Multiple factor analysis: principal component analysis for multitable and multiblock data sets

Hervé Abdi[1], Lynne J. Williams[2] and Domininique Valentin[3]

Multiple factor analysis (MFA, also called multiple factorial analysis) is an extension of principal component analysis (PCA) tailored to handle multiple data tables that measure sets of variables collected on the same observations, or, alternatively, (in dual-MFA) multiple data tables where the same variables are measured on different sets of observations. MFA proceeds in two steps: *First* it computes a PCA of each data table and 'normalizes' each data table by dividing all its elements by the first singular value obtained from its PCA. *Second*, all the normalized data tables are aggregated into a grand data table that is analyzed via a (non-normalized) PCA that gives a set of factor scores for the observations and loadings for the variables. In addition, MFA provides for each data table a set of partial factor scores for the observations that reflects the specific 'view-point' of this data table. Interestingly, the common factor scores could be obtained by replacing the original normalized data tables by the normalized factor scores obtained from the PCA of each of these tables. In this article, we present MFA, review recent extensions, and illustrate it with a detailed example. © 2013 Wiley Periodicals, Inc.

INTRODUCTION

Multiple factor analysis (MFA, also sometimes named 'multiple factorial analysis' to avoid the confusion with Thurstone's multiple factor analysis described in Ref 1) is a generalization of principal component analysis (PCA). Its goal is to analyze several data sets of variables collected on the same set of observations, or—as in its dual version—several sets of observations measured on the same set of variables

(see Ref 2). As such, MFA is part of the multitable (also called multiblock or consensus analysis[3–20]) PCA family which comprises related techniques such as STATIS, multiblock correspondence analysis (MUDICA), and SUM-PCA.

MFA is a recent technique (ca 1980) that originated from the work of the French statisticians Brigitte Escofier and Jérôme Pagès (see Refs 14,21,22, for an introduction and for example see Ref 23, for an extensive and comprehensive review see Ref 24). The goals of MFA are (1) to analyze several data sets measured on the same observations; (2) to provide a set of common factor scores (often called 'compromise factor scores'); and (3) to project each of the original data sets onto the compromise to analyze communalities and discrepancies. The main

*Correspondence to: herve@utdallas.edu

[1]Department of Brain and Behavioral Sciences, The University of Texas at Dallas, Richardson, TX, USA

[2]Rotman Research Institute, Baycrest, Toronto, ON, Canada

[3]Université de Bourgogne, Dijon, France

idea behind MFA is remarkably simple and akin to the idea behind the *Z*-score normalization that makes variables comparable by dividing each element of a variable by the variable standard deviation (i.e., the square root of the variance) of this variable. For a PCA a notion similar to the standard deviation is the singular value which is the square root of an eigenvalue (which can be seen as a variance). So, in MFA each data table is normalized by dividing all of its elements by the first singular value of this data table. This transformation ensures that the length (i.e., the singular value) of the first principal component of each data table is equal to 1 and therefore that no data table can dominate the common solution only because it has a larger inertia on its first dimension.

MFA is a popular method for analyzing multiple sets of variables measured on the same observations and it has been recently used in various domains, such as sensory and consumer science research (a domain where MFA applications and developments have been particularly rich and varied, see Refs 9,21,25–40), chemometry and process monitoring,[9,41–43] ecology,[44–53] agriculture,[54,55] broadcasting,[56] geology,[57] neuroimaging,[4,5,58–60] medicine and health,[61–64] genetics,[54,65,66] statistical quality control,[27,67–72] economy,[73] and molecular biology to name but a few.[46,50] In addition to being used in several domains of applications, MFA is also a vigourous domain of theoretical developments that are explored later in this article.

When to Use It

MFA is used when several sets of variables have been measured on the same set of observations. The number and/or nature of the variables used to describe the observations can vary from one set of variables to the other, but the observations should be the same in all the data sets.

For example, the data sets can be measurements taken on the same observations (individuals or objects, e.g., students) at different occasions (e.g., semesters). In this case, the first data set corresponds to the data collected at time 1 (e.g., the first semester), the second one to the data collected at time 2 (e.g., the second semester) and so on. The goal of the analysis, then is to evaluate how the positions of the observations change over time (note, however, that MFA does not *explicitly* models the time variable as it does not make assumptions about the relationships between the measurements).

In another example, the different data sets can be the same observations (e.g., wines) evaluated by different subjects (e.g., wine experts) or groups of subjects with different variables (e.g., each wine expert evaluates the wines with his/her own set of scales). In this case, the first data set corresponds to the first subject, the second one to the second subject and so on. The goal of the analysis, then, is to evaluate if there is an agreement between the subjects or groups of subjects.

Alternatively, dual-MFA can be used when the *same* variables are measured on different populations or on different participants. When both observations and variables are the same for all data tables the technique could be called (by analogy with STATIS, see, Ref 3) 'partial triadic MFA'.

