflects the quadratic polynomial of the variable and is also weighted with respect to its marginal proportions. As the data are

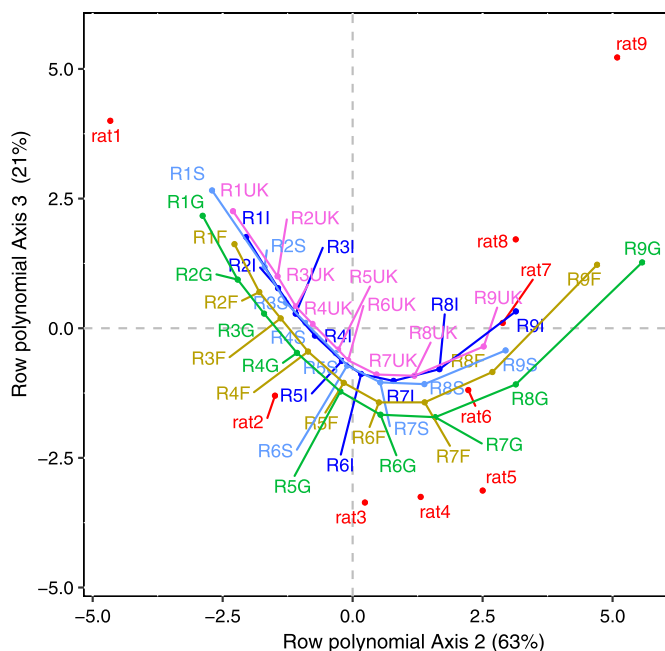


FIG. 2. Interactively-coded polynomial biplot from an ordinal symmetrical three-way correspondence analysis.

centred at the origin of the plot, and the rows are depicted using standard coordinates, the origin of the plot corresponds to the position of the mean rating; which is shown to be 3 ($rat3$). We note that the categories of the *Ratings* variable do not follow a perfect parabolic curve since there are large differences in the linearity (or mean values) of its categories. Based on the principal coordinates only, distances between the interactively-coded *Rankings-Countries* categories can be measurable and interpretable. For example, the highest ranking values for Germany and France, $R9G$ and $R9F$, are close to each other showing the greatest consistency than any other *Rankings-Countries* pair.

When examining the rankings, the ranks of 4, 5, and 6 across all of the countries are associated with a low rating value of 2 and 3, showing that there exist inconsistencies in how the different countries rate. The mean value of the two lowest rating levels, $rat1$ and $rat2$, are less than the overall mean rating (since these categories are located on the left side of the plot) while $rat4$ through to $rat9$ have a mean value that is higher than the overall mean.

We can also observe some differences across countries in terms of ranking variability. For example, rankings are higher for Germany and France (note the large coordinates on the second axis for $R9G$ and $R9F$) while the rankings are lower for Spain, Italy and UK. Given the large coordinate values along the first and second axis, note that there is a strong association between the highest level of rating ($rat9$) and the highest level of ranking given by Germany and France ($R9G$ and $R9F$). One can also see that there exists an association between the lowest level of rating ($rat1$) and the lowest level of ranking for Germany, Spain and UK ($R1G, R1S$ and $R1UK$). Figure 2 also shows that the increasing rating levels are associated with increasing levels of ranking, but not at the same level (e.g., $rat3$ is strongly associated with $R6G, R6F$). This is consistent with the two-way association terms summarised in Table 1; recall that the inertia for *Ratings* \times *Rankings* and *Ratings* \times *Countries* accounted for 86% and 3%, respectively.

5.3 Single-Variable Polynomial Biplot for the Symmetrical Hybrid Decomposition

The biplots constructed under a single-variable isometry are interpreted in a similar way to interactively-coded biplots. However, it is the basis of the space in which the categories of the variables are displayed that is different. The variable coordinates are now determined by the axes of the interactively-coded *Rankings* and *Countries* variables.

