Food Quality and Preference 112 (2023) 105028

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Why use component-based methods in sensory science?

Tormod Næs [a,*], Paula Varela [a], John C. Castura [b], Rasmus Bro [c], Oliver Tomic [d]

[a] Nofima, Ås, Norway
[b] Compusense Inc., 255 Speedvale Av. W., Guelph, Ontario N1H 1C5, Canada
[c] Dept. Food Science, Univ. Copenhagen, Denmark
[d] Realtek, University of Life Sciences, Ås, Norway

ARTICLE INFO

Keywords:
Component-based method
Principal component analysis (PCA)
Partial least squares regression (PLSR, PLS)
Multiple factor analysis (MFA)
Parallel factor analysis (PARAFAC)
Temporal check-all-that-apply (TCATA)
Projective mapping (PM)
Quantitative descriptive analysis (QDA)

ABSTRACT

This paper discusses the advantages of using so-called component-based methods in sensory science. For instance, principal component analysis (PCA) and partial least squares (PLS) regression are used widely in the field; we will here discuss these and other methods for handling one block of data, as well as several blocks of data. Component-based methods all share a common feature: they define linear combinations of the variables to achieve data compression, interpretation, and prediction. The common properties of the component-based methods are listed and their advantages illustrated by examples. The paper equips practitioners with a list of solid and concrete arguments for using this methodology.

1. Introduction

A *component-based method* or just *component method* is a method in which a low number of components or latent variables are used to approximate the original multivariate data (data reduction). These methods are often used when there are only a few underlying sources of variation even though many attributes have been measured. These data reduction methods provide component scores and loadings where the former represent the 'new variables' (or coordinates) defined as combinations of the original variables and the latter describe how the components themselves relate to the original variables. Typical examples of component methods used in sensory science are principal component analysis (PCA; Hotelling (1933) and Pearson, (1901)) partial least squares regression (PLSR/PLS; see e.g. Wold & Martens (1983) and Martens & Næs (1989)), and correspondence analysis (CA; Hirschfeld (1935)). Through visualization, e.g. in scatter plots, the components lend themselves to discovery and/or confirmation of relationships/trends across samples which can be easily communicated within a research and development team, and, with care to avoid misinterpretation, also outside of the sensory community (i.e. marketing, leadership in a company, or to researchers in other disciplines).

Component methods are used in sensory science for various types of data sets and situations. For a long time, in particular PCA and variants thereof have been used extensively for understanding and interpreting variability in data sets from, for instance, QDA (Stone et al. (1974), Lawless et al. (2010)) with trained panels and consumer liking studies. The focus has been on establishing a low-dimensional representation of the data which is easier to interpret than considering all the variables simultaneously, but also studies of individual differences as well as outlying and deviating assessors have been important.

With the emergence of many new sensory methods, in particular the so-called rapid descriptive profiling methods (e.g. projective mapping (PM; Risvik et al. (1994)), check-all-that-apply (CATA; see e.g. Sinopoli & Lawless (2012)), sorting) and those that take time into account, such as temporal check-all-that-apply (TCATA; Castura, Antúnez, Giménez & Ares, (2016a)) and temporal dominance of sensations (TDS; Pineau et al. (2009)), the number of possible applications has grown enormously. For rapid descriptive profiling, the focus of the component methods has generally been the same as for QDA. For some of the rapid profiling methodologies, extensions and modifications of PCA, like for instance multiple factor analysis (MFA, see Section 4.3) and correspondence analysis (CA, see Section 4.2) are needed. For the methods involving a temporal profile, a major issue is to visualise how sensory perception changes with time. A typical question is what product differences are most pronounced in the beginning or the end of the eating process.

The versatility of component methods has been demonstrated by their use in many scientific publications (see references in this paper and references therein), as well as in the industry, but as far as we know a

* Corresponding author.
E-mail address: tormod.naes@nofima.no (T. Næs).

https://doi.org/10.1016/j.foodqual.2023.105028
Received 9 May 2023; Received in revised form 17 October 2023; Accepted 18 October 2023
Available online 20 October 2023

thorough and broad presentation of pros and cons is so far unavailable. This paper addresses this challenge, gives a broad overview and discusses the main reasons why component methods are important and useful in sensory science. The ideas will be illustrated by examples. The main target group of readers is applied scientists that already have some experience in the use of one or more of these methods, but who still want a deeper understanding of their properties and also better arguments for choosing a component method. Some warnings and pitfall will be pointed out when appropriate (see also Næs et al (2020) for further discussion on these issues).

The structure of the paper is as follows. First we give a conceptual introduction to what a component method is. Then we describe some typical data structures met in sensory science where component methods are useful. After that we discuss a number of much used and typical component methods and mention sensory tests they are useful for and also point to which data structure it is useful for. Finally we give a number of illustrative examples and end up with a conclusion listing the main advantages of component methods and also some possible drawbacks.

2. What are component methods and why use them?

2.1. Component methods for one data set

Fig. 1 illustrates graphically the idea behind component methods for a single data set. A data matrix or table **X** (dimension $I \times J$, J can be smaller or larger than I) is approximated by a set of component scores **T** (dimension $I \times A$), where A is the number of components, which is smaller than (or equal to) J, and that describe (account for) the main (interesting or systematic) variability in the data set. Often the rows of **X** represent products (I) and the columns (i.e. variables) represent sensory attributes or chemical/physical properties (J), but other possibilities also exist: for instance in internal preference mapping, the columns of **X** represent different consumers.

The main idea behind component methods is that the results in **X** give an unnecessarily complex and redundant representation of the situation. Reducing the large number of variables to a few components in **T** may improve interpretation, as well as simplify and accelerate further analyses considerably. In most cases in sensory science, the columns of **X** are centred prior to analysis. Exceptions exist in for instance spectroscopy.

2.2. Scores and loadings

The components (scores) in the matrix **T** are usually defined as linear combinations of the measured variables, i.e. $\mathbf{T} = \mathbf{XW}$ (see Fig. 1) and will typically represent the variability in **X** one is primarily interested in (depending on criterion used). The matrix **W** (dimension $J \times A$) of so-called loading weights define how the components are computed. These loading weights depend on the criterion that is used, for instance variance or covariance explained.

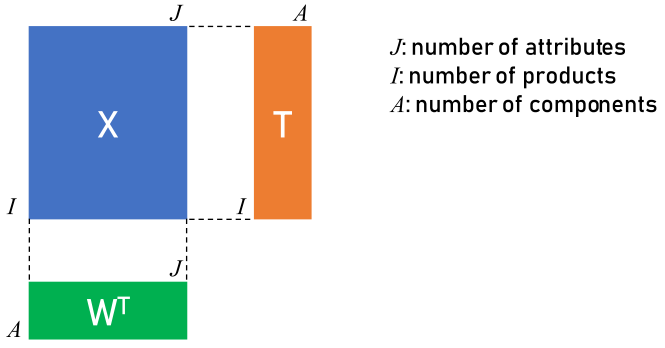


Fig. 1. The idea behind component methods. The idea is to extract a few linear combinations of **X**, i.e. scores **T**, to represent/approximate **X** as closely as possible. The loading weights **W** here describe how the linear combinations as $\mathbf{T} = \mathbf{XW}$ are obtained.

In addition to loadings weights which describe how the components are computed, it is useful to look at the loadings which describe how each of the variables in **X** relate to the components. The loadings **P** (dimension $J \times A$) are usually obtained by regressing the data set **X** onto the scores **T** using the regression equation.

$$\mathbf{X} = \mathbf{TP}^{\mathrm{T}} + \mathbf{E} \tag{1}$$

For PCA, the **P** will be the same as **W**, otherwise they will be different. The **E** represents the residuals, i.e. the part of **X** not explained by the components **T**.

The most important aspect of the model represented in equation 1 is that the scores **T** and the loadings **P (**and also **W)** can be easily visualised in scatter and line plots (see examples in Section 5) which are immensely useful for understanding complex data. The full benefit of these plots is obtained when considered together as will be discussed in the examples in Section 5. The visualization possibilities are an important reason why these methods are useful and why they are so frequently used in sensory and other sciences.

2.3. Component methods for two data sets

Component methods are also useful for decomposing more than one data set. The simplest extension is when investigating the relationship between two data sets. For instance, measurements of the same I products in a chemistry data set (input data **X**, dimension $I \times J$) with J predictor variables and a sensory data set (output, **Y**, dimension $I \times K$) with K response variables can be investigated to understand the relationships and to obtain a model for making sensory predictions based on chemistry measurements. In such cases, one is often interested in estimating the regression coefficients **B** (dimension $J \times K$) in the equation.

$$\mathbf{Y} = \mathbf{XB} + \mathbf{F} \tag{2}$$

using some type of regression method. The **F** represents the residuals with the same number of columns and rows as **Y**. If **X** has more columns than rows or the columns in **X** are collinear, then standard least squares regression can be unstable (Jolliffe, 2010). Component methods can be used in these cases. The idea is to reduce the input data **X** using a component method (equation 1) using a **W** to obtain the **T**, and then using only the components **T** in the regression equation 2 instead of the original variables **X**.

The advantages of component methods in regression are two-fold. First, the prediction equations become more stable because only a few stable components are used in the regression. Second, the information captured by the components can be visualised and interpreted using standard scatter plots.

Fig. 2 illustrates how the relationship between two data sets can be investigated using a component method. First, **T** is obtained by some criterion (using loading weights) before the **Y**- and **X**-variables are regressed onto **T** to obtain the loadings **P** and **Q** (see equation (5). The criterion for extracting components in prediction is often a compromise between prediction quality (strong correlations between input and output) and stability (high explained variance of input data, as in for instance PLS regression; see Section 5.3).

2.4. Outlier detection

In Fig. 3 is illustrated conceptually how component methods can be useful for detection of outliers, which is always an important part of a full validation of a model. The ellipse represents the area where the normal observations lie in the space spanned by **T**. The point marked with C is a 'normal' observation lying in the space spanned by **T** and well within the range of variability. Point A on the other hand is a point that

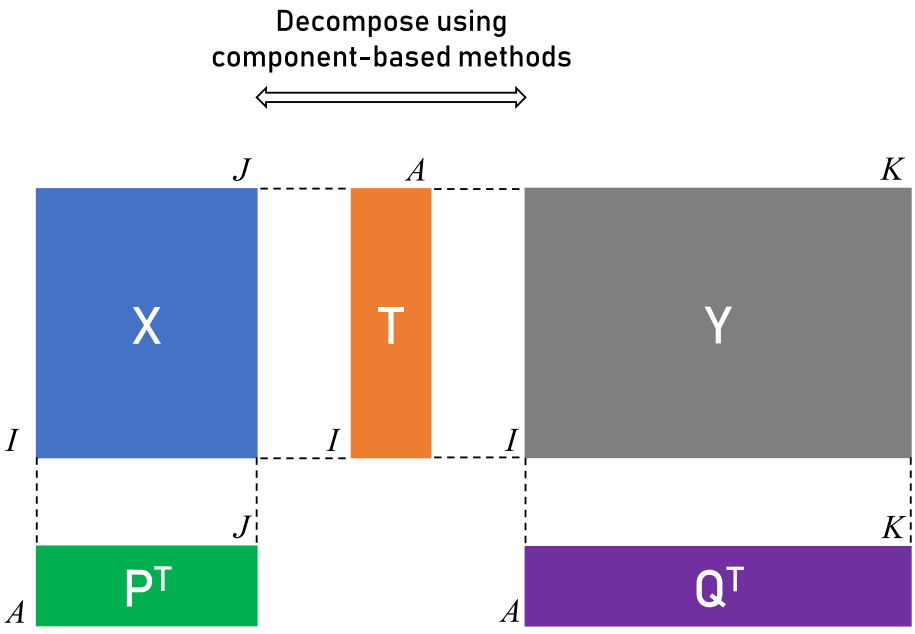


Fig. 2. The idea of relating two data sets using component methods. This is an extenuation of the idea in Fig. 1.. The additional element is that now there is a link between the scores of the two blocks, often obtained by some regression or correlation method. The T represents the scores that link the blocks. P and Q are the loadings obtained by regressing X and Y onto T, respectively.

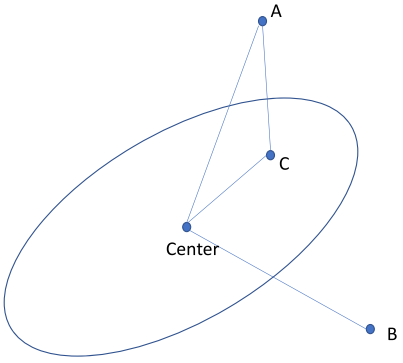


Fig. 3. Leverage and residuals. The ellipse represents the area in the subspace (defined by **T**) where most observations are located. The point A has a large residual since it does not fit to the space spanned by the majority of samples (i.e. by **T**). Point B is a so-called leverage point; it lies in the space spanned by **T**, but is positioned far from the centre within the space. The point C lies within the normal range within the plane. Points A and B will in practice be detected by using residuals and leverage measures.

does not fit to the space spanned by **T**. These types of outliers can often be detected by using/plotting the residuals E either columns-wise or row-wise, or sometimes also functions thereof as will be illustrated in the example in Fig. 5d. The point marked B is a point lying in the space spanned by **T**, but has scores far outside the normal range of variability. These types of outliers can sometimes be detected by using the score plots directly or by using various types of so-called leverage measures (see e.g. Martens and Næs (1989)). If an outlier is detected, it is not always easy to know what to do about them (leave them out or keep them), but the first thing to do is to look for possible errors. If no error is found, one should always compare solutions with and without the samples and in this way judge their importance. Apart from this, there is no overall generally valid approach and the scientist must decide based on prior knowledge and insight.