The Main Idea

The general idea behind MFA is to normalize each of the individual data sets so that their first principal component has the same length (as measured by the first singular value of each data table) and then to combine these data tables into a common representation of the observations sometimes called the *compromise*, or the *consensus*. This compromise is obtained from a (non-normalized) PCA of the grand table obtained from the concatenation of the normalized data tables. This PCA decomposes the variance of the compromise into a set of new orthogonal variables (i.e., the principal components also often called dimensions, axes, factors, or even latent variables) ordered by the amount of variance that each component explains. The coordinates of the observations on the components are called *factor scores* and these can be used to plot maps of the observations in which the observations are represented as points such that the distances in the map best reflect the similarities between the observations. The positions of the observations 'as seen by' each data set are called *partial factor scores* and can be also represented as points in the compromise map. The average of the factor scores of all the tables gives back the factor score of the compromise. A pictorial sketch of the technique is provided in Figure 1.

As the components are obtained by combining the original variables, each variable 'contributes' a certain amount to each component. This quantity, called the *loading* of a variable on a component, reflects the importance of that variable for this component and can also be used to plot maps of the variables that reflect their association. Squared loadings can also be used to evaluate the importance of variables. A variation over squared loadings, called *contributions* evaluate the importance of each variable as the proportion of the explained variance of the component by the variable. The contribution of a data table to a component can be obtained by adding the contributions of its variables. These contributions can then be
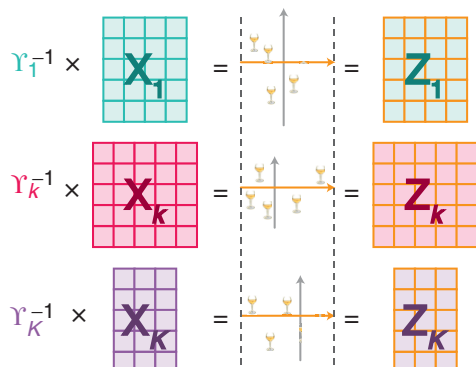
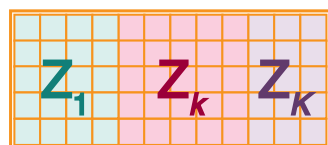**Step 1:** K tables of J_k variables collected on the same observations
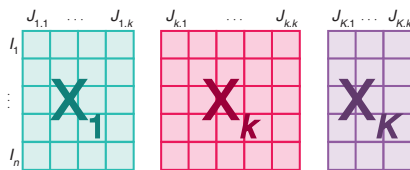
Step 2: Compute generalized PCA on each of the K tables (where Υ is the first singular value of each table)
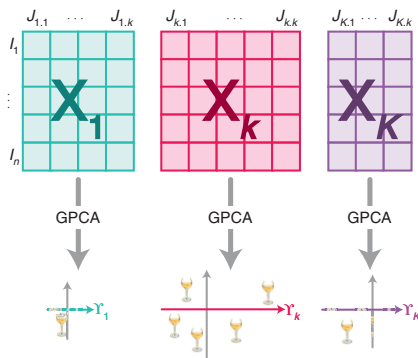
Step 3: Normalize each table by dividing by its first singular value (Υ)

Step 4: Concatenate the K normalized tables

Step 5: Compute a generalized PCA on the concatenated table
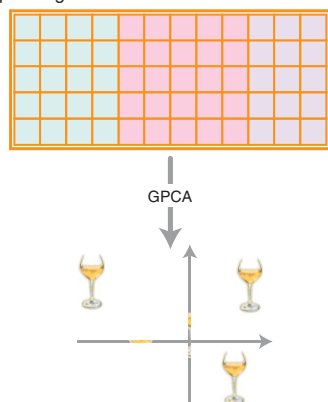
FIGURE 1 | The different steps of MFA.

used to draw plots expressing the importance of the data tables in the common solution.

NOTATIONS AND PRELIMINARIES

Matrices are denoted by boldface uppercase letters (e.g., $\mathbf{X}$), vectors by boldface lowercase letters (e.g., $\mathbf{q}$), elements of vectors and matrices are denoted by italic lower case letters with appropriate indices if needed (e.g., x_{ij} is an element of $\mathbf{X}$). Blocks of variables (i.e., tables) are considered as sub-matrices of larger matrices and are represented in brackets separated by vertical bars (e.g., a matrix $\mathbf{X}$ made of two sub-matrices $\mathbf{X}_{[1]}$ and $\mathbf{X}_{[2]}$ is written $\mathbf{X} = [\mathbf{X}_{[1]}|\mathbf{X}_{[2]}]$). The identity matrix is denoted by $\mathbf{I}$, a vector of ones is denoted by $\mathbf{1}$ (indices may be used to specify the dimensions if the context is ambiguous). The transpose of a matrix is denoted by $^\top$. The inverse of a matrix is denoted by $^{-1}$. When applied to a square matrix, the diag operator takes the diagonal elements of this matrix and stores them into a column vector; when applied to a vector, the diag operator stores the elements of this vector on the diagonal elements of a diagonal matrix. The trace operator computes the sum of the diagonal elements of a square matrix. The vec operator transforms a matrix into a vector by stacking the elements of a matrix into a column vector. The standard product between matrices is implicitly denoted by simple juxtaposition or by $\times$ when it needs to be explicitly stated (e.g., $\mathbf{XY} = \mathbf{X} \times \mathbf{Y}$ is the product of matrices $\mathbf{X}$ and $\mathbf{Y}$). The Hadamard or element-wise product is denoted by $\circ$ (e.g., $\mathbf{X} \circ \mathbf{Y}$).