5.3.1 *Linear-by-quadratic polynomial biplots.* For the single-variable polynomial biplot of Figure 3, the horizontal axis consists of the combination of the linear *Rankings* polynomial, β_1 , and the first generalised axis

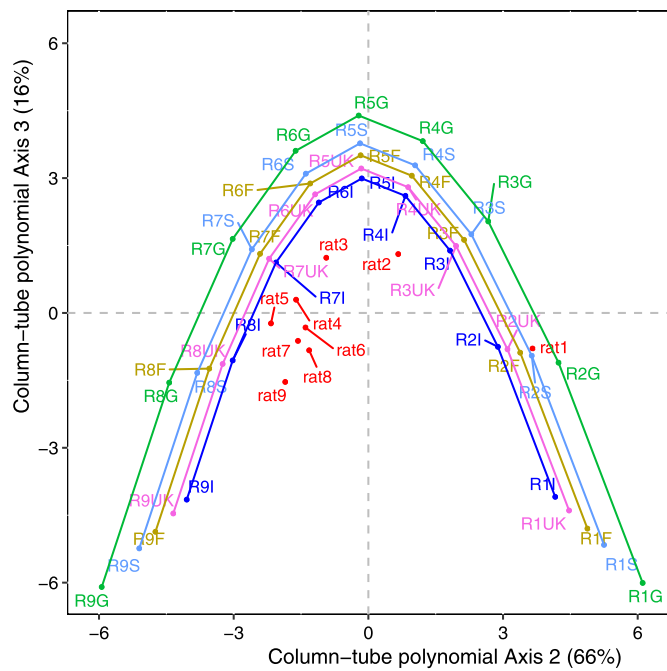


FIG. 3. Single-variable polynomial biplot from an ordinal symmetrical three-way correspondence analysis.

(Tucker3 component) for *Countries*, c_1 , while the vertical axis stems from the combination of the quadratic *Rankings* polynomial, β_2 , and, again, the first generalised *Countries* axis, c_1 . Therefore, we note that there are parabolic trends between the interactively-coded column-tube variables and the row variable (see Section 4.5.2). In Figure 3, for all countries, the parabolas are concave. Since the narrowness of the parabolas is determined by the size of the quadratic polynomial coordinates of *Rankings*, higher coordinates produce narrow parabolic trends (e.g., UK and Italy) and coordinate values close to zero lead to flat parabolas. Here, unlike the interactively-coded biplot of Figure 2, the *Rankings-Countries* categories are depicted using standard polynomial coordinates and the rating categories are depicted using principal polynomial coordinates. Therefore, only distances between rating category points can be properly assessed. As the column-tube categories have standardised coordinates in Figure 3, the origin of the plot represents the mean of the rankings; it shows that the mean ranking across all *Countries* is 5 (R_5). This mean ranking is strongly associated with ratings 2 and 3 (rat_2 and rat_3).

As we described for Figure 1, it can be observed from Figure 3 that the countries have been placed in the same order for each ranking score. However in contrast to Figure 1, participants in Italy show less discrimination in their rankings as Italy is located closest to the origin followed by UK, France, Spain and Germany, where the latter is always the furthest country from the origin. Furthermore, given the isometry of the *Ratings* variable, Figure 3 shows that the distance of rat_9 from rat_8 is much shorter

than the distance of rat_1 from rat_2 . Therefore the participants show less discrimination in their high ratings than in the lowest ratings. Figure 3 accounts for 82% of the total inertia which is commensurate with Figure 2.

Note that, for all countries, say for example Germany (G), the *Ranking* categories lie in a linearly decreasing order along the horizontal axis of Figure 3. The low levels of *Rankings*, R_1 and R_2 , across the countries are primarily associated with the low level of *Ratings* (rat_1) that lie on the right side of the horizontal axis below the mean of the linear polynomial of *Rankings*. Furthermore, items considered least important with a rating of 2 (rat_2) are located on the right-hand side of the biplot, showing a strong association with R_4 , while ratings of 6, 7, 8 and 9 (rat_6 , rat_7 , rat_8 , rat_9) are located on the left of the plot, above the mean of the linear polynomial of *Rankings* (R_5). However, items with ratings of 3, 4 and 5 (rat_3 , rat_4 , rat_5) are on the same side of the plot showing some inconsistencies between rankings and ratings, and appear more evident than what Figure 1 shows.