2.5. *Further extensions*

Further extensions for analysing more complicated data sets, including multiblock data (Smilde et al., 2022) and three-way data (Smilde et al., 2004), will be discussed in Section 4. The methods of this type to be discussed are MFA (Section 4.3), STATIS (Section 4.4), SO-PLS (Section 4.6), PARAFAC and Tucker-2 (Section 4.7), INDSCAL (Chapter 4.8), L-PLS (Section 4.9) and GCA (Section 4.10).

3. Typical data structures suitable for component methods

Four different and typical examples of data sets that are often analysed by component methods are given in Fig. 4. The first example (Fig. 4a) shows a three-way QDA data set (dimension $I \times J \times D$) that can be averaged over assessors to provide a standard two-way matrix (dimension $I \times J$). In the second example (Fig. 4b), this same QDA data set is unfolded so that the matrix of responses from each assessor is concatenated horizontally to create a wide supermatrix (dimension $I \times DJ$). This can also be done in cases with different number of variables in each block as for instance in free choice profiling (see e.g. Lawless and Heymann (2010). Note that this QDA data set can also be analysed by a three-way method directly, such as PARAFAC (see Smilde et al., 2004; Bro et al., 2008). Fig. 4c illustrates three different data blocks (X_{CC}, X_{CL} and X_{PI}) from a consumer liking study and how they are related. These three blocks are linked in a so-called L-structure named after the shape of the full data set. One of the blocks (X_{PI}) represents sensory descriptors or product information (PI) (usually obtained by averaging the data set over assessors as in Fig. 4a), a second data set (X_{CL}) contains consumer liking (CL) values of the same products, and a third data set (X_{CC}) represents information about the consumers or consumer characteristics (CC), for instance gender, age, attitudes and habits. For this type of structure, so-called L-shape component methods are needed (see Section 4.10). Fig. 4d illustrates schematically a section of a CATA data set. The entries here are the number of checks over the panel for each of the products. This type of structure may call for correspondence analysis. Examples of all these methods will be presented in Section 4.

Links between the various component methods and the data structure in Fig. 4 will be emphasised in each particular case,

Usually, the data from the various situations are numerical, or at

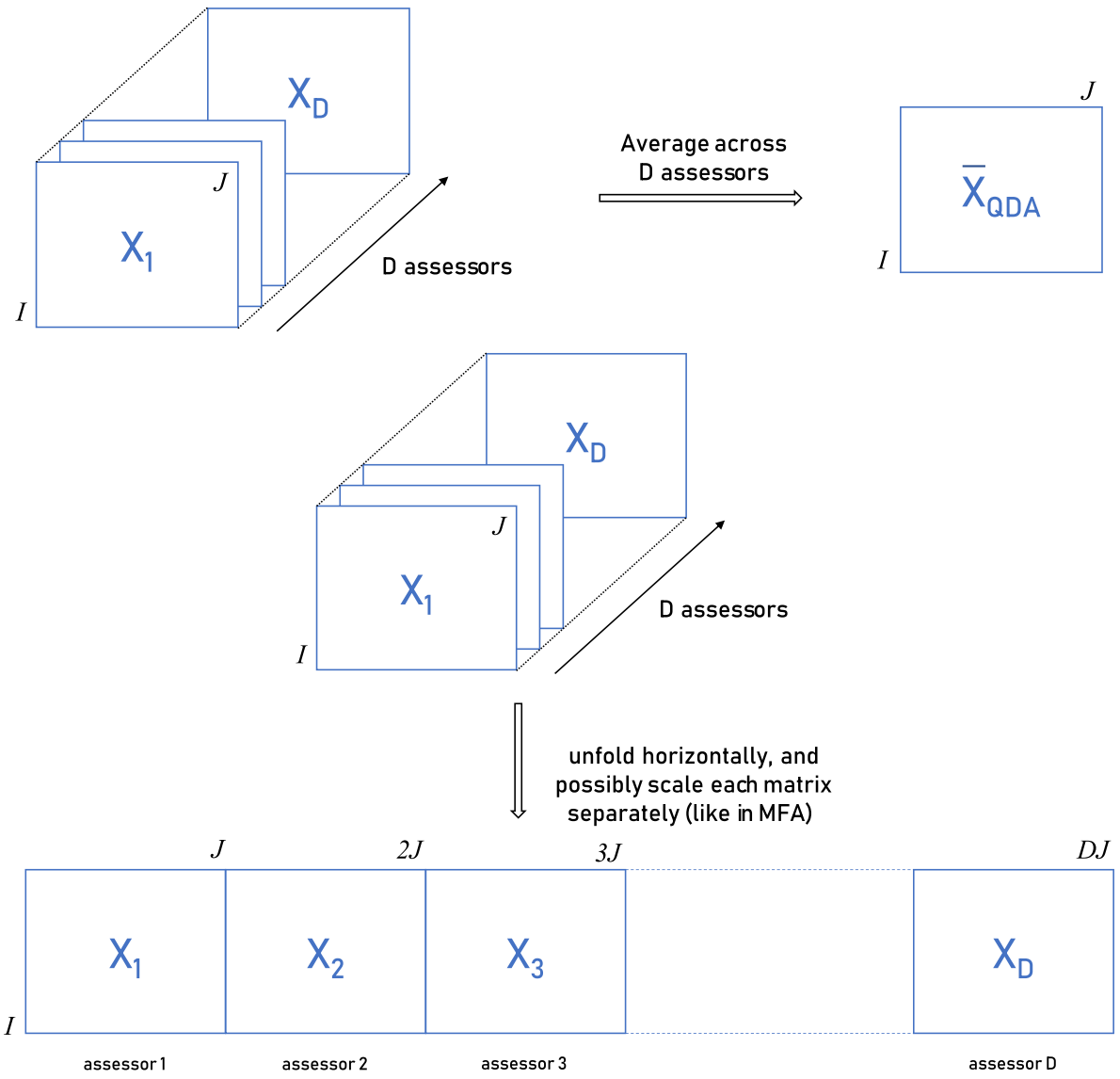


Fig. 4. a, b, c, d. Data sets from related areas. i.e. data collection methods in sensory science. 4a) represents a typical data structure for QDA, in which assessors score products on different attributes. Very often the average over assessors is used for further analysis. 4b) shows how a three-way data set, for instance from projective mapping or QDA, can be unfolded for simpler analysis. 4c) is an example of an L-shape data set with consumer liking, product attributes and consumer characteristics. 4d) shows a typical shape of a contingency table for CATA.

least treated as if they are numerical, but in some cases, the data are categorical. In this paper, we will only consider numerical data and linear methods which are the most widespread and which work very well in most cases in sensory science.

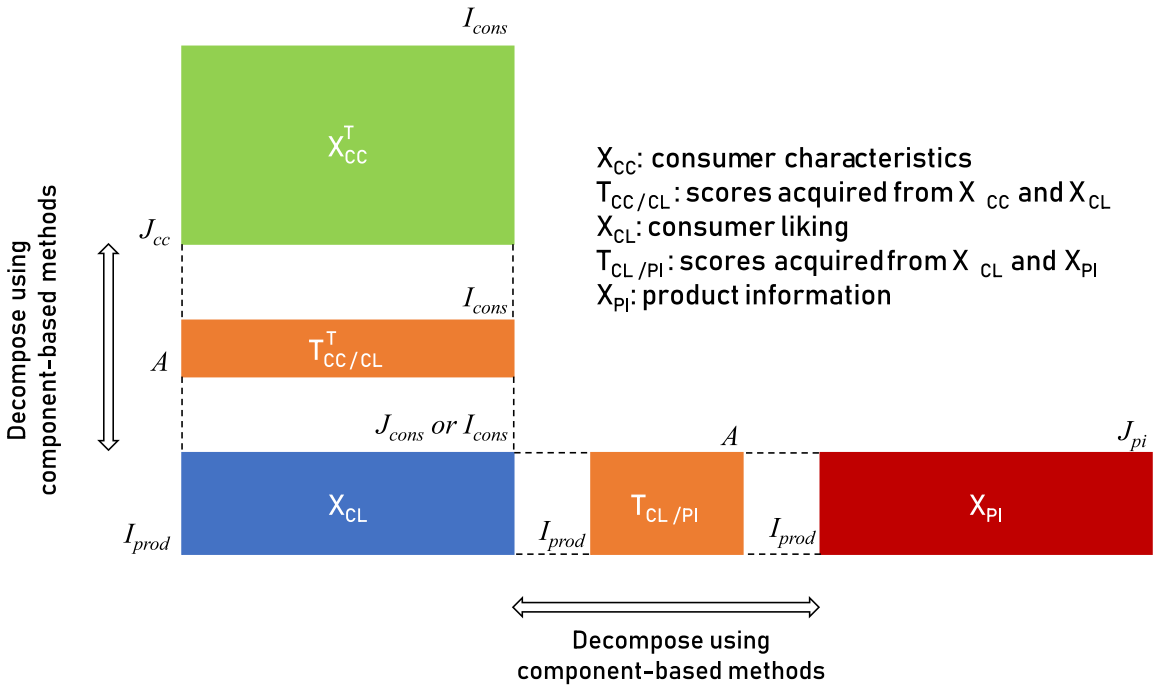
4. Selected examples of methods

In the following we will present a list of the most commonly used component methods used in sensory analysis. In each case we will mention typical sensory tests that they are useful for and also point to which data structure in Fig. 4 they represent. The link between the statistical methods and sensory tests is emphasised in Table 1.

We begin our discussion of selected multivariate analyses by providing some historical context. The earliest proposed method was Principal component analysis (PCA; Hotelling (1933), Pearson (1901)). This method extracts components **T** that describe (account for) as much of the variation in **X** as possible. When first introduced, PCA represented a new way of looking at a data set. Although it was extremely useful and versatile, it did not reach widespread use before the development of efficient computers and software in the 1970 s. The ideas behind PCA grew into a large set of alternatives for different situations and based on criteria other than explained variance.

4.1. Invariance and the need for scaling

Most of the component methods are not invariant to the scaling of the data. The reason is that variability is sought where it can be found, and variables with the largest variance will therefore have the largest influence on the solution, regardless of their importance. The sensitivity to scaling can be seen both as an advantage and a disadvantage. An advantage of sensitivity to scaling is that if there are certain variables that are considered more important than others, it is possible to highlight these by multiplying them with constants larger than one before analysis. The disadvantage is that one must be careful about which units to use for the variables. If the variables have very different variances, one will usually standardise the data in some way, typically by dividing



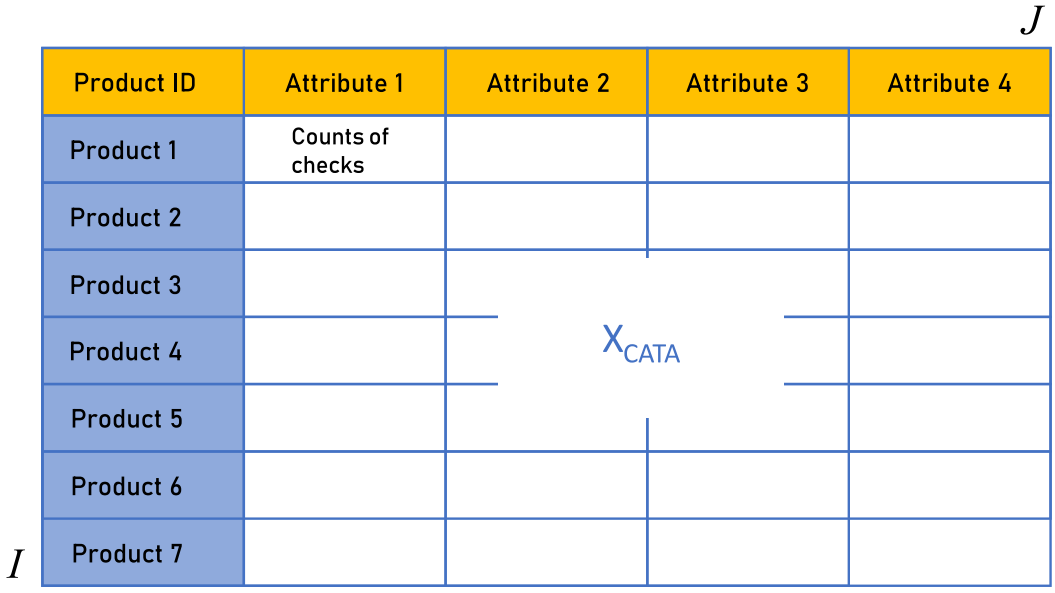


Fig. 4. (*continued*).

Table 1
The table shows the most common links between application area and component method used. Other less common possibilities exist.

		Application area							
Method	Section	QDA	PM	CATA	TCATA	TDS	Sorting	LIKING	Consumer attributes
PCA	4.2	x	x	x	x	x	x	x	x
CA	4.3			x	x		x		x
MFA	4.4	x	x	x				x	
STATIS./DISTATIS	4.5	x	x				x		
PCR/PLS	4.6	x	x	x	x	x	x	x	x
Multiblock regression	4.7	x						x	
PARAFAC, Tucker-2	4.8	x				x			
MDS	4.9		x	x			x	x	
INDSCAL	4.9		x					x	
LPLS	4.10	x	x					x	x
GCA	4.11	x	x					x	

each variable by its own standard deviation. It is assumed that all variables are centred prior to standardisation. After standardisation, all variables will have the same weight in the analysis. It should be mentioned that this may be problematic in some case ssince any variable with a small variance that represents mostly noise can then be inflated such that it becomes more important than it should be. An important possibility discussed in Næs et al (2020) for QDA is to test for significance of each attribute separately (using simple two-way ANOVA tests) and then eliminate attributes that do not show any significant relation to the design of the study.