The raw data consist of K data sets collected on the same observations. Each data set is also called a table, a sub-table, or sometimes also a block or a study (in this article we prefer the term table or occasionally block). The data for each table are stored in an $I \times J_{[k]}$ rectangular data matrix denoted by $\mathbf{Y}_{[k]}$, where I is the number of observations and $J_{[k]}$ the number of variables collected on the observations for the k-th table. The total number of variables is denoted by J (i.e., $J = \sum J_{[k]}$). Each data matrix is, in general, preprocessed (e.g., centered, normalized) and the preprocessed data matrices actually used in the analysis are denoted by $\mathbf{X}_{[k]}$ (the preprocessing steps are detailed below in section *More on Preprocessing*).

The K data matrices $\mathbf{X}_{[k]}$, each of dimensions I rows by $J_{[k]}$ columns, are concatenated into the complete I by J data matrix denoted by $\mathbf{X}$:

$$\mathbf{X} = \left[\mathbf{X}_{[1]}|\dots|\mathbf{X}_{[k]}|\dots|\mathbf{X}_{[K]}\right]. \qquad (1)$$

A mass, denoted by m_i, is assigned to each observation. These masses are collected in the mass

vector, denoted by **m**, and in the diagonal elements of the mass matrix denoted by **M**, which is obtained as

$$\mathbf{M} = \text{diag}\{\mathbf{m}\}. \tag{2}$$

Masses are non-negative elements whose sum equals one. Often, equal masses are chosen with $m_i = \frac{1}{I}$.

To each matrix $\mathbf{X}_{[k]}$, we associate its *cross-product* matrix defined as

$$\mathbf{S}_{[k]} = \mathbf{X}_{[k]}\mathbf{X}_{[k]}^{\mathsf{T}}. \tag{3}$$

A cross-product matrix of a table expresses the pattern of relationships between the observations as seen by this table. Note that because of the block structure of **X**, the cross product of **X** can be expressed as

$$
\begin{aligned}
\mathbf{X}\mathbf{X}^{\mathsf{T}} &= \left[\mathbf{X}_{[1]}|\ldots|\mathbf{X}_{[k]}|\ldots|\mathbf{X}_{[K]}\right] \\
&\quad \times \left[\mathbf{X}_{[1]}|\ldots|\mathbf{X}_{[k]}|\ldots|\mathbf{X}_{[K]}\right]^{\mathsf{T}} \\
&= \sum_k \mathbf{X}_{[k]}\mathbf{X}_{[k]}^{\mathsf{T}} \\
&= \sum_k \mathbf{S}_{[k]}.
\end{aligned}
\tag{4}
$$

Singular Value Decomposition and Generalized Singular Value Decomposition

MFA is part of the PCA family and therefore its main analytical tool is the singular value decomposition (SVD) and the generalized singular value decomposition (GSVD) of a matrix (see for tutorials, Refs 74–80). We briefly describe these two methods below.

SVD

Recall that the SVD of a given $I \times J$ matrix **Z** decomposes it into three matrices as:

$$\mathbf{X} = \mathbf{U}\boldsymbol{\Gamma}\mathbf{V}^{\mathsf{T}} \quad \text{with} \quad \mathbf{U}^{\mathsf{T}}\mathbf{U} = \mathbf{V}^{\mathsf{T}}\mathbf{V} = \mathbf{I} \tag{5}$$

where **U** is the I by L matrix of the normalized left singular vectors (with L being the rank of **X**), **V** the J by L matrix of the normalized right singular vectors, and $\boldsymbol{\Gamma}$ the L by L diagonal matrix of the L singular values; also γ_ℓ, $\mathbf{u}_\ell$, and $\mathbf{v}_\ell$ are, respectively, the ℓth singular value, left, and right singular vectors. Matrices **U** and **V** are orthonormal matrices (i.e., $\mathbf{U}^{\mathsf{T}}\mathbf{U} = \mathbf{V}^{\mathsf{T}}\mathbf{V} = \mathbf{I}$). The SVD is closely related to and generalizes the well-known *eigendecomposition* as **U** is also the matrix of the normalized eigenvectors of $\mathbf{X}\mathbf{X}^{\mathsf{T}}$, **V** is the matrix of the normalized eigenvectors of $\mathbf{X}^{\mathsf{T}}\mathbf{X}$, and the singular values are the square root of

the eigenvalues of $\mathbf{X}\mathbf{X}^{\mathsf{T}}$ and $\mathbf{X}^{\mathsf{T}}\mathbf{X}$ (these two matrices have the same eigenvalues).