5.3.2 Linear-by-linear polynomial biplots. Figure 4 is the polynomial biplot of linear-by-linear hybrid polynomials. Therefore, the configuration of the plot is akin to the spokes of a cartwheel, where all lines go through the origin due to the centring of the orthogonal polynomials (see Section 3.4.3). In Figure 4, the horizontal axis is the combination of the linear ranking polynomial, β_1 , and the first generalised singular axis (Tucker3 component) for *Countries*, c_1 , while the vertical axis is the combination of the linear *Rankings* polynomial, β_1 , and the second generalised singular *Countries* axis, c_2 . From Table 7, we

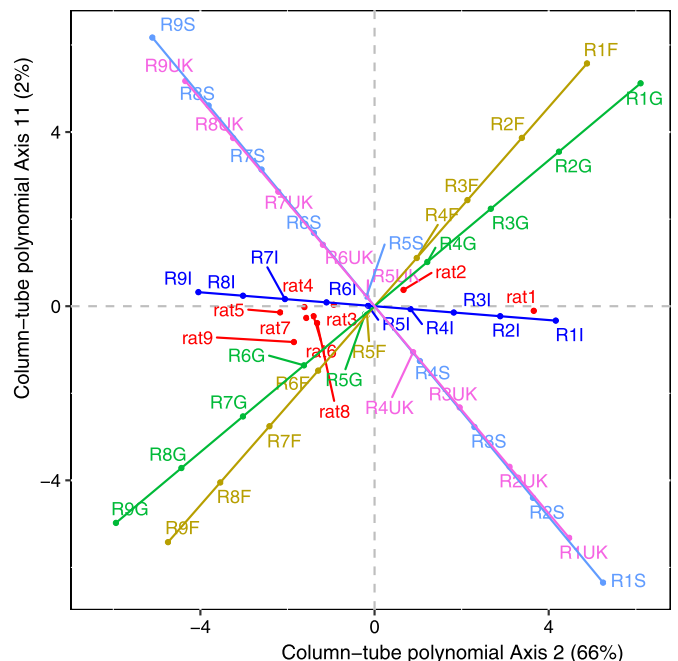


FIG. 4. Single-variable polynomial biplot by ordinal symmetrical three-way correspondence analysis.

can see that, on this second Tucker3 axis, UK and Spain have a larger score than they do on axis c_1 and are opposite in sign, while France and Germany have the same score sign on the two Tucker3 axes, c_1 and c_2 . Indeed, we also observe that the score of Italy along c_2 is close to zero, and so the ranking of Italy lies almost adjacent to the first axis. All ranking categories are linearly ordered on both the horizontal and vertical axis, while the interactively coded categories $R5G$, $R5F$, $R5S$, $R5I$, $R5UK$ all have a zero coordinate as the mean ranking is 5.

As in Figure 3, the explained inertia along the horizontal axis is 66%; see also Table 10. Along the vertical axis, the explained inertia is 2%. Figure 4 accounts for 68% of the total inertia. This reduction in quality, when compared with Figures 2 and 3, can be explained because of the large number of polynomials required to optimally depict the association. However, Figure 4 does provide an interesting representation of the linearity of the interactively-coded ranking-country categories.

In the interactively-coded biplot of Figure 4, we observe that all categories for Spain and UK lie on the same line and the ranking categories ($R9$ through $R1$) lie in a linearly decreasing order from the top-left to the bottom-right. We also note a reversal of ordering of rankings for Germany and France, when compared to Spain and UK; see Table 7.

In contrast to Figure 1, when only considering the linearity of ranking categories (ignoring its variability unlike Figure 2), Figure 4 shows that the participants of Spain and UK are very consistent in their rankings of items. It also shows that the participants in Germany and France have similar rankings. However, Italian participants exhibit very little variation in how they rank, highlighting a very different approach when ranking items.

6. CONCLUSION

The methodology of both symmetrical and non-symmetrical ordered three-way correspondence analysis has been discussed with the aim of showing how both ordering and dependency can be incorporated into the modelling of the association of the three variables using three-way correspondence analysis. The major benefits of including such design aspects into the analysis include its ability to incorporate pre-existing information about the data in the modelling phase.

Including the order of the categories has been achieved by replacing the singular vectors with orthogonal polynomials. Generating orthogonal polynomials (Favard, 1935, Chihara, 1978, 1990) using the recurrence formulae described for categorical variables by Emerson (1968) has been shown to be especially useful in this respect. These polynomials provide information about the sources of variation in the variables by examining their linear, quadratic or higher-order trends, which is difficult to do

using the Tucker3 approach or other similar nominally structured methods of three-way decomposition.