4.2. Principal component analysis (PCA)

The PCA (Hotelling (1933), Pearson (1901)) of a matrix **X** with *I* rows and *J* columns can be represented as shown in equation 1 (see the part to the right in Fig. 4a). The **T** (obtained from $\mathbf{T} = \mathbf{XW}$, as in Section 2.1.) is the matrix of scores, **P** represents the loadings, and **E** is the residual matrix after the first few most dominating components (represented by **T**) have been calculated. The components **T** are assumed to be uncorrelated ($\mathbf{T}^{\mathbf{T}}\mathbf{T} = \mathbf{I}$) and defined in such a way that they explain as much of the variability in **X** as possible. More precisely, the first component (first column in **T**) is defined as the linear function $\mathbf{t} = \mathbf{Xw}$, ($\|\mathbf{w}\| = 1$) which has the maximum variance (i.e. $\mathbf{t}^{\mathbf{T}}\mathbf{t}$) possible. Each subsequent component is based on a similar criterion, but is assumed uncorrelated with the scores in all previous components. This process continues until the desired number of components have been extracted. The maximum number of components is limited by the lower of the number of samples or variables (rows and columns), but usually in sensory science very few components will be enough. In practice, it is difficult to interpret a large set of components using scores and loadings plots. In most cases, at most three or four components are used in the area of sensory and consumer science. The **P** obtained from equation 1 is in this case the same as **W.** The $\mathbf{TP}^{\mathbf{T}}$ can also be considered the best possible approximation to **X** in a least squares sense. The challenge is to select an adequate number of components for the purpose of summarising **X**.

Validation, for the purpose of determining the number of important components, can be done by permutation testing as suggested in Endrizzi et al. (2014). Standard simple cross-validation (CV) eliminating samples successively is more problematic here since the explained variance will always increase as the number of components increases, see e.g. Section 5.1. Methods for more elaborate removal of observations in CV are developed to circumvent some of these problems (Wold (1978)). Ways of assessing uncertainty of PCA plots can be found in for instance Castura et al. (2023) and references therein.

An important aspect to bear in mind is that one should be careful about interpreting too many components. After 2–3 components calculated for sensory science data, it is often in our experience quite difficult to find important information. This comment is valid for all methods discussed in this paper.

4.3. Correspondence analysis

A method which is very similar to PCA, but which is developed primarily for categorical table data (Fig. 4d) is correspondence analysis (see e.g. Greenacre (1993)). A typical data set from sensory analysis often analysed by correspondence analysis is check-all-that-apply (CATA see example in Section 5.4) data given as an example in Fig. 4d. Correspondence analysis can be presented in different ways, but it is essentially a method which uses a truncated singular value decomposition (with properties similar to PCA) on a contingency table of counts that are centred and standardised in a specific way, i.e.

$$\mathbf{D}_{\mathbf{r}}^{-\frac{1}{2}}\left(\mathbf{X} - \mathbf{rc}^{\mathbf{T}}\right)\mathbf{D}_{\mathbf{c}}^{-\frac{1}{2}} \tag{3}$$

where **r** and **c** are row and column sums respectively and $\mathbf{D}_{\mathrm{r}}$ and $\mathbf{D}_{\mathrm{c}}$ are diagonal matrices with diagonal elements **r** and **c**, respectively. Scores and loadings from correspondence analysis can be plotted as described by Abdi and Valentin (2007).

4.4. Multiple factor analysis (MFA)

MFA (Pagès, (2005)) is a multiblock method that applies PCA to a concatenated results matrix (Fig. 4b). In sensory evaluation, MFA can be used to analyse results from projective mapping (Risvik et al. (1994)) studies (also called napping (Pagès (2005)). In this case, the individual data set blocks, which have products in rows and the two-dimensional (x and y) coordinates in columns, are concatenated to form a supermatrix (see Fig. 4b). After column centring, each of the D blocks is weighted by the reciprocal of its first singular value. This scaling/weighting is done to make the data sets comparable. Then, PCA is conducted on the supermatrix (see Fig. 4b)

$$(\mathbf{X}_1, \mathbf{X}_2, \cdots\cdots, \mathbf{X}_{\mathrm{D}}) = \mathbf{TP}^{\mathrm{T}} + \mathbf{E} \tag{4}$$

Here the $\mathbf{X}_{\mathrm{d}}$ matrices are the individual data sets after scaling. The different **X**-blocks can have different number of variables /columns.

The so-called consensus scores **T** from (4) are considered to be a compromise for the different individual blocks. In projective mapping data, the original dimensions (i.e. dimension on the paper sheet where the samples are put) are rarely of interest, so **P** is not interpreted, but rather used to project each individual down onto the consensus space.

The individual scores for each assessor ($\mathbf{X}_{\mathrm{d}}$) can be obtained as $\mathbf{T}_{\mathrm{d}} = \mathbf{DX}_{\mathrm{d}}\mathbf{P}_{\mathrm{d}}$ where the $\mathbf{P}_{\mathrm{d}}$ are the individual parts/contributions of **P**. The individual scores ($\mathbf{T}_{\mathrm{d}}$ submatrices) can be plotted in the same plot as the original scores **T** to visualise individual differences. The average of the individual scores is equal to the consensus scores, **T**.

Often, the projective mapping task is done in combination with ultra-flash profiling (UFP; see e.g. Santos et al (2013)). After positioning samples on the table (sheet), the assessor adds words which characterize the samples. These words are summarised over the assessors in a contingency table. Typically, the relationship between two data sets, the consensus scores from (4) and the results from UFP, is analysed using projection (regression analysis), but other possibilities also exist.

The same concatenation procedure is sometimes used for panel checking (quality control, see PanelCheck software (Tomic et al. (2010); Tomic et al. (2007)) using the so-called correlation loadings plots for visualizing differences between assessors, as will be illustrated in Fig. 6. These are plots which present the correlations between the scores and the original variables. In the PanelCheck software and also in the aforementioned publications, the PCA of concatenated data (supermatrix) is called a Tucker-1 analysis named after the inventor (see Tucker (1964)).

Multi-block analyses may be used to find common and distinct components between the blocks as defined in Smilde et al. (2017). A definition of common components in the strict sense for two blocks without noise are linear combinations of the data blocks with a perfect canonical correlation; i.e., the correlation between latent variables in the two blocks equals one. Due to noise, a perfect correlation is not achievable in practice, so the definition is softened to include a relatively small number of common components with sufficiently large canonical correlations. The approach is also extended in various ways to obtain common components for more than two blocks. The distinct components can be defined as linear combinations of the two blocks that are orthogonal to the common components. The method has been used for exploring common and distinct variability in, for example, TCATA and TDS studies (see, e.g. Berget et al. (2020)).

It is important to mention that MFA (and also STATIS in Section 4.4.) can be used for blocks that do not necessarily share the same variables or number of variables. For instance, for free choice profiling (see e.g. Lawless and Heymann (2010)), where each assessor uses his/her own vocabulary, both these methods can equally well be applied.

Other perspectives on relations between datasets can be found in

Hanafi and Kiers (2006), Smilde et al. (2017) and Smilde et al. (2022),

4.5. STATIS (structuration des tableaux à trois indices de la Statistique)

STATIS is a multi-block method (typically used for QDA data, see also Schlich (1996)) based on a weighted average of the cross-product matrices $\mathbf{X_d X_d^T}$ of the individual blocks. The underlying idea is to calculate a consensus without too much influence of deviating assessors. The weights are obtained by first calculating the RV coefficients (Robert & Escofier (1976)) between each pair of blocks, then putting the weights into a symmetric matrix. This first eigenvector is calculated for this matrix (essentially a non-centred PCA). The weights for the blocks are the elements of this eigenvector divided by the sum of these elements, such that the weights sum to one. Finally, a PCA of the weighted average matrix is used to calculate the components that are used for interpretation. The individual blocks can be projected onto the same space in order to analyse individual differences. A variant of STATIS called DISTATIS (Abdi et al. (2007)) can work on distance matrices. There exists also a variant of MFA that works on the original **X**-matrices which is very similar to STATIS (Smilde et al (2022)).

4.6. Principal component regression (PCR) / partial least square regression (PLSR) for prediction

Prediction is another important application area for component methods, for instance prediction of sensory responses (e.g., QDA results) from chemistry results or consumer liking responses (hedonic responses) from sensory results (see, e.g., Martens & Martens, (2001)). Predicting sensory outcomes (**Y**) based on chemistry results (**X**) may be important for further interpretation or for replacement of simple sensory measurements in a production environment. Relating sensory QDA data (**X**) to consumer data (**Y**) by regression may be important for identifying drivers of liking. In such cases, PLS or PCR are natural methods to use. These methods are based on the model (see also Fig. 2)

$$\mathbf{X} = \mathbf{TP^T} + \mathbf{E} \tag{5}$$

$$\mathbf{Y} = \mathbf{TQ^T} + \mathbf{F}$$

where the **T** carries information from the input **X** to the regression equation, which has the output **Y**. The **E** and **F** are residuals. The components **T** are first calculated by for instance PCA or PLS and used directly as input in the regression equation with **Y** as response. The **T** is as usual obtained as $\mathbf{T} = \mathbf{XW}$ where **W** is different for the two methods. In both cases, the **P** and **Q** are estimated by least squares regression. It is well documented both theoretically and empirically that both these methods are excellent when the variables in **X** exhibit multicollinearity. The estimated version of **Q** and **P** can be combined into a matrix of regression coefficients $\mathbf{B} = \mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}\mathbf{Q}^T$ to be used in prediction (see equation 2).

Validation in this case is often done using cross-validation or prediction testing with focus on the root mean square error of prediction (RMSEP). Too many components in a model lead to an overfitted solution, while too few components do not fit the data properly. The smaller value of RMSEP, the better the predictions.

Extensions of PLS regressions to groups of data, organised horizontally in the data set, can be found in for instance Eslami et al (2014).

4.7. Multiblock regression analysis

The same methodology can be extended to multi-block input data. The model is typically written

$$\mathbf{Y} = \mathbf{X}_1\mathbf{B}_1 + \mathbf{X}_2\mathbf{B}_2 + \ldots\mathbf{X_d B_d} + \mathbf{F} \tag{6}$$

where the **X**-blocks represent different input variable blocks (see Fig. 4b) and **Y** is the output. As above, the **F** represents residuals. The simplest solution is to concatenate the input data shown in Eq. (6), then to use a method such as a standard PLS regression, but other methods are also available, such as the Sequential and Orthogonalized PLS (SO-PLS; Næs et al. (2020)) method which incorporates blocks sequentially. Another is Parallel and Orthogonalized PLS (PO-PLS; Måge et al. (2008), Måge et al. (2012)) which identifies both distinct and common variability in the **X** for prediction of **Y**. Both of these methods are developed for shedding additional light on how the input blocks effect the output block. A closely related method is multiblock redundancy analysis (see e.g. Bougeard et al (2011))

4.8. Three-way methods, PARAFAC – Tucker-2

It is possible to analyse the three-way data in Fig. 4a,b without unfolding or averaging by using a three-way method. The advantage of using a three-way method to analyse for instance QDA data is that such methods explicitly incorporate information about individual differences. In sensory science, both parallel factor analysis (PARAFAC see e.g. Bro et al. (2008) and Tucker-2 (Dahl and Næs (2009)) analyses have been applied. They are both based on the model

$$\mathbf{X}_d = \mathbf{TV}_d\mathbf{P}^T + \mathbf{E_d} \tag{7}$$

where the **T** and **P** are the common consensus scores and loadings respectively and the $\mathbf{V}_d$ matrices represent the individual differences. The **E**-matrices are the residuals matrices. The component scores (**T**) and loadings (**P**) are the same for all assessors while the individual differences are modelled in their responses to each underlying dimension. For Tucker-2, the **V** matrices are unconstrained. For PARAFAC, the **V** matrices are assumed to be diagonal, such that the individual differences amount only to different weights for the different dimensions.

4.9. Methods for distance based data, DISTATIS and INDSCAL

In some cases, sensory data are given in terms of distances between samples. For example, in a sorting task (Faye et al.(2004)), the number of times two samples appear in the same groups is used as a basis for calculating distances. Projective mapping data can also be considered to be distance data because the direction on the sheet is of minor interest.

This type of distance data can be handled by for instance multidimensional scaling (MDS). MDS transforms distances between for instance products into coordinates along component directions. The first step is to double centre the distance matrix and then a PCA is used to calculate the coordinates. A major difference is that MDS focuses only on scores, not on loadings, since for distance data there are normally no original variables for loadings to be calculated from.

In many cases, the distance data are provided at an individual level. The DISTATIS (see above) and individual differences scaling (INDSCAL; Carroll and Chang, (1970); Husson & Pagès, (2006)) are examples of component methods that operate on individual distance data. The INDSCAL is based on the assumption that the distance between two objects *i* and *j* for assessor *d* can be written

$$distance^2_{ijd} = \sum_{a=1}^{A} \left(\mathbf{t}_i^a - \mathbf{t}_j^a\right)^T v_{da}(\mathbf{t}_i^a - \mathbf{t}_j^a) \tag{8}$$

The **t**-vectors are the scores for *A* components and the weights v represent the individual differences. The model can be formulated in a way that is very similar to PARAFAC and the solution can be obtained by the PARAFAC algorithm. INDSCAL gives common component scores **T** and individual differences represented in the elements v_{da}. An applications of INDSCAL for projective mapping data can be found in Næs et al (2017).