Key property: the SVD provides the best reconstitution (in a least squares sense) of the original matrix by a matrix with a lower rank.

GSVD

The GSVD generalizes the SVD of a matrix by incorporating two additional positive definite matrices (recall that a positive definite matrix is a square symmetric matrix whose eigenvalues are all positive) that represent 'constraints' to be incorporated in the decomposition (formally, these matrices are constraints on the orthogonality of the singular vectors, see Refs 75,78 for more details). Specifically let **M** denote an I by I positive definite matrix representing the 'constraints' imposed on the rows of an I by J matrix **X**, and **A** a J by J positive definite matrix representing the 'constraints' imposed on the columns of **X**. Matrix **M** is almost always a diagonal matrix of the 'masses' of the observations (i.e., the rows); whereas matrix **A** implements a metric on the variables and if often but not always diagonal. Obviously, when $\mathbf{M} = \mathbf{A} = \mathbf{I}$, the GSVD reduces to the plain SVD. The GSVD of **X**, taking into account **M** and **A**, is expressed as (compare with Eq. (5)):

$$\mathbf{X} = \mathbf{P}\boldsymbol{\Delta}\mathbf{Q}^{\mathsf{T}} \quad \text{with} \quad \mathbf{P}^{\mathsf{T}}\mathbf{M}\mathbf{P} = \mathbf{Q}^{\mathsf{T}}\mathbf{A}\mathbf{Q} = \mathbf{I} \tag{6}$$

where **P** is the I by L matrix of the normalized left generalized singular vectors (with L being the rank of **X**), **Q** the J by L matrix of the normalized generalized right singular vectors, and $\boldsymbol{\Delta}$ the L by L diagonal matrix of the L generalized singular values. The GSVD implements the whole class of generalized PCA which includes (with a proper choice of matrices **M** and **A** and preprocessing of **X**) techniques such as discriminant analysis, correspondence analysis, canonical variate analysis, etc. With the so called 'triplet notation', that is used as a general framework to formalize multivariate techniques, the GSVD of **X** under the constraints imposed by **M** and **A** is equivalent to the statistical analysis of the triplet $(\mathbf{X}, \mathbf{A}, \mathbf{M})$ (see, Refs 81–85).

Key property: the GSVD provides the best reconstitution (in a least squares sense) of the original matrix by a matrix with a lower rank under the constraints imposed by two positive definite matrices. The generalized singular vectors are orthonormal with respect to their respective matrix of constraints.

THE DIFFERENT STEPS OF MFA

MFA comprises three main steps: In the *first* step, a PCA of each data table is performed and the first

singular value of each table recorded. In the *second* step a grand matrix is created by concatenating all the data tables and a non-normalized generalized PCA is performed by decomposing the grand matrix with a GSVD where the column weights are obtained from the first (squared) singular value of each data table. An equivalent way of performing this second step is to divide all the elements of a data table by the table's first singular value, concatenate the data tables into a grand data table and then perform a non-normalized PCA of this grand data table. In the *third* step, the observation partial factor scores for each table are computed by projecting each data table onto the common space.

As MFA boils down to the PCA of the grand data table, the usual PCA indices can be computed to identify the important components, observations, and variables (see Ref 75). In addition some indices—specific to MFA—can also be derived to quantify the importance of each table in the common solution.

Step 1: PCA of Each Data Table

In the first step of MFA each data table is analyzed via a standard PCA. Specifically, each table is expressed via its SVD as

$$\mathbf{X}_{[k]} = \mathbf{U}_{[k]}\mathbf{\Gamma}_{[k]}\mathbf{V}_{[k]}^{\mathsf{T}} \text{ with } \mathbf{U}_{[k]}^{\mathsf{T}}\mathbf{U}_{[k]} = \mathbf{V}_{[k]}^{\mathsf{T}}\mathbf{V}_{[k]} = \mathbf{I}. \quad (7)$$

From the SVD of each table we also obtain its *factor scores* (as in any standard PCA) that are computed as

$$\mathbf{G}_{[k]} = \mathbf{U}_{[k]}\mathbf{\Gamma}_{[k]}. \quad (8)$$

The matrices $\mathbf{U}_{[k]}$ and $\mathbf{V}_{[k]}$ store (respectively) the left and right singular vectors of the table $\mathbf{X}_{[k]}$, whose singular values are stored in the diagonal of the (diagonal) matrix $\mathbf{\Gamma}_{[k]}$

$$[\gamma_{1,k}, \ldots, \gamma_{\ell,k}, \ldots, \gamma_{L,k}] = \text{diag}\{\mathbf{\Gamma}_{[k]}\}. \quad (9)$$