The orthogonal polynomials (and/or the singular vectors) used for the partition of the three-way measures of association are computed here for the *complete* model. However, the axes derived from these polynomial/singular vectors are the same as those used in the decomposition of the various association terms derived from the partition. To further understand what is gained by considering ordered and/or hybrid variants of three-way correspondence analysis, further investigations can be made to analyse those association terms that come from the familywise decompositions of Pearson's Φ^2 statistic (Lombardo, Takane and Beh, 2020).

We have described how the technical issues of the variants of three-way correspondence analysis presented in this paper can be applied to a three-way contingency table with a non-symmetrical association structure. A comprehensive discussion of all these variants including the issue of the best model dimension choice shall be left for future consideration.

ACKNOWLEDGMENTS

The authors thank Hester van Herk and Michel van de Velden for the data and the anonymous reviewers for their helpful comments.

REFERENCES

- AGRESTI, A. and GOTTARD, A. (2007). Independence in multi-way contingency tables: S. N. Roy's breakthroughs and later developments. *J. Statist. Plann. Inference* **137** 3216–3226. MR2363251 <https://doi.org/10.1016/j.jspi.2007.03.006>
- ALWIN, D. F. and KROSNICK, J. A. (1985). The measurement of values in surveys: A comparison of ratings and rankings. *Public Opin. Q.* **49** 535–552.
- BEH, E. J. (1997). Simple correspondence analysis of ordinal cross-classifications using orthogonal polynomials. *Biom. J.* **39** 589–613.
- BEH, E. J. (1998). A comparative study of scores for correspondence analysis with ordered categories. *Biom. J.* **40** 413–429.
- BEH, E. J. and DAVY, P. J. (1998). Partitioning Pearson's chi-squared statistic for a completely ordered three-way contingency table. *Aust. N. Z. J. Stat.* **40** 465–477. MR1664197 <https://doi.org/10.1111/1467-842X.00050>
- BEH, E. J. and DAVY, P. J. (1999). Partitioning Pearson's chi-squared statistic for a partially ordered three-way contingency table. *Aust. N. Z. J. Stat.* **41** 233–246. MR1705401 <https://doi.org/10.1111/1467-842X.00077>
- BEH, E. J. and LOMBARDO, R. (2014). *Correspondence Analysis: Theory, Practice and New Strategies*. Wiley Series in Probability and Statistics. Wiley, Chichester. MR3288854 <https://doi.org/10.1002/9781118762875>
- BEH, E. J. and LOMBARDO, R. (2019). Multiple and multiway correspondence analysis. *Wiley Interdiscip. Rev.: Comput. Stat.* **11** e1464. MR3999531 <https://doi.org/10.1002/wics.1464>
- BEH, E. J. and LOMBARDO, R. (2021). Features of the polynomial biplot of ordered contingency tables. *J. Comput. Graph. Statist.* To appear.