Other important references in this area are Kruskal (1964) and Schiffman et al. (1981).

4.10. Methods for analysis of L-shape data (L-PLS and two-step procedures)

In sensory science, the most typical example of L-shape data is the one illustrated in Fig. 4c. The three blocks are consumer liking ($\mathbf{X}_{CL}$), consumer attributes/characteristics (for instance age, gender, attitudes and habits; $\mathbf{X}_{CC}$), and sensory properties for a number of products ($\mathbf{X}_{PI}$). The link between consumer liking and sensory properties represents the horizontal axis in the L where both matrices share the product dimension. The link between consumer liking and information about the consumers represents the vertical dimension where both matrices share the consumers. Important goals here are understanding how the sensory attributes influence the liking and how this liking pattern relates to information about the consumers. As a possible way to analyse these data, one can first use PLS to model the relationship between sensory results and consumer liking, then use another PLS to model the relationship between the consumer loadings from this first analysis with the consumer information. Another possibility is to use a simultaneous approach called L-PLS (Martens et al. (2005)) which is based on calculated components from the singular value decomposition of the product of the three matrices. The L-PLS is based on extracting components from

$$\mathbf{svd}(\mathbf{X}_{\mathbf{CC}}^{\mathbf{T}}\mathbf{X}_{\mathbf{CL}}\mathbf{X}_{\mathbf{PI}}^{\mathbf{T}}) \tag{9}$$

with a deflation of the blocks for each component. The SVD here stands for the singular value decomposition (which is one way of doing PCA) of the matrix in parenthesis.

4.11. Generalised canonical analysis (GCA)

The GCA (Van der Burg and Dijksterhuis (1996)) is a multiblock method for analysing several data sets ($\mathbf{X}_1$, …, $\mathbf{X}_D$). It has a strong structural similarity with, for instance, MFA. The difference is that while the methods above try to explain the different blocks as well as possible, the GCA optimizes the correlation between the components extracted from the different blocks to obtain the consensus scores $\mathbf{T}$. These scores can be obtained in different ways, but the most common definition is to minimise the following expression:

$$\text{argmin}\sum_{d=1}^{D}\|\mathbf{X_d}\mathbf{U}_d-\mathbf{T}\|^2 \tag{10}$$

under the constraint that $\mathbf{T}^{\mathrm{T}}\mathbf{T} = \mathbf{I}$. The $\mathbf{U}$ matrices show how the $\mathbf{X}$ matrices relate to the consensus scores in $\mathbf{T}$. Methods for regularizing the GCA solution can be found in Tenenhaus and Tenenhaus (2014).

4.12. Validation

Validation is a very important aspect of component methods as also discussed briefly for some of the methods above. This can be done in many different ways, dependent on whether one is interested in internal (for the actual data set) or external validations. It is, however, not obvious how to do it for small data sets. The most commonly used methods are cross-validation and permutation testing. A full discussion of this will not be given here, instead we refer to Næs et al. (2020). We will touch on the topic for a number of the examples in Section 5.

It is important to mention that a careful look at possible outliers and abnormalities in the data sets, using for instance the methodology discussed in connection with Fig. 3 (Section 2.3), should be an integral part of a full validation of a model.

4.13. Software

Calculations for all the above methods can be done using for instance various open-source software packages such as Scikit-learn, Hoggorm, Prince, FactoMineR, SensoMineR or commercial packages such as The Unscrambler or PLS Toolbox. Most of the methods are also available in standard statistical software packages such as for instance SAS and Minitab.

5. Illustrations

Table 1 presents some of the most common links between methods described and application areas. Some of the them will be presented in some detail in the examples to come.

5.1. Illustration of PCA and PARAFAC for QDA data

In this example, a data set from QDA of olive oil is used for illustration. There are 11 commercially available olive oils measured on 20 sensory variables. For more detailed analysis we refer to Næs et al. (2020). The data are available from the first author on request.

5.1.1. PCA for data averaged over assessors

The scores plot and loadings plot for the raw average sensory (averaged as in Fig. 4a) data are given in Fig. 5a and 5b together with the explained variances in Fig. 5c. Leave-one-out CV and permutation testing (see Næs et al (2020)) indicate that a three-component solution is possible, although component 3 describes very little of the variation. We will here for simplicity concentrate on the first two components describing more than 95 % of the variance. In this case we see that the leave-one-out CV gives an indication of the number of components which is similar to the permutation testing results. We refer to Section 4.1 for a warning regarding leave-one-out CV for PCA.

The number of dimensions in the data set, represented by principal components, is much smaller than the number of variables (Fig. 5c), which facilitates interpretation. In other words, the original 20 variables are here reduced to two components (latent variables) which can be plotted in scatter plots (Fig. 5a and 5b) and interpreted visually.

Additionally, we plotted the squared residuals for each product for different number of components (Fig. 5d). This plot shows that the residuals after the first component (blue), the squared residuals are largest for oils 6 and 10. This means that the two samples are not fitted well by the first component. These are the samples responsible for the main variability along component two (Fig. 5a). After two components have been fitted (dotted orange), the residuals are relatively small for all objects. This supports the finding above that only two components are needed for describing most of the variability.

The first component is dominated by a contrast of fruity, flower and sweet with the burning, grass and unripe flavours. In other words, the main dimension is characterized by a contrast between the fruity/sweet and more grass-oriented flavours. The second component contrasts trigeminal sensations (burning, astringent), typical descriptors of fresh-pressed olive oil, as opposed to grass, nutty and chemical flavours that may be linked to older, more rancid, or lower-quality olive oils. The samples in the score plot are characterised by higher values of the variables in the same area in the loadings plot. For instance, sample ten has characteristics of a high quality, freshly pressed olive oil, characterized by fruity flavours and astringent and burning notes, while sample six is nuttier. From the joint interpretation of the scores plot and the loadings plot, it is clear which of the samples are characterized by the different sensory properties.

This example illustrates that a small number of components describe almost all the variability of the original attributes. It also demonstrates how these dimensions can be interpreted and validated. No abnormalities were detected using the residuals and scores plots.

5.1.2. PCA for individual assessor data

Next, we apply a Tucker-1 analysis of the horizontally concatenated data set (see Fig. 4b) in order to shed some light on individual differences. Here we show only two of the many possible plots that can be used with a Tucker-1 analysis (Tomic et al., 2007). First, we focus on the

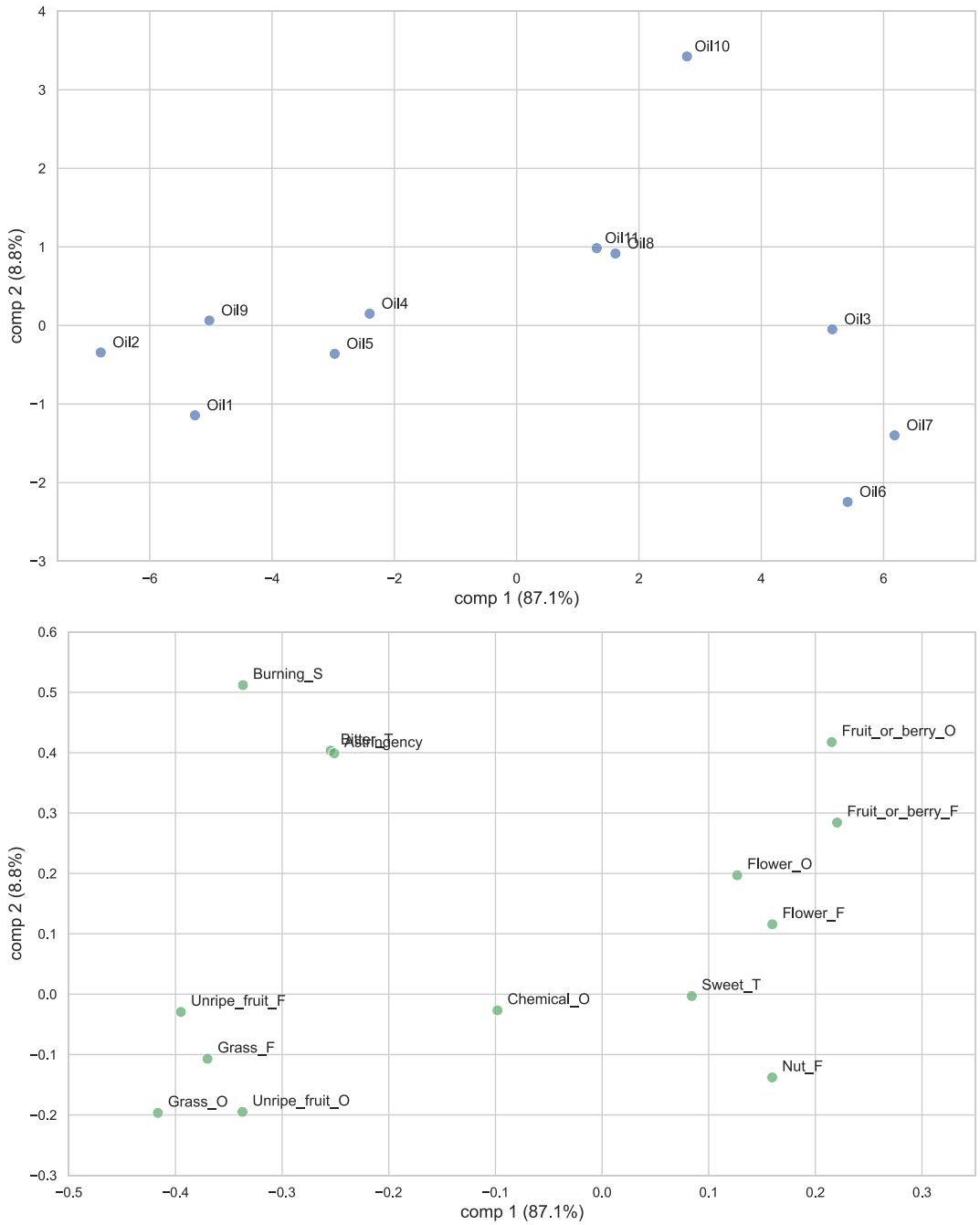


Fig. 5. A, b, c, d. Example of the use of PCA for averaged (over assessors) QDA data (olive oil, only significant attributes), showing 5a) scores (oil 8 and oil11 overlap in the plot) 5b) loadings (Bitter_taste and Astringency overlap in the plot), 5c) explained variance (based on data fitting and cross-validation), and 5d). squared residuals for different number of components for the different samples.

correlation loadings plot (see e.g. Fig. 6a and Martens and Martens (2001)) for all combinations of assessors and attributes. All points are marked (i.e. each combination of attribute and assessor). Since attributes are usually investigated one at a time, the different individuals are highlighted for one variable (fruit or berry odour) to allow for comparison.

The correlation loadings are defined as the correlations between the scores and the individual variables. The main advantage is that one gets information on the explained variance for each individual variable based on the two components the plot is based on. The circles in the correlation loadings plot represent 50 % and 100 % explained variance of the different variables. Other than that, the overall interpretation is in many cases similar to using the standard loadings plot.

One of the individuals (assessor 105) is deviating strongly from the rest along the first component. From the residual plot (Fig. 6b) it can be seen that in this case the same attribute and also its flavour counterpart ('fruit or berry flavour') are among the two attributes least described by the first two components. This indicates that these attributes are among those which fit the least to the Tucker-1 model. The correlation loadings plots show a need for retraining this particular assessor; either this person is not sensitive to this particular attribute, or he/she has misunderstood it. When data are averaged over assessors. the influence of this single assessor may be quite small, but there may still be good reasons for assessor retraining. In other cases there may be more than