In MFA, the weight of a table is obtained from the first singular value of its PCA. This weight, denoted by α_k, is equal to the inverse of the first squared singular value:

$$\alpha_k = \frac{1}{\gamma_{1,k}^2} = \gamma_{1,k}^{-2}. \quad (10)$$

For convenience, the α weights can be gathered in a J by 1 vector denoted $\mathbf{a}$ where each variable is assigned the α weight of the matrix to which it belongs. Specifically, $\mathbf{a}$ is constructed as:

$$\mathbf{a} = \left[\alpha_1\mathbf{1}_{[1]}^{\mathsf{T}}, \ldots, \alpha_k\mathbf{1}_{[k]}^{\mathsf{T}}, \ldots, \alpha_K\mathbf{1}_{[K]}^{\mathsf{T}}\right], \quad (11)$$

where $\mathbf{1}_{[k]}$ stands for a $J_{[k]}$ vector of ones. Alternatively, the weights can be stored as the diagonal elements of a diagonal matrix denoted by $\mathbf{A}$ obtained as

$$\mathbf{A} = \text{diag}\{\mathbf{a}\}$$
$$= \text{diag}\left\{\left[\alpha_1\mathbf{1}_{[1]}^{\mathsf{T}}, \ldots, \alpha_k\mathbf{1}_{[k]}^{\mathsf{T}}, \ldots, \alpha_K\mathbf{1}_{[K]}^{\mathsf{T}}\right]\right\}. \quad (12)$$

Step 2: Generalized PCA of X
GSVD of X

After the weights have been collected, they are used to compute the GSVD of $\mathbf{X}$ under the constraints provided by $\mathbf{M}$ (masses for the observations) and $\mathbf{A}$ (squared singular value derived weights for the K tables). This GSVD is expressed as:

$$\mathbf{X} = \mathbf{P}\mathbf{\Delta}\mathbf{Q}^{\mathsf{T}} \text{ with } \mathbf{P}^{\mathsf{T}}\mathbf{M}\mathbf{P} = \mathbf{Q}^{\mathsf{T}}\mathbf{A}\mathbf{Q} = \mathbf{I}. \quad (13)$$

This GSVD corresponds to a generalized PCA of matrix $\mathbf{X}$ and, consequently, will provide *factor scores* to describe the observations and *factor loadings* to describe the variables. Each column of $\mathbf{P}$ and $\mathbf{Q}$ refers to a *principal component* also called a *dimension* (because the numbers in these columns are often used as coordinates to plot maps, see 75 for more details). In PCA, Eq. (13) is often rewritten as

$$\mathbf{X} = \mathbf{F}\mathbf{Q}^{\mathsf{T}} \text{ with } \mathbf{F} = \mathbf{P}\mathbf{\Delta} \quad (14)$$

where $\mathbf{F}$ stores the factor scores (describing the observations) and $\mathbf{Q}$ stores the loadings (describing the variables). Note, incidentally, that in the triplet notation, MFA is equivalent to the statistical analysis of the triplet $(\mathbf{X}, \mathbf{A}, \mathbf{M})$.

Because the matrix $\mathbf{X}$ concatenates K tables, each of them, in turn, comprising $J_{[k]}$ variables, the matrix $\mathbf{Q}$ of the left singular vectors can be partitioned in the same way as $\mathbf{X}$. Specifically, $\mathbf{Q}$ can be expressed as a column block matrix as:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{[1]} \\ \vdots \\ \mathbf{Q}_{[k]} \\ \vdots \\ \mathbf{Q}_{[K]} \end{bmatrix} = \left[\mathbf{Q}_{[1]}^{\mathsf{T}}|\ldots|\mathbf{Q}_{[k]}^{\mathsf{T}}|\ldots|\mathbf{Q}_{[K]}^{\mathsf{T}}\right]^{\mathsf{T}}, \quad (15)$$

where $\mathbf{Q}_{[k]}$ is a $J_{[k]}$ by L (with L being the rank of $\mathbf{X}$) matrix storing the right singular vectors corresponding to the variables of matrix $\mathbf{X}_{[k]}$. With this in mind,

Eq. (13) can re-expressed as:

$$\mathbf{X} = \left[\mathbf{X}_{[1]} | \ldots | \mathbf{X}_{[k]} | \ldots | \mathbf{X}_{[K]}\right] = \mathbf{P}\boldsymbol{\Delta}\mathbf{Q}^{\mathsf{T}}$$

$$= \mathbf{P}\boldsymbol{\Delta} \left(\left[\mathbf{Q}_{[1]}^{\mathsf{T}} | \ldots | \mathbf{Q}_{[k]}^{\mathsf{T}} | \ldots | \mathbf{Q}_{[K]}^{\mathsf{T}} \right]^{\mathsf{T}} \right)^{\mathsf{T}}$$

$$= \mathbf{P}\boldsymbol{\Delta} \left[\mathbf{Q}_{[1]}^{\mathsf{T}} | \ldots | \mathbf{Q}_{[k]}^{\mathsf{T}} | \ldots | \mathbf{Q}_{[K]}^{\mathsf{T}} \right]$$

$$= \left[\mathbf{P}\boldsymbol{\Delta}\mathbf{Q}_{[1]}^{\mathsf{T}} | \ldots | \mathbf{P}\boldsymbol{\Delta}\mathbf{Q}_{[k]}^{\mathsf{T}} | \ldots | \mathbf{P}\boldsymbol{\Delta}\mathbf{Q}_{[K]}^{\mathsf{T}} \right]. \qquad (16)$$