- BEH, E. J., SIMONETTI, B. and D'AMBRA, L. (2007). Partitioning a non-symmetric measure of association for three-way contingency tables. *J. Multivariate Anal.* **98** 1391–1411. MR2364125 <https://doi.org/10.1016/j.jmva.2007.01.011>
- BENZÉCRI, J.-P. (1973). *L'analyse des Données. Vols. I, II*. Dunod, Paris.
- BENZÉCRI, J. P. (1977). Histoire et préhistoire de l'analyse des données. Partie V: L'analyse des correspondances. *Cah. Anal. Données* **2** 9–40.
- BÖCKENHOLT, U. and BÖCKENHOLT, I. (1990). Canonical analysis of contingency tables with linear constraints. *Psychometrika* **55** 633–649.
- BREIMAN, L. (2001). Statistical modeling: The two cultures. *Statist. Sci.* **16** 199–231. MR1874152 <https://doi.org/10.1214/ss/1009213726>
- CARLIER, A. and KROONENBERG, P. M. (1996). Biplots and decompositions in two-way and three-way correspondence analysis. *Psychometrika* **61** 355–373.
- CARLIER, A. and KROONENBERG, P. M. (1998). Three-way correspondence analysis: The case of the French cantons. In *Visualization of Categorical Data* (J. Blasius and M. Greenacre, eds.) 253–275. Academic Press, New York.
- CARROLL, J. D. and CHANG, J. J. (1970). Analysis of individual differences in multidimensional scaling via an n -way generalization of Eckart–Young decomposition. *Psychometrika* **35** 283–319.
- CEULEMANS, E. and KIERS, H. A. L. (2006). Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method. *Br. J. Math. Stat. Psychol.* **59** 133–150. MR2246998 <https://doi.org/10.1348/000711005X64817>
- CEULEMANS, E. and KIERS, H. A. L. (2009). Discriminating between strong and weak structures in three-mode principal component analysis. *Br. J. Math. Stat. Psychol.* **62** 601–620. MR2750421 <https://doi.org/10.1348/000711008X369474>
- CEULEMANS, E., TIMMERMAN, M. and KIERS, H. A. L. (2011). The CHULL procedure for selecting among multilevel component solutions. *Chemom. Intell. Lab. Syst.* **106** 12–20.
- CHIHARA, T. S. (1978). *An Introduction to Orthogonal Polynomials. Mathematics and Its Applications* **13**. Gordon and Breach Science Publishers, New York. MR0481884
- CHIHARA, T. S. (1990). The three term recurrence relation and spectral properties of orthogonal polynomials. In *Orthogonal Polynomials (Columbus, OH, 1989)*. NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **294** 99–114. Kluwer Academic, Dordrecht. MR1100289
- CHOUKAKIAN, V. (1988). Exploratory analysis of contingency tables by loglinear formulation and generalizations of correspondence analysis. *Psychometrika* **53** 235–250. MR0955468 <https://doi.org/10.1007/BF02294135>
- CLOGG, C. C. (1982). Some models for the analysis of association in multiway cross-classifications having ordered categories. *J. Amer. Statist. Assoc.* **77** 803–815. MR0686403
- D'AMBRA, L. and LAURO, N. (1989). Nonsymmetrical analysis of three-way contingency tables. In *Multiway Data Analysis (Rome, 1988)* 301–315. North-Holland, Amsterdam. MR1088959
- EMERSON, P. L. (1968). Numerical construction of orthogonal polynomials from general recurrence formula. *Biometrics* **24** 696–701.
- FAVARD, J. (1935). Sur les polynomes de Tchebicheff. *C. R. Acad. Sci.* **200** 2052–2053.
- GINI, C. (1912). *Variabilità e Mutuabilità [Variability and Mutability]*. Contributo Allo Studio delle Distribuzioni e delle Relazioni Statistiche. Cuppini Press, Bologna, Italy.