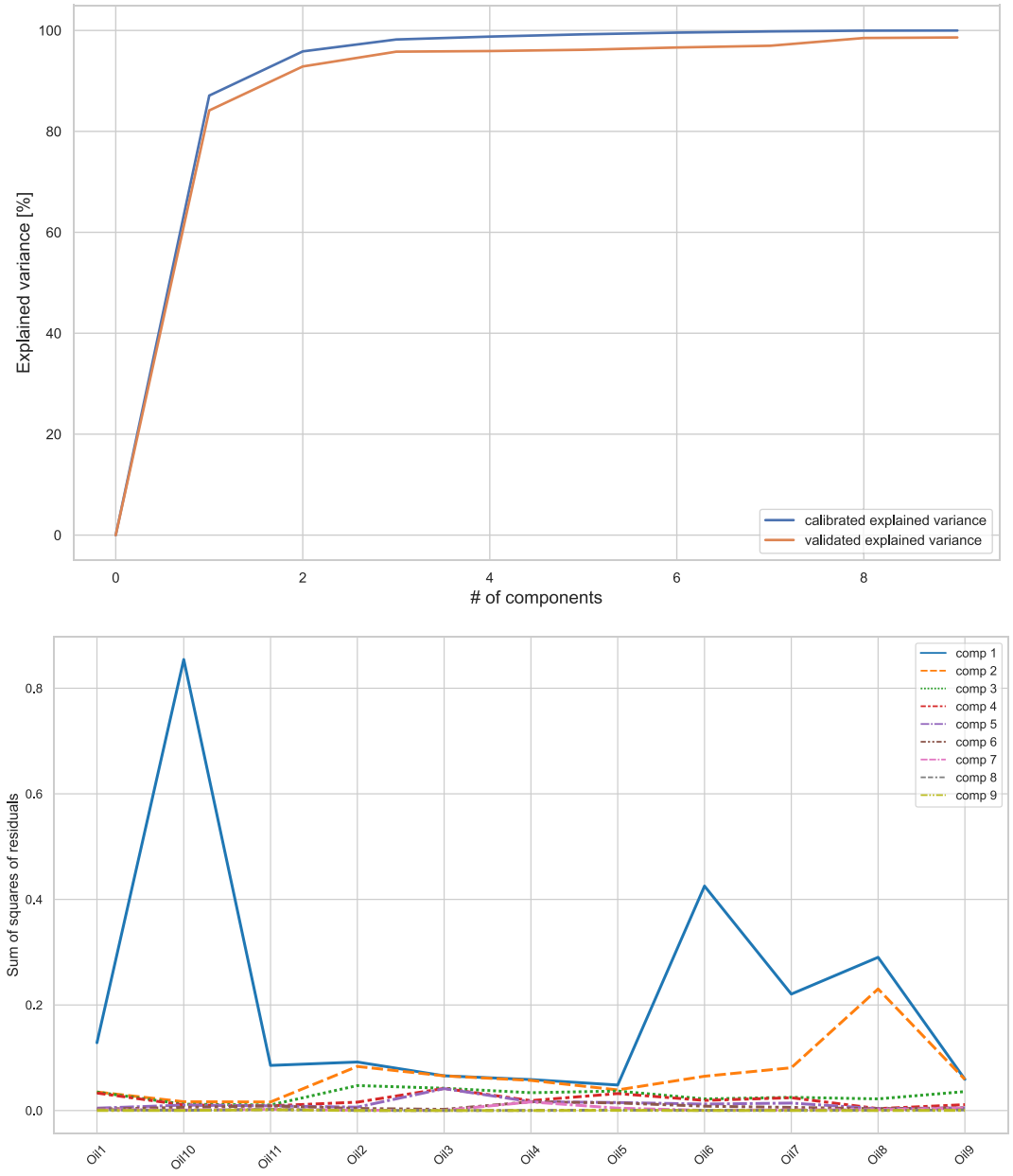


Fig. 5. (*continued*).

one deviating assessors, making it even more important to check assessor performance and take proper action.

This shows that Tucker-1 component plots can be useful for detecting outlying observations, in this case non-conforming assessors in sensory analysis.

5.1.3. PARAFAC for individual assessor data

The data (to the left in Fig. 4a) were centred in the same way as above, i.e. the average was subtracted for each combination of assessor and attribute. The two first components explained about 27 % of the variation. This relatively low number is quite natural given the strong model assumptions underlying the method and the fact that the percentage refers to variability of all assessors. On the other hand, the solution represents an optimal result when scores and loadings are assumed equal for all assessors but allowing for a different weight for each assessor.

When interpreting a PARAFAC solution, it is important to be aware of an indeterminacy in the signs of the components, i.e. multiplying two of the 'ways', for instance assessor and attribute components, with −1 gives the same fit. In this paper we have chosen a solution that resembles as much as possible the PCA scores and loadings. As can be seen from Fig. 7, the oil mode component (Fig. 7b) and attribute mode (Fig. 7c) show some similarities with the PCA scores and loadings, but also some differences. The assessor mode plot (Fig. 7a) gives information about which assessors that have the largest contributions to the different PARAFAC axes.. For instance, assessors 3 and 4 (Fig. 7a) seem to give a high score for the fruit-oriented attributes in particular for samples 10 and 11. Having access to this type of information and being able to allow for individual difference is one of the advantages of PARAFAC.

5.2. Illustration of MFA based for projective mapping data

The olive oil samples for this projective mapping example were the same as for the QDA example. MFA was run on the concatenated data set (see Fig. 4b). A panel of ten assessors participated in the study, the same as were used for QDA. They also conducted an ultra-flash profiling (UFP)

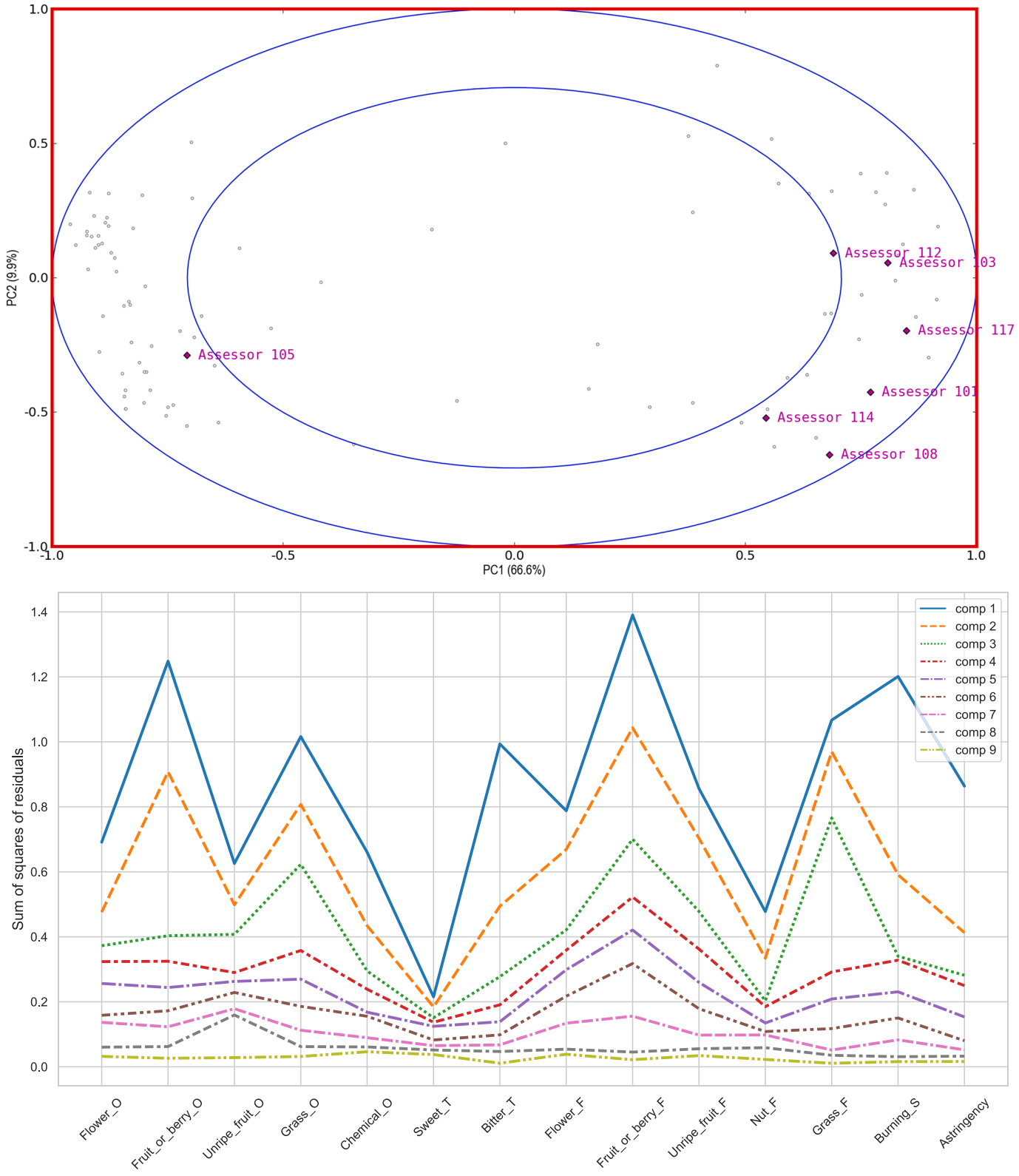


Fig. 6. a,b. Analysis of three-way QDA data structure (olive oil, samples assessors and attributes) by Tucker-1 using only significant variables – 6a) Correlation loadings plot for one of the attributes (Fruit_or_berry odour) and 6b) sum (over assessors and samples) of squared residuals for different components and attributes.

in order to gain better insight into what are the different product characteristics. The UFP data were accumulated in a table and regressed onto the consensus score. This regression gives loadings in the same style as standard PCA loadings. The individual scores were visualised (as described in Section 4.4) to investigate individual disagreement.

The consensus scores (Fig. 8a) resemble to a large extent the scores plot for the QDA data, as can be seen by tilting the QDA plot slightly counter clockwise. The explained variances (inertia) are lower in MFA

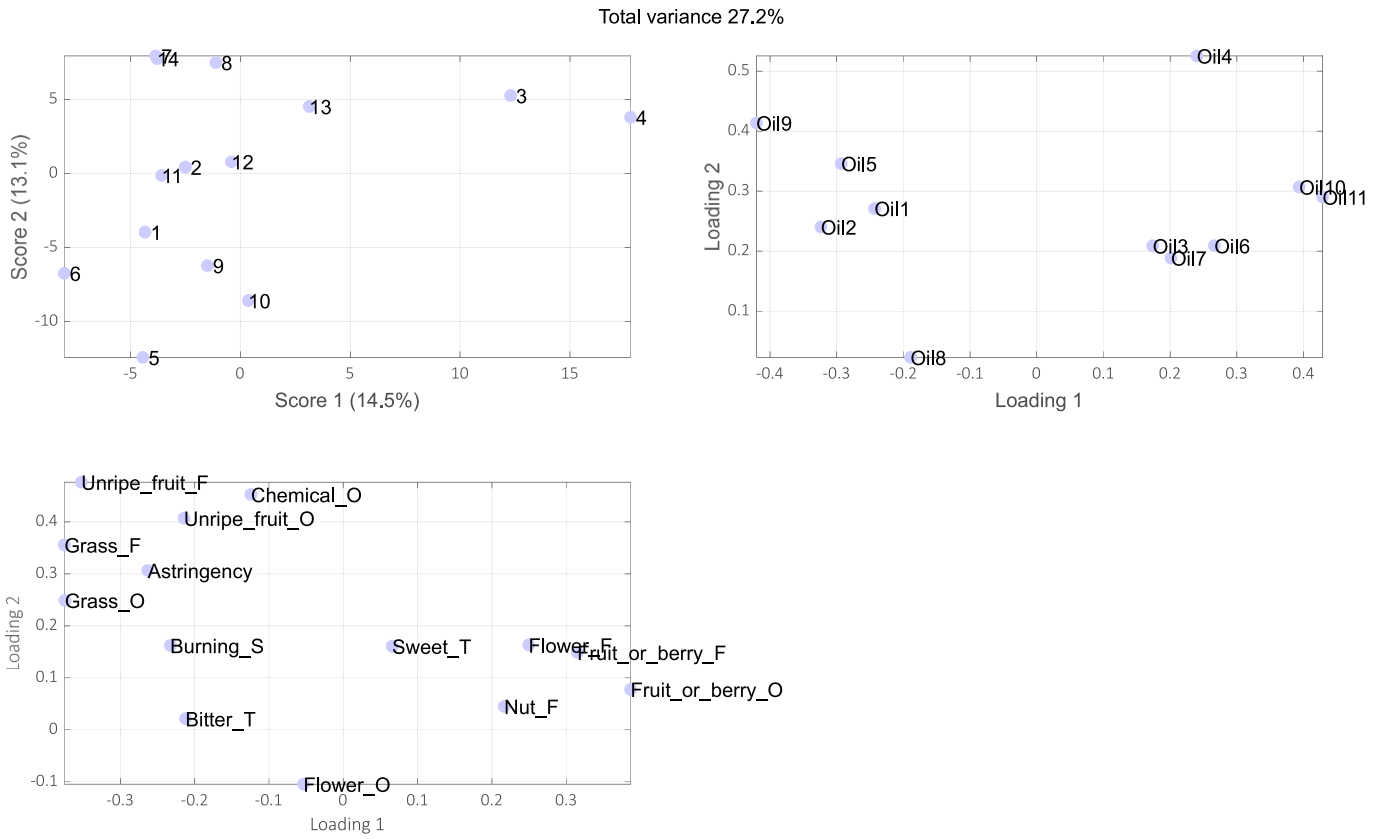


Fig. 7. PARAFAC results (2 components) for the olive oil data. The estimated values are presented for 7a) sensory assessors, 7b) samples, and 7c) attributes.

based on projective mapping data from untrained assessors than in the PCA of the QDA data from trained assessors.

The samples are characterized by properties with similar, but not identical, meaning in both UFP (Fig. 8b) and in QDA.

The individual differences in the projective mapping (PM) study are quite large (Fig. 8c), as expected for this type of analysis. It is clear, however, that assessor agreement for the first components is strong enough to clearly distinguish two main groups of oil (as for QDA above). For the second axis the disagreement is relatively larger, so the difference between these groups of samples is less clear compared to the first axis.

This result shows that component methods are also useful for analysing individual differences in sensory data. The two examples (QDA and PM) support the fact that there is some underlying dimensionality which is conveyed from the assessors who used two very different sensory methodologies.

5.3. *Illustration of regression by PLS*

Here we will present an illustration of the use of PLS regression for relating consumer liking data for a number of samples to sensory characteristics (QDA) of the same samples. The approach taken here is simple so-called internal preference mapping (McEwan (1996)). This is a method where the sensory profile acts as the output and the consumer liking is used as input to the regression. The goal is to identify which samples that are most liked and also to characterize these samples by the use of QDA. Note that the concept of internal mapping is sometimes used for analysis of consumer liking data with PCA alone.

The data are the same as those used in Agudelo et al. (2015) and Næs et al. (2018). One hundred consumers were asked to give their overall liking of six samples of fruit fillings, formulated with three different hydrocolloid systems: tapioca starch (TS), modified waxy corn starch (C), and a mixed system with tapioca starch plus pectin (P). Each of these was prepared with sugar (S) or with polydextrose and artificial sweeteners (PD). In other words, the samples are based on an experimental design with two factors with 3 (TS, C and P) and 2 (S and PD) levels. (i.e. a 3×2 design) In addition to the consumer liking the same samples were characterized by QDA.