Note, that, the pattern in Eq. (13) does not completely generalize to Eq. (16) because, if we define $\mathbf{A}_{[k]}$ as

$$\mathbf{A}_{[k]} = \alpha_k \mathbf{I}, \qquad (17)$$

we have, in general, $\mathbf{Q}_{[k]}^{\mathsf{T}}\mathbf{A}_{[k]}\mathbf{Q}_{[k]} \neq \mathbf{I}$.

Factor Scores

The factor scores for $\mathbf{X}$ represent a compromise (i.e., a common representation) for the set of the K matrices. Recall that these compromise factor scores, are computed (cf., Eqs (13) and (14)) as

$$\mathbf{F} = \mathbf{P}\boldsymbol{\Delta}. \qquad (18)$$

Factor scores can be used to plot the observations as done in standard PCA for which each column of $\mathbf{F}$ represents a dimension. Note that the variance of the factor scores of the observations is computed using their masses (stored in matrix $\mathbf{M}$) and can be found as the diagonal of the matrix $\mathbf{F}^{\mathsf{T}}\mathbf{M}\mathbf{F}$. This variance is equal, for each dimension, to the square of the singular value of this dimension as shown by

$$\mathbf{F}^{\mathsf{T}}\mathbf{M}\mathbf{F} = \boldsymbol{\Delta}\mathbf{P}^{\mathsf{T}}\mathbf{M}\mathbf{P}\boldsymbol{\Delta} = \boldsymbol{\Delta}^2. \qquad (19)$$

As in standard PCA, $\mathbf{F}$ can be obtained from $\mathbf{X}$ by combining Eqs (13) and (18) to get:

$$\mathbf{F} = \mathbf{P}\boldsymbol{\Delta} = \mathbf{X}\mathbf{A}\mathbf{Q}. \qquad (20)$$

Taking into account the block structure of $\mathbf{X}$, $\mathbf{A}$, and $\mathbf{Q}$, Eq. (13) can also be rewritten as (cf., Eq. (17)):

$$\mathbf{F} = \mathbf{X}\mathbf{A}\mathbf{Q} = \left[\mathbf{X}_{[1]} | \ldots | \mathbf{X}_{[k]} | \ldots | \mathbf{X}_{[K]}\right] \times \mathbf{A} \times \begin{bmatrix} \mathbf{Q}_{[1]} \\ \vdots \\ \mathbf{Q}_{[k]} \\ \vdots \\ \mathbf{Q}_{[K]} \end{bmatrix}$$

$$= \sum_k \mathbf{X}_{[k]}\mathbf{A}_{[k]}\mathbf{Q}_{[k]} = \sum_k \alpha_k \mathbf{X}_{[k]}\mathbf{Q}_{[k]}. \qquad (21)$$

This equation suggests that the *partial factor scores* for a table can be defined from the projection of this table onto its right singular vectors (i.e., $\mathbf{Q}_{[k]}$). Specifically, the partial factor scores for the kth table are stored in a matrix denoted by $\mathbf{F}_{[k]}$ computed as

$$\mathbf{F}_{[k]} = K \times \alpha_k \times \mathbf{X}_{[k]}\mathbf{Q}_{[k]}. \qquad (22)$$

Note that the compromise factor scores matrix is the *barycenter* (also called centroid or center of gravity see Ref 86) of the partial factor scores because it is the average of all K partial factor scores (cf., Eq. (20)):

$$\frac{1}{K}\sum_k \mathbf{F}_{[k]} = \frac{1}{K}\sum_k K\alpha_k \mathbf{X}_{[k]}\mathbf{Q}_{[k]} = \sum_k \alpha_k \mathbf{X}_{[k]}\mathbf{Q}_{[k]} = \mathbf{F}. \qquad (23)$$

Also as in standard PCA, the elements of $\mathbf{Q}$ are *loadings* and can be plotted either on their own or along with the factor scores as a biplot (see Refs 87,88). As the loadings come in blocks (i.e., the loadings correspond to the variables of a table), it makes sense to create a biplot with the partial factor scores (i.e., $\mathbf{F}_{[k]}$) for a block and the loadings (i.e., $\mathbf{Q}_{[k]}$) for this block. In doing so, it is often practical to normalize the loadings such that their variance is commensurable with the variance of the factor scores. This can be achieved, for example, by normalizing, for each dimension, the loadings of a block such that their variance is equal to the square of the singular value of the dimension or even to the singular value itself (as illustrated in the example that we present in a following section). These biplots are helpful for understanding the statistical structure of each block, even though the relative positions of the factor scores and the loadings are not directly interpretable because only the projections of observations on the loading vectors can be meaningfully interpreted in a biplot (cf., Refs 87,88).