- GOODMAN, L. A. and KRUSKAL, W. H. (1954). Measures of association for cross classifications. *J. Amer. Statist. Assoc.* **49** 732–764.
- GOWER, J. C., LE ROUX, N. J. and GARDNER-LUBBE, S. (2016). Biplots: Qualitative data. *Wiley Interdiscip. Rev.: Comput. Stat.* **8** 82–111. MR3466000 <https://doi.org/10.1002/wics.1377>
- GREENACRE, M. J. (1984). *Theory and Applications of Correspondence Analysis*. Academic Press [Harcourt Brace Jovanovich, Publishers], London. MR0767260
- GREENACRE, M. J. (1990). Some limitations of multiple correspondence analysis. *Comput. Stat. Q.* **3** 249–256.
- GREENACRE, M. J. (2017). *Correspondence Analysis in Practice*, 3rd ed. CRC Press/CRC, Barcelona.
- GREENACRE, M. J. and BLASIUS, J. (2006). *Multiple Correspondence Analysis and Related Methods*. Chapman & Hall/CRC, Boca Raton.
- HARSHMAN, R. A. (1970). Foundation of the PARAFAC procedure: Models and conditions for an explanatory multi-modal factor analysis. *UCLA Work. Pap. Phon.* **16** 1–84.
- KAHLE, L. R. (1983). *Social Values and Social Change: Adaptation to Life in America*. Praeger, New York.
- KATERI, M. (2014). *Contingency Table Analysis: Methods and Implementation Using R. Statistics for Industry and Technology*. Birkhäuser/Springer, New York. MR3290014 <https://doi.org/10.1007/978-0-8176-4811-4>
- KIERS, H. A. L. (1989). *Three-Way Methods for the Analysis of Qualitative and Quantitative Two-Way Data*. DSWO Press, Leiden, NL.
- KIERS, H. A. L. (2000). Towards a standardized notation and terminology in multiway analysis. *Chemom. Intell. Lab. Syst.* **14** 105–122.
- KROONENBERG, P. M. (1983). *Three-Mode Principal Component Analysis: Theory and Applications*. DSWO Press, Leiden, NL. (Errata, 1989; available from the author).
- KROONENBERG, P. M. (1992). Multilinear models: Applications in spectroscopy: Comment: PARAFAC in three-way land. *Statist. Sci.* **7** 312–314.
- KROONENBERG, P. M. (2008). *Applied Multiway Data Analysis. Wiley Series in Probability and Statistics*. Wiley Interscience, Hoboken, NJ. MR2378349 <https://doi.org/10.1002/9780470238004>
- KROONENBERG, P. M. (2014). History of multiway component analysis and three-way correspondence analysis. In *Visualization and Verbalization of Data* (M. Greenacre and J. Blasius, eds.) 78–93. CRC Press, Boca Raton, FL.
- KROONENBERG, P. M. (2020). Multiway extensions of the SVD. In *Advanced Studies in Behaviormetrics and Data Science* (T. Imaizumi, A. Nakayama and S. Yokoyama, eds.) 141–157. Springer, Singapore.
- KROONENBERG, P. M. and ANDERSON, C. J. (2006). Additive and multiplicative models for three-way contingency tables: Darroch (1974) revisited. In *Multiple Correspondence Analysis and Related Methods* (M. Greenacre and J. Blasius, eds.) 455–486. Chapman & Hall, London.
- KROONENBERG, P. M. and DE LEEUW, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. *Psychometrika* **45** 69–97. MR0570771 <https://doi.org/10.1007/BF02293599>
- KROONENBERG, P. M. and LOMBARDO, R. (1999). Non-symmetric correspondence analysis: A tool for analysing contingency tables with a dependence structure. *Multivar. Behav. Res.* **34** 367–397.
- KROONENBERG, P. M. and OORT, F. J. (2003). Three-mode analysis of multimode covariance matrices. *Br. J. Math. Stat. Psychol.* **56** 305–335. MR2101772 <https://doi.org/10.1348/000711003770480066>
- LANCASTER, H. O. (1951). Complex contingency tables treated by the partition of χ^2 . *J. Roy. Statist. Soc. Ser. B* **13** 242–249. MR0051473