In Næs et al. (2018), the boxplot of the liking of the six samples shows clearly that the samples with sugar (S) are more liked. There is, however, limited information available in this plot about the individual differences among consumers and also about which sensory properties are involved in the liking or lack of liking.

The first two components of PLS described about 50 % of consumer liking data and almost 90 % of the QDA data. This indicates that there is good correspondence between liking and sensory properties. With so few samples, proper multivariate validation, for instance cross-validation, is problematic. Therefore, we evaluate these results by considering related information and context. Do the results make sense based on what is known about the samples? And how clear are the results in terms of interpretation?

The PLS plots (Fig. 9a, b, c) show the same overall tendencies as the boxplots: The samples formulated with sugar are generally more liked than the other three samples. The samples TS-S and P-S are particularly well liked. The consumer/assessor plot (Fig. 9b) shows that consumers are scattered, indicating some disagreement about which products they like. It can be seen from the QDA loadings (Fig. 9c) that the samples with sugar are generally fruitier and have less strange taste. The best-liked samples (upper left in Fig. 9b) are generally fruitier than the rest. All these results are in good correspondence with what could be expected; the results are very explainable and clear. As can also be seen, the design variable S/PD is clearly a separating axis in the plots which supports the validity of the analysis.

This example emphasises that the component method PLS can be used to investigate the variability among consumers in a data matrix based on a large number of consumers and its relation to sensory

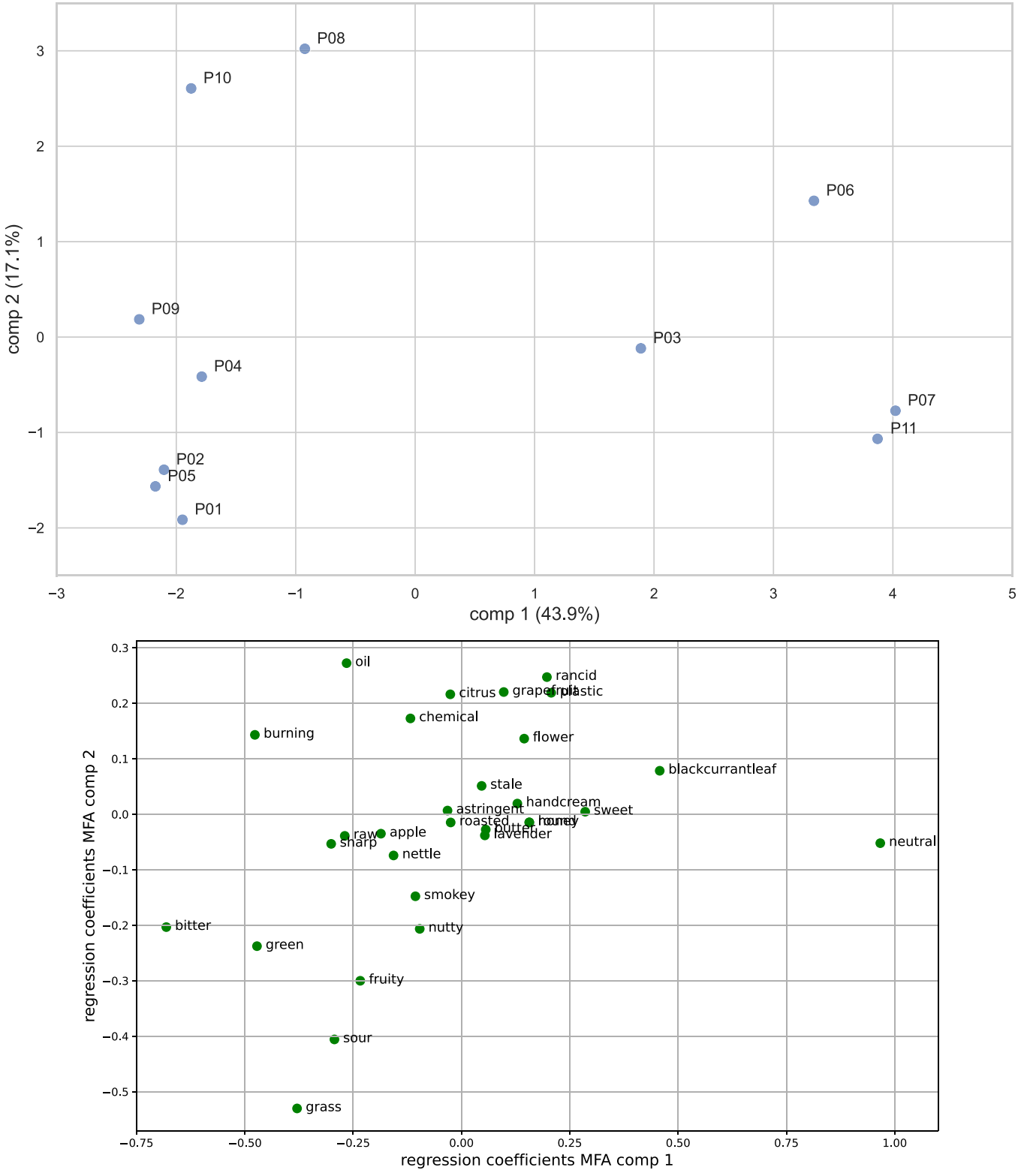


Fig. 8. a, b, c. MFA for projective mapping data (olive oil samples) showing 8a) the consensus profile, 8b) results from ultra flash profiling (Honey and Round overlap in the plot), and 8c) individuals projected onto the consensus plot.

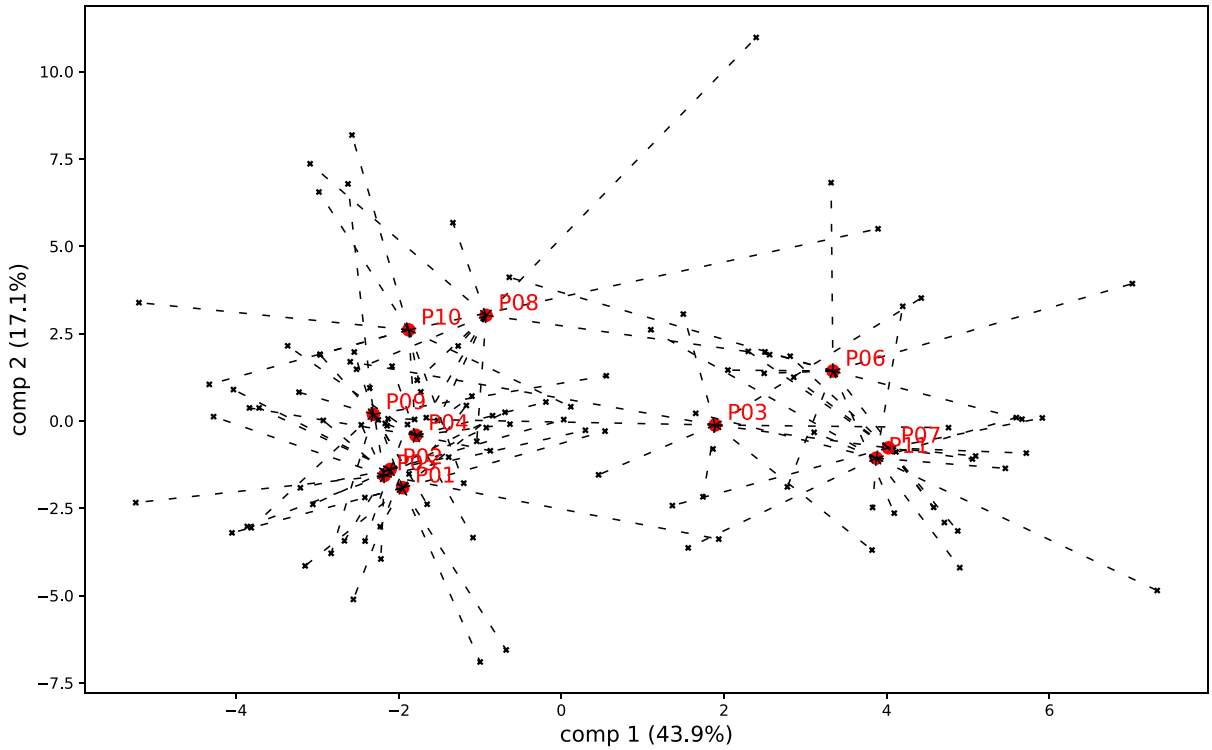


Fig. 8. (*continued*).

characteristics of the samples. The input data matrix has so many (collinear) columns and few rows that an ordinary multiple least squares (LS) regression method cannot be used. Important information can be gained by reducing many variables to only a few underlying components. One can also see that the underlying strategy and tools for interpreting PLS results are the same as used for PCA.

5.4. *Illustration of the use of PCA for TCATA data*

The temporal check-all-that-apply (TCATA; Castura et al., (2016a)) method allows assessors to characterize how each sample is perceived during the evaluation. In one variant of this method, TCATA Fading (Ares et al., (2016)), a list of words or phrases (called attributes) are presented to assessors on a screen. When the evaluation begins, assessors tick (check) attributes that characterize the sample. When an attribute is ticked, it is emphasized visually, but this visual emphasis fades until the appearance of the attribute returns to its unticked state. Assessors can re-tick a fading attribute at any time. Sometimes an attribute returns to its unticked state before it can be re-ticked; this produces small endorsement "gaps" that can be filled prior to data analysis (Vidal et al.(2017)). The TCATA data for the sample are coded 1 at timepoints when an attribute is ticked, including while it is fading, and coded 0 when it is not ticked.

Recently, the TCATA Fading method was used by assessors to characterize sensations elicited by de-alcoholized beer over a 90-seconds duration using seven attributes: *astringent*, *bitter*, *carbonation*, *fruity/hops*, *malty*, *sour*, *sweet* (Mitchell, Castura, Thibodeau & Pickering, (2019). Tasting conditions from a 2 × 2 factorial design were evaluated. The first factor was beverage temperature; the two levels were refrigerated, 6 °C, and room temperature, 21 °C. The second factor was sound augmentation, in which a 7-seconds audio clip of carbonation sounds was either played (at 40 dB) or not played at the start of the evaluation on headphones worn by subjects. The audio clip (available in Mitchell et al.,(2019, Supplemental Materials) contained fizzing and popping sounds characteristic of carbon dioxide effervescence.

Test results indicate that about one-third of people perceive one or more "phantom tastes" (e.g. bitter, sour, sweet, metallic) when the tongue changes temperature; an individual who perceives hot or cold on the tongue in this way is called a thermal taster (TT; Cruz & Green, (2000); Bajec and Pickering (2010); Yang, Hollowood & Hort (2014), Thibodeau et al. (2018)). A person who does not experience a taste when their tongue changes temperature is called a thermal non-taster (TnT). In the two-independent groups study by Mitchell et al. (2019), 21 TTs and 20 TnTs underwent training and practice with the evaluation protocols, then evaluated beer samples in duplicate for each combination of conditions. Results showed that TTs experienced and ticked more attributes than TnTs in all of the conditions tested. Readers are referred to Mitchell et al. (2019) for further details of the study.

Here, we show a multivariate analysis that complements the analyses presented by Mitchell et al. (2019). We organized the results then applied PCA in the same manner as in Castura, Baker and Ross (2016b). Specifically, TCATA results at 0.1-second intervals were summarized in a (7208 × 7) matrix of mean citation rates (i.e., average tick rates, over the assessors), where rows corresponded to all combinations of the 2 TT statuses, 2 beverage temperatures, 2 sound conditions, and 901 time points (0 s to 90 s by 0.1-s increments), and where columns corresponded to the 7 mean-centred sensory attributes. PCA results were plotted in R 4.1.3 (R Core Team 2022) using the R package tempR (Castura (2022).