An alternative pictorial representation of the variables and the components plots the correlations between the original variables of $\mathbf{X}$ and the factor scores. These correlations are plotted as two-dimensional maps in which a circle of radius one (called the *circle of correlation*[75,89]) is also plotted. The closer to the circle a variable is, the better this variable is 'explained' by the components used to create the plot (see Refs 23,24 for examples). Loadings and correlations as often used interchangeably because these two concepts are very similar and, sometimes the names loading is used for both concepts (see Ref 75). In fact, loadings and correlation differ only by a normalization factor: the sum of the squared loadings of all the variables for a given dimension is equal to

one whereas the sum of the squared correlations of all the dimensions for a given variable is equal to one (and therefore it is always possible to transform one set into the other).

HOW TO FIND THE IMPORTANT ELEMENTS: CONTRIBUTIONS, ETC.

Contributions of Observations, Variables, and Tables to a Dimension

In MFA, just like in standard PCA, the importance of a dimension (i.e., principal component) is reflected by its eigenvalue which indicates how much of the total inertia (i.e., variance) of the data is explained by this component.

To better understand the relationships between components, observations, variables, and tables and also to help interpret a component, we can evaluate how much an observation, a variable, or a whole table contribute to the inertia extracted by a component. In order to do so, we compute descriptive statistics, called *contributions* (see Refs 78,89–91 and Ref 75, p. 437ff.). The stability of these descriptive statistics can be assessed by cross-validation techniques such as the bootstrap whose results can be used to select the relevant elements for a dimension.

Contribution of an Observation to a Dimension

As stated in Eq. (19), the variance of the factor scores for a given dimension is equal to its eigenvalue (i.e., the square of the singular value) associated with this dimension. If we denote λ_ℓ, the eigenvalue of a given dimension, we can rewrite Eq. (19) as

$$\lambda_\ell = \sum_i m_i \times f_{i,\ell}^2 \qquad (24)$$

where m_i and $f_{i,\ell}$ are, respectively, the mass of the *i*th observation and the factor score of the *i*th observation for the ℓth dimension. As all the terms $m_i \times f_{i,\ell}^2$ are positive or null, we can evaluate the contribution of an observation to a dimension as the ratio of the squared weighted factor score by the dimension eigenvalue. Formally, the contribution of observation *i* to component ℓ, denoted $\mathrm{ctr}_{i,\ell}$, is computed as

$$\mathrm{ctr}_{i,\ell} = \frac{m_i \times f_{i,\ell}^2}{\lambda_\ell} . \qquad (25)$$

Contributions take values between 0 and 1, and for a given component, the sum of the contributions of all observations is equal to 1. The larger a contribution, the more the observation contributes to the component. A useful heuristic is to base the interpretation of a component on the observations that have contributions larger than the average contribution. Observations with high contributions and whose factor scores have different signs can then be contrasted to help interpreting the component. Alternatively (as described in a later section) we can derive pseudo *t* statistics (called bootstrap ratios) in order to find the observations important for a given dimension.

Contributions of a Variable to a Dimension

As we did for the observations, we can find the important variables for a given dimension by computing variable contributions. The variance of the loadings for the variables is equal to one when the α weights are taken into account (cf., Eq. (13)). So if we denote by a_j the α weight for the *j*th variable (recall that all variables from the same table share the same α weight cf., Eq. (11)), we have

$$1 = \sum_j a_j \times q_{j,\ell}^2 \qquad (26)$$

where $q_{i,\ell}$ is the loading of the *j*th variable for the ℓth dimension. As all terms $a_j \times q_{i,\ell}^2$ are positive or null, we can evaluate the contribution of a variable to a dimension as its squared weighted loading for this dimension. Formally, the contribution of variable *j* to component ℓ, denoted $\mathrm{ctr}_{j,\ell}$, is computed as

$$\mathrm{ctr}_{j,\ell} = a_j \times q_{j,\ell}^2 . \qquad (27)$$

Variable contributions take values between 0 and 1, and for a given component, the contributions of all variables sum to 1. The larger a contribution of a variable to a component the more this variable contributes to this component. Variables with high contributions and whose loadings have different signs can then be contrasted to help interpreting the component.