- LAURO, N. and D'AMBRA, L. (1984). L'analyse non symétrique des correspondances. In *Data Analysis and Informatics, III (Versailles, 1983)* 433–446. North-Holland, Amsterdam. MR0787651
- LIGHT, R. J. and MARGOLIN, B. H. (1971). An analysis of variance for categorical data. *J. Amer. Statist. Assoc.* **66** 534–544. MR0494714
- LIN, Y. and ZHANG, H. H. (2006). Component selection and smoothing in multivariate nonparametric regression. *Ann. Statist.* **34** 2272–2297. MR2291500 <https://doi.org/10.1214/009053606000000722>
- LOISEL, S. and TAKANE, Y. (2016). Partitions of Pearson's chi-square statistic for frequency tables: A comprehensive account. *Comput. Statist.* **31** 1429–1452. MR3573085 <https://doi.org/10.1007/s00180-015-0619-1>
- LOMBARDO, R. (1994). *Modelli di Decomposizione per L'analisi della Dipendenza Nelle Tabelle di Contingenza a Tre-Vie*. [Decomposition Models for the Analysis of Three-Way Contingency Tables]. Tesi di Dottorato di Ricerca in Statistica Computazionale e Applicazioni VI Ciclo. Università di Napoli, Italy.
- LOMBARDO, R. and BEH, E. J. (2017). Three-way correspondence analysis for ordinal-nominal variables. In *SIS 2017 Statistics and Data Science: New Challenges, New Generations, 28–30 June 2017 Florence (Italy). Proceedings of the Conference of the Italian Statistical Society* 613–620. ISBN:978-88-6453-521-0.
- LOMBARDO, R., BEH, E. J. and KROONENBERG, P. M. (2016). Modelling trends in ordered correspondence analysis using orthogonal polynomials. *Psychometrika* **81** 325–349. MR3505369 <https://doi.org/10.1007/s11336-015-9448-y>
- LOMBARDO, R., CARLIER, A. and D'AMBRA, L. (1996). Nonsymmetric correspondence analysis for three-way contingency tables. *Methodologica* **4** 59–80.
- LOMBARDO, R., TAKANE, Y. and BEH, E. J. (2020). Family-wise decompositions of Pearson's chi-square statistic in the analysis of contingency tables. *Adv. Data Anal. Classif.* **14** 629–649. MR4150431 <https://doi.org/10.1007/s11634-019-00374-7>
- LORENZO-SEVA, U., TIMMERMAN, M. E. and KIERS, H. AL. (2011). The Hull method for selecting the number of common factors. *Multivar. Behav. Res.* **46** 340–364. 2011. <https://doi.org/10.1080/00273171.2011.564527>
- MARCOTORCHINO, F. (1985). *Utilisation des Comparaisons Par Paires en Statistique des Contingencies: Partie III*. [Use of Paired Comparisons in Contingency Statistics. Part III.] Étude du Centre Scientifique No F 081. IBM, Paris, France.
- NISHISATO, S. (1980). *Analysis of Categorical Data: Dual Scaling and Its Applications. Mathematical Expositions* **24**. Univ. Toronto Press, Toronto, Canada. MR0600656
- RAYNER, J. C. W. and BEH, E. J. (2009). Towards a better understanding of correlation. *Stat. Neerl.* **63** 324–333. MR2751894 <https://doi.org/10.1111/j.1467-9574.2009.00425.x>
- RAYNER, J. C. W. and BEST, D. J. (1996). Smooth extensions of Pearson's product moment correlation and Spearman's rho. *Statist. Probab. Lett.* **30** 171–177. MR1417004 [https://doi.org/10.1016/0167-7152\(95\)00216-2](https://doi.org/10.1016/0167-7152(95)00216-2)
- RODRIGUE, N., GUILLET, M., FORTIN, J. and MARTIN, J. F. (2000). Comparing information obtained from ranking and descriptive tests of four sweet corn products. *Food Qual. Prefer.* **11** 47–54.
- SMILDE, A. K., BRO, R. and GELADI, P. (2004). *Multi-Way Analysis: Applications in the Chemical Sciences*. Wiley, Chichester.
- TAKANE, Y. and JUNG, S. (2009). Tests of ignoring and eliminating in nonsymmetric correspondence analysis. *Adv. Data Anal. Classif.* **3** 315–340. MR2570509 <https://doi.org/10.1007/s11634-009-0054-7>
- TAKANE, Y., YANAI, H. and MAYEKAWA, S. (1991). Relationships among several methods of linearly constrained correspondence analysis. *Psychometrika* **56** 667–684. MR1149337 <https://doi.org/10.1007/BF02294498>
- TUCKER, L. R. (1966). Some mathematical notes on three-mode factor analysis. *Psychometrika* **31** 279–311. MR0205395 <https://doi.org/10.1007/BF02289464>
- VAN HERK, H. and VAN DE VELDEN, M. (2007). Insight into the relative merits of rating and ranking in a cross-national context using three-way correspondence analysis. *Food Qual. Prefer.* **18** 1096–1105.
- VAN DER HEIJDEN, P. G. M., DE FALGUEROLLES, A. and DE LEEUW, J. (1989). A combined approach to contingency table analysis using correspondence analysis and log-linear analysis. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **38** 249–292. MR1001477 <https://doi.org/10.2307/2348058>
- WAHBA, G. (1990). *Spline Models for Observational Data. CBMS-NSF Regional Conference Series in Applied Mathematics* **59**. SIAM, Philadelphia, PA. MR1045442 <https://doi.org/10.1137/1.9781611970128>
- WAHBA, G., WANG, Y., GU, C., KLEIN, R. and KLEIN, B. (1995). Smoothing spline ANOVA for exponential families, with application to the Wisconsin Epidemiological Study of Diabetic Retinopathy. *Ann. Statist.* **23** 1865–1895. MR1389856 <https://doi.org/10.1214/aos/1034713638>