The full results matrix for the 90-seconds evaluation was submitted to PCA. The perceptions evolve rapidly in the first 30 s, then decay over the next 60 s. For brevity and clean presentation, only two of the eight trajectories are shown in Fig. 10, and only for the first 25 s, just before all sensory attributes attenuate and the attribute citations rates return gradually to zero. The trajectories show how perceptions evolve for the TT and TnT groups until the 25-seconds timepoint for the refrigerated beverage evaluated by subjects who heard the audio clip of carbon dioxide effervescence on headphones. So, as expected, this sample is characterized initially by the carbonation (trigeminal sensation), which was often the first attribute that subjects ticked. Before 25 s, various other sensory attributes become noticeable. At the end of the evaluation, almost no attributes are ticked. So if the full 90-seconds trajectories were

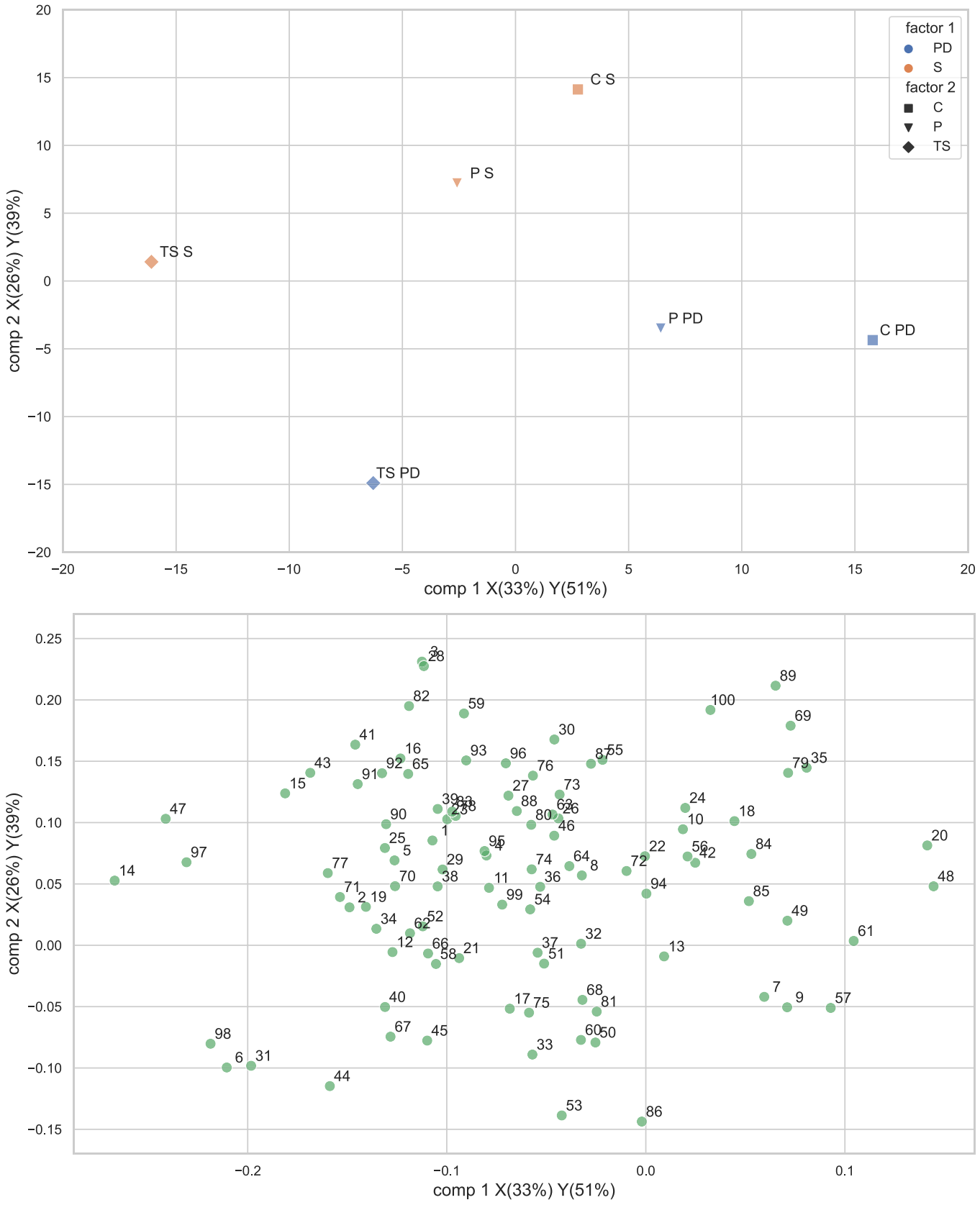


Fig. 9. a, b, c. Preference mapping. PLS used for the fruit filling example, product attributes vs. consumer liking — 9a) score plot, 9b) loadings for the consumer data, and 9c) loadings for the sensory data (Mouthcoating and Resistance time overlap in the plot).

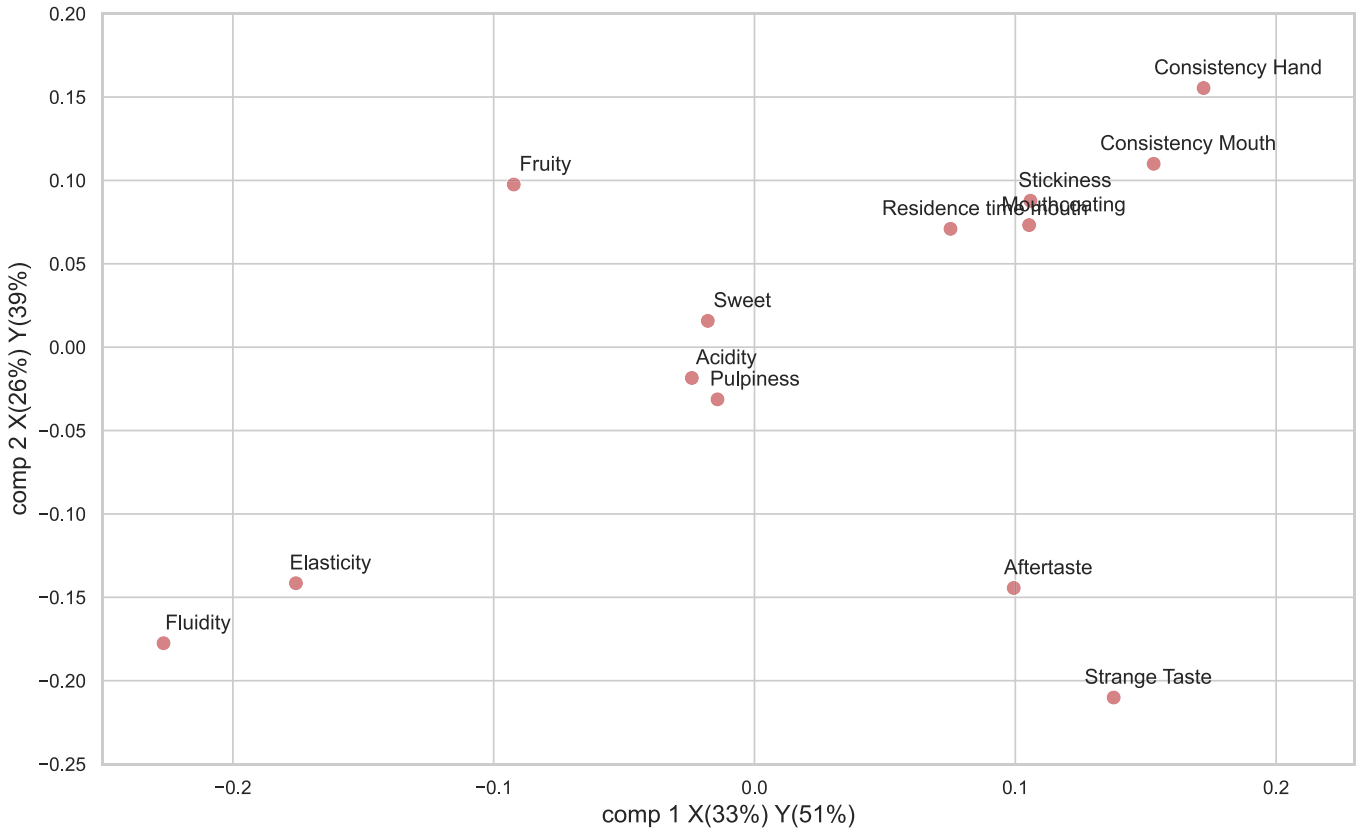


Fig. 9. (*continued*).

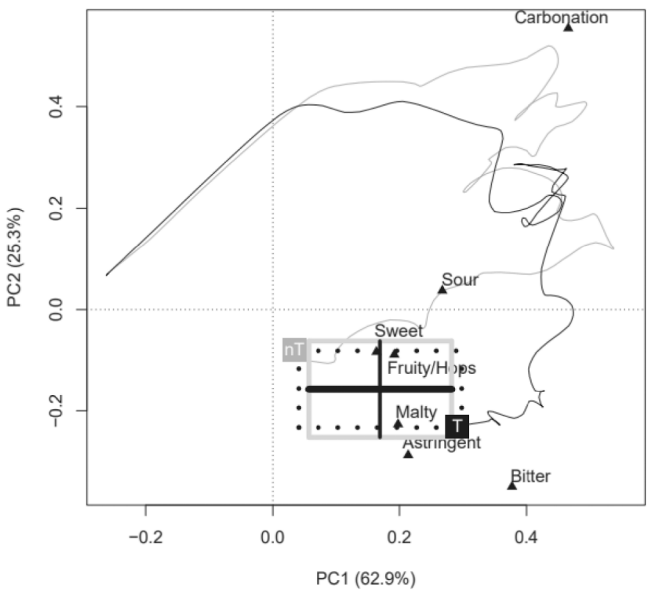


Fig. 10. Trajectories for thermal tasters (in the plot denoted only T) and non-thermal tasters (in the plot denoted only nT) for the first 25 s (of the 90-s evaluation) for the non-alcoholic beer evaluated under the refrigerated temperature condition with sound augmentation present (i.e. headphones played the audio clip of carbonation sounds). Trajectories are labelled at their positions at 25 s (with T and nT). The rectangle superimposed is a helping tool to show significance or not. The dotted rectangle shows the actual differences along the first and second axis. The grey solid rectangle represents the lines for comparison to assess significance. The difference between the TT (denoted by T in the plot) and TnT (denoted by nT in the plot) groups at 25 s is significant in PC1 (heavy horizontal line) but not in PC2 (thin vertical line).

plotted, then the trajectories would return approximately to the point from which the trajectories originated.

Some visual aids (grey box) are added to evaluate whether the trajectories differ in each principal component at 25 s. The approach comes from a bootstrap-derived procedure (see Castura, Rutledge, Ross & Næs (2022). The dotted line shows the paired difference between TT and TnT at 25 s. This paired difference is considered to be significantly different if it is larger than the half-width of the 95 % confidence interval for this paired comparison, which is represented by the grey box. The paired difference is larger than that half-width of the 95 % confidence interval in principal component 1 but not in principal component 2, indicating that the TT and TnT groups differ mainly in their overall citation rates (principal component 1) rather than in a shift away from carbonation towards tastes and mouthfeels that tend to be perceived later (principal component 2). As mentioned previously, thermal tasters experience phantom tastes when the tongue experiences temperature changes. As expected, the TT subjects experience more sensations than TnT subjects. Conclusions that we obtain from this analysis concur with conclusions presented by Mitchell et al. (2019).

6. Conclusions

This paper discusses the benefits of using component methods in sensory and consumer sciences. A large number of component methods used in sensory analysis are presented and discussed briefly. How these methods are linked to various types of sensory and consumer studies is discussed and illustrated by examples. It is also emphasized that the component methods are versatile and represent the same way of thinking regardless of which of the methods is used. The most important common feature is the use and interpretation of scatter plots of scores and loadings in a low dimensional space. Calculations are generally simple and can be done by standard software packages.

The main advantages of component methods

The following points summarize some of the most important arguments in favour of component methods.

- The component idea fits well with the fact that, in practice, the number of sources of variability in a data set are often much smaller than the number of measured variables (see e.g. examples in Section 5.1 and Section 5.3).
- The concept of underlying dimensions (or components or latent variable) is a useful concept, even as an approximation, for understanding complex data sets (see e.g. Section 5.1).
- Component methods are suitable for simple interpretation and communication of complex intercorrelations among samples as well as variables. (see e.g. Section 5.1)
- Subsequent data analysis is often more stable and reliable if the components are chosen according to a sensible criterion. For instance, standard multiple least squares regression analysis with many collinear input variables can be made more stable and reliable if replaced by a component-based regression method such as principal component regression (PCR, see e.g. Martens and Næs (1989)) (see e.g. Section 5.3).
- The models can easily be validated empirically by permutation testing (Endrizzi et al (2014)), cross-validation (CV, Stone (1974)), or prediction testing on new data. This is useful for avoiding overfitting and determining how many components one can rely on before the noise becomes dominating (see e.g. Section 5.1). This is, however, not always trivial.
- Component methods are useful for reducing noise, i.e. calculating components can often be seen as a kind of filtering, i.e. separating the signal from the noise (see equation 1 and Section 5.1).
- The component methods lend themselves to detection of outliers (see e.g. Martens & Næs (1989). Outlier detection can be enabled by tools that measure distances between objects and the centre (average) of the data (see e.g. Fig. 3 for a conceptual illustration of different of outliers and also the figure caption for an indication of how they can be detected). See also Section 5.1.
- Component methods are versatile and can be used in many types of multivariate data analysis for exploring a data set, generating hypotheses and confirming what is already known. They are not always optimal, but in most cases they are useful. See the ensemble of examples in Section 5.
- All component methods allow to interpret data similarly, from an applied point of view. This consistency may be an advantage when moving between different application areas and methods. See the ensemble of exanples in Section 5.

In addition to these concrete aspects covered in this paper, we would like to mention a few other and also more general aspects that hold for component method.

- Component methods can be used for selecting samples for more elaborate studies. One can for instance use PCA scores of sensory panel data to select a few relevant samples for more elaborate and time-consuming consumer testing (see e.g. Helgesen and Næs (1995) and Moskowitz et al. (2017)).
- Many component methods can handle missing information quite efficiently without needing to remove samples or variables (examples are PCA and PLS regression), hence making full use of the available information. Otherwise, the missing values have to be imputed by a suitable imputation method.
- The component models allow for understanding even complex reasons for observations being outliers, for example through the use of contribution plots (see e.g. Miller et al. (1998)).
- For instance PCA is very useful as a first step in a data analysis process since it can often be used to detect mistakes and problems with the data.
- Component methods are excellent for efficient feature extraction, for example before fitting complex nonlinear models.
- Most component methods are implemented and available in standard software packages.

It should be mentioned, though, that there are aspects of these methods that may be considered disadvantages by some scientific communities and researchers.

- Incorporating prior information directly in the data analysis is sometimes difficult. Although possible to incorporate prior information in cases such as multivariate curve resolution (Tauler, 2001), in most other cases it is impractical to do so. When available, prior information can be used informally when reviewing/interpreting scores and loadings plots.
- Non-linear and other more complex relationships are not an integral part of the standard component methods. For some methods, non-linear analogues exist.
- Statistical probability models are usually not involved in component extraction, making classical statistical inference more difficult. Bootstrap methods (see e.g. Castura et al (2023)) can, however, be very useful for making inference in component models.
- The methods are not always suitable for situations with limited correlation among variables, since in such cases a larger number of components is needed.

Author statement

Tormod Næs had the idea and also has been the driving force in the whole structuring and writing process.

Oliver Tomic, Rasmus Bro and John Castura have done the calculations and contributed in writing and editing of the manuscript.

Paula Varela has contributed with writing and interpretation of results.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Tormod Naes reports financial support was provided by Norwegian Institute of Food Fisheries and Aquaculture Research.

Data availability

Data will be made available on request.

Acknowledgements

The authors would also like to thank for the financial support received from Research council of Norway and the Norwegian Foundation for Research Levy on Agricultural Products FFL, through the research program “Food for future’ and ‘Precision’ (2021-2024).

References

Abdi, H. & Valentin, D. (2007). Multiple correspondence analysis. In Encyclopedia of measurement and statistics. Thousand Oaks, (CA): Sage.

Abdi, H., Valentin, D., Chollet, S., Chrea, C., Abdi, H., Valentin, D., et al. (2007). Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. *Food Quality and Preference, 18*, 627–650.

Agudelo, A., Varela, P., & Fiszman, S. (2015). Fruit Filling development. A multiparametric approach. *LWT – Food Science Technol., 61*, 564–572.

Ares, G., Castura, J. C., Antúnez, L., Vidal, L., Giménez, A., Coste, B., et al. (2016). Comparison of two TCATA variants for dynamic sensory characterization of food products. *Food Quality and Preference, 54*, 160–172.

Bajec, M. R., & Pickering, G. J. (2010). Association of thermal taste and PROP responsiveness with food liking, neophobia, body mass index, and waist circumference. *Food Quality and Preference, 21*, 589–601.

Berget, I., Castura, J. C., Ares, G., Næs, T., & Varela, P. (2020). Exploring the common and unique variability in TDS and TCATA data – a comparison using canonical correlation and orthogonalization. *Food Quality and Preference, 79*, Article 103790.
Bougeard, S., Qannari, E. M., Lupo, C., & Hanafi, M. (2011). From Multiblock Partial Least Squares to Multiblock Redundancy Analysis. *A Continuum Approach. Informatica, 22*(1), 11–26.
Bro, R., Qannari, E. M., Kiers, H. A. L., Næs, T., & Frøst, M. B. (2008). Multi-way Models for Sensory Profiling Data. *Journal of Chemometrics, 22*, 36–45.
Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling vi an N-way generalisation of "Eckart-Young" decomposition. *Psykometrika, 35*, 238–319.
Castura, J.C. (2022). tempR: Temporal Sensory Data Analysis. R package version 0.9.9.20. http://www.cran.r-project.org/package=tempR/.
Castura, J. C., Antúnez, L., Giménez, A., & Ares, G. (2016). Temporal Check-All-That-Apply. (TCATA). A novel dynamic method for characterizing products. *Food Quality and Preference, 47*, 79–90.
Castura, J. C., Baker, A. K., & Ross, C. F. (2016). Using contrails and animated sequences to visualize uncertainty in dynamic sensory profiles obtained from temporal check-all-that-apply (TCATA) data. *Food Quality and Preference, 54*, 90–100.
Castura, J. C., Rutledge, D. N., Ross, C. F., & Næs, T. (2022). Discriminability and uncertainty in principal component analysis (PCA) of temporal check-all-that-apply (TCATA) data. *Food Quality and Preference, 96*, Article 104370.
Castura, J. C., Varela, P., & Næs, T. (2023). Evaluation of complementary numerical and visual approaches for investigating pairwise comparisons after principal component analysis. *Food Quality and Preference, 107*, Article 104843.
Cruz, A., & Green, B. G. (2000). Thermal stimulation of taste. *Nature, 403*(6772), 889–892. https://doi.org/10.1038/35002581
Dahl, T., & Næs, T. (2009). Identifying outlying assessors in sensory profiling using fuzzy clustering and multiblock methodology. *Food Quality and Preference, 30*(4), 287–294.
Endrizzi, I., Gasperi, F., Rødbotten, M., & Næs, T. (2014). Interpretation, validation and segmentation of preference mapping models. *Food Quality and Preference, 32*, 198–209.
Eslami, A., Qannari, E. M., Kohler, A., & Bougeard, S. (2014). Algorithms for multi-group PLS. *Journal of Chemometrics, 28*(3), 192–201.
Faye, P., Brémaud, D., Daubin, M. D., Courcoux, P., Giboreau, A., & Nicod, H. (2004). Perceptive free sorting and verbalization tasks with naive subjects: An alternative to descriptive mappings. *Food Quality and Preference, 15*(7–8), 781–791.
Greenacre, M. J. (1993). *Correspondence analysis in practice*. London: Academic Press.
Hanafi, M., & Kiers, H. A. L. (2006). Analysis of k sets of data, with differential emphasis on agreement between and within sets. *Computational Statistics & Data Analysis, 51*(3), 1491–1508.
Helgesen, T., & Næs, T. (1995). Selection of dry fermented lamb sausages for consumer testing. *Food Quality and Preference, 6*, 109–120.
Hirschfeld, H. O. (1935). A connection between correlation and contingency. *Mathematical Proceedings of the Cambridge Philosophical Society, 31*, 520–524.
Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology, 24*, 417–441.
Husson, F., & Pagès, J. (2006). INDSCAL model: Geometrical interpretation and methodology. *Computational Statistics and Data Analysis, 50*, 358–378.
Kruskal, J. B. (1964). Nonmetric multidimensionnal scaling: A numerical method. *Psychometrika, 29*, 115–129.
Lawless, H. T., & Heymann, H. (2010). *Sensory Evaluation of Food: Principles and Practices* (2nd edition). New York: Springer.
Jolliffe, I. T. (2010). *Principal component analysis*. Berlin: Springer.
Martens, H., & Næs, T. (1989). *Multivariate calibration*. Chichester, UK: John Wiley and Sons.
Martens, H. & Martens, M (2001). *Multivariate analysis of quality. An introduction* .John Wiley and Sons, Chichester, UK.
Martens, H., Anderssen, E., Flatberg, A., Gidskehaug, L. H., Høy, M., Westad, F., et al. (2005). Regression of a data matrix on descriptors of both rows and of its columns via latent variables: L-PLSR. *Computational Statistics and Data analysis, 48*, 103–123.
McEwan, J. A. (1996). Preference mapping for product optimisation. In T. Naes, & E. Risvik (Eds.), *Multivariate analysis of data in sensory science* (pp. 71–100). Amsterdam: Elsevier.
Miller, P., Swanson, R. E., & Heckler, C. E. (1998). Contribution plots: A missing link in multivariate quality control. *Applied Mathematics & Computer Science, 8*(4), 775–792.
Moskowitz, D., Moskowitz, H.R., & Maier, A. (2017). Category appraisal and ingredient search: identifying key sensory factors and product features at the early development stage. In J.H. Beckley, L.J. Herzog & M.M. Foley (eds.): *Accelerating New Food Product Design and Development* (pp. 259-297). https://doi.org/10.1002/9781119149330.ch17.
Mitchell, J., Castura, J. C., Thibodeau, M., & Pickering, G. (2019). Application of TCATA to examine variation in beer perception due to thermal taste status. *Food Quality and Preference, 73*, 135–142.
Måge, I., Mevik, B. H., & Næs, T. (2008). Regression models with process variables and parallel blocks of raw material measurements. *Journal of Chemometrics, 22*, 443–456.
Måge, I., Menichelli, E., & Næs, T. (2012). Preference mapping by PO-PLS: Separating common and unique information in several data blocks. *Food Quality and Preference, 24*, 8–16.
Næs, T., Berget, I., Liland, K. H., Ares, G., & Varela, P. (2017). Estimating and interpreting more than two consensus components in projective mapping: INDSCAL vs. Multiple Factor Analysis (MFA). *Food Quality and Preference, 58*, 45–60.
Næs, T., Varela, P., & Berget, I. (2018). *Individual differences in sensory and consumer science*. UK: Elsevier.
Næs, T., Romano, R., Tomic, O., Måge, I., Smiles, A. K., & Liland, K. H. (2020). Sequential and orthogonalized PLS (SO-PLS) regression for path analysis. Order of blocks and relations between effects. *Journal of Chemometrics.*, e3243.
Næs, T., Tomic, O., Endrizzi, I., & Varela, P. (2020). Principal components analysis of descriptive sensory data; reflections, challenges and suggestions. *Journal of Sensory Studies*, e12692. https://doi.org/10.1111/joss.12692
Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire valley. *Food Quality and Preference, 16*, 652–659.
Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine, 2*, 559–572.
Pineau, N., Schlich, P., Cordelle, S., Mathonniere, C., Issanchou, S., Imbert, A., et al. (2009). Temporal dominance of sensations: Construction of the TDS curves and comparison with time-intensity. *Food Quality and Preference., 20*, 450–455.
R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.
Risvik, E., MacEwan, J. A., Colwill, J. S., Rogers, R., & Lyon, D. H. (1994). Projective mapping: A tool for sensory analysis and consumer research. *Food Quality and Preference, 5*, 263–269.
Robert, P., & Escofier, Y. (1976). A unifying tool for linear multivariate statistical methods: The RV-coefficient. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 25*(3), 257–265.
Santos, B. A., Pollonio, M. A. R., Cruz, A. G., Messias, V. C., Monteiro, R. A., Oliveira, T. L. C., Faria, J. A. F., Freitas, M. Q., & Bolini, H. M. A. (2013). Ultra-flash profile and projective mapping for describing sensory attributes of prebiotic mortadellas. *Food Research International, 54*(2), 1705–1711.
Schlich, P. (1996). Defining and validating assessor compromises about product distances and attribute correlations. In T. Naes, & E. Risvik (Eds.), *Multivariate analysis of data in sensory science* (pp. 259–306). Amsterdam: Elsevier.
Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). Introduction to multidimensional scaling. *Theory, methods and applications*. Orlando: Academic Press.
Sinopoli, D. A., & Lawless, H. T. (2012). Taste Properties of Potassium Chloride Alone and in Mixtures with Sodium Chloride Using a Check - All - That - Apply Method. *Journal of Food Science, 77*, S319–S322.
Smilde, A.K., Bro, R., & Geladi, P. (2004). Multiway analysis. *Applications in the chemical sciences*. John Wiley and sons, NY.
Smilde, A. K., Måge, I., Næs, T., Hankemeier, T., Lips, M. A., Kiers, H. A., et al. (2017). Common and distinct components in data fusion. *Journal of Chemometrics, 31*(7), e2900.
Smilde, A. K., Næs, T., & Liland, K. H. (2022). *Multiblock data fusion in statistics and machine learning*. UK: Wiley.
Stone, M. (1974). Cross-validatory choice and assessment of statistical prediction. *Journal of the Royal Statistical Society, B*, 111–113.
Stone, H., Sidel, J., Oliver, S., Woolsey, A., & Singleton, R. C. (1974). Sensory evaluation by quantitative descriptive analysis. *Food Technology, 28*(11), 24–34.
Tauler, R. (2001). Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution. *Journal of Chemometrics, 15*(8), 627–646.
Tenenhaus, A., & Tenenhaus, M. (2014). Regularized generalized canonical correlation analysis for multiblock or multigroup data analysis. *European Journal of Operational Research, 238*(2), 391–403.
Thibodeau, M., Saliba, A., Bajec, M., & Pickering, G. (2019). Examination and validation of classification schema for determining thermal taste status. *Chemosensory Perception, 12*, 69–89.
Tomic, O., Nilsen, A., Martens, M., & Næs, T. (2007). Graphical techniques for detection of individual differences among assessors in sensory profiling. *LWT Food Science and Technology, 40*, 262–269.
Tomic, O., Luciano, G., Nilsen, A., Hyldig, G., Lorensen, K., & Næs, T. (2010). Analysing sensory panel performance in a proficiency test using PanelCheck software. *European Food Research and Technology, 230*(3), 497–511.
Tucker, L. (1964). The extension of factor analysis to three-dimensional matrices. *Contributions to Mathematical Psychology, 110119*.
Van der Burg, E & Dijksterhuis, G. (1996). Generalised canonical analysis of individual sensory profiles and instrument data. In. Næs, T. and Risvik, E. (Eds.): *Multivariate analysis of data in sensory science*. Elsevier. 221-258.
Vidal, L., Castura, J. C., Coste, B., Picallo, A., Jaeger, S. R., & Ares, G. (2017). Analysis of TCATA Fading data: Imputation of gaps in temporal profiles. *Food Quality and Preference, 59*, 114–122.
Yang, Q., Hollowood, T., & Hort, J. (2014). Phenotypic variation in oronasal perception and the relative effects of PROP and thermal taster status. Food Quality and Preference, 38, 83–91.Wold, S. (1978). Cross-validatory estimation of the number of components in factors and principal components models. *Technometrics, 20*, 397–405.
Wold, S., Martens, H., & Wold, H. (1983). The multivariate calibration problem in chemistry solved by the PLS method. Proc. Conf. Matrix Pencils (A. Ruhe, B. Kågstrøm eds.), March 1982. Lecture notes in mathematics, Springer Verlag, Heidelberg, 286-293.

Further reading

Næs, T., Brockhoff, P., & Tomic, O. (2010). *Statistics for sensory and consumer science*. UK: Wiley.