Contribution of a Table to a Dimension

Specific to multiblock analysis is the notion of a table contribution. As a table comprises several variables, the contribution of a table can simply be defined as the sum of the contributions of its variables (a simple consequence of the Pythagorean theorem that states that *squared* lengths are additive). So the contribution of table *k* to component ℓ is denoted $\mathrm{ctr}_{k,\ell}$ and is defined as

$$\mathrm{ctr}_{k,\ell} = \sum_j^{J_{[k]}} \mathrm{ctr}_{j,\ell} . \qquad (28)$$

Table contributions take values between 0 and 1, and for a given component, the contributions of all tables sum to 1. The larger a contribution of a table to a component, the more this table contributes to this component. The contributions of the tables for a given dimension sum to one, an alternative approach re-scales the contributions so that the sum of the contributions for a dimension is now equal to the eigenvalue of this dimension. These re-scaled contributions are called *partial inertias* and are denoted $\mathcal{I}_{\text{partial}}$. The partial inertias are obtained from the contributions by multiplying the contributions for a dimension by the dimension eigenvalue.

Table contributions and partial inertias can be used to create plots that show the importance of these tables for the components. These plots can be drawn one component at a time or two (or rarely three) components at a time in a manner analogous to factor maps.

How to Analyze the Between-Table Structure

To evaluate the similarity between two tables one can compute coefficients of similarity between data tables. A traditional coefficient is Escoufier's R_V coefficient (see Refs 92,93, see, also Refs 94 and 95 for alternatives), which can be interpreted as a non centered squared coefficient of correlation between two matrices. The R_V coefficient varies between 0 and 1 and reflects the amount of variance shared by two matrices. Specifically, the R_V coefficient between data tables k and k' is computed as

$$R_{Vk,k'} = \frac{\text{trace}\left\{\left(\mathbf{X}_{[k]}\mathbf{X}_{[k]}^{\mathsf{T}}\right) \times \left(\mathbf{X}_{[k']}\mathbf{X}_{[k']}^{\mathsf{T}}\right)\right\}}{\sqrt{\begin{array}{l}\text{trace}\left\{\left(\mathbf{X}_{[k]}\mathbf{X}_{[k]}^{\mathsf{T}}\right) \times \left(\mathbf{X}_{[k]}\mathbf{X}_{[k]}^{\mathsf{T}}\right)\right\} \\ \times \text{trace}\left\{\left(\mathbf{X}_{[k']}\mathbf{X}_{[k']}^{\mathsf{T}}\right) \times \left(\mathbf{X}_{[k']}\mathbf{X}_{[k']}^{\mathsf{T}}\right)\right\}\end{array}}}. \tag{29}$$

A slightly different coefficient, called the L_g coefficient, is also often used in the context of MFA. This coefficient reflects the MFA normalization and takes positive values. Specifically, the L_g coefficient between data tables k and k' is computed as

$$\begin{aligned} L_{g(k,k')} &= \frac{\text{trace}\left\{\left(\mathbf{X}_{[k]}\mathbf{X}_{[k]}^{\mathsf{T}}\right) \times \left(\mathbf{X}_{[k']}\mathbf{X}_{[k']}^{\mathsf{T}}\right)\right\}}{\gamma_{1,k}^2 \times \gamma_{1,k'}^2} \\ &= \text{trace}\left\{\left(\mathbf{X}_{[k]}\mathbf{X}_{[k]}^{\mathsf{T}}\right) \times \left(\mathbf{X}_{[k']}\mathbf{X}_{[k']}^{\mathsf{T}}\right)\right\} \\ &\quad \times (\alpha_k \times \alpha_{k'}). \end{aligned} \tag{30}$$

An eigen-decomposition of the K by K between-table of R_V or L_g coefficients can provide factor scores for the tables that can be used to plot maps of the tables in an analogous way to the STATIS method (see Ref 3).

ALTERNATIVE PRESENTATION OF MFA

MFA with Cross-Product Matrices

An alternative computational approach to MFA—particularly useful for generalizing to other data types and for comparison to other methods—uses the cross-product matrices $\mathbf{S}_{[k]}$. In this context, the first step is to compute an average cross-product matrix (called a 'compromise' by analogy with the STATIS method, see 3) which is computed as the weighted average of the $\mathbf{S}_{[k]}$ matrices with weights provided by the elements of $\boldsymbol{\alpha}$. Specifically, the compromise cross-product matrix is denoted $\mathbf{S}_{[+]}$ and is computed as

$$\mathbf{S}_{[+]} = \sum_{k}^{K} \alpha_k \mathbf{S}_{[k]}. \tag{31}$$

Note that $\mathbf{S}_{[+]}$ can also be directly computed from $\mathbf{X}$ as

$$\mathbf{S}_{[+]} = \mathbf{X}\mathbf{A}\mathbf{X}^{\mathsf{T}}. \tag{32}$$

The compromise matrix being a weighted sum of cross-product matrices, is also a cross-product matrix (i.e., it is a positive semi-definite matrix) and therefore its eigendecomposition amounts to a PCA. The generalized eigendecomposition under the constraints provided by matrix $\mathbf{M}$ of the compromise gives:

$$\mathbf{S}_{[+]} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{\mathsf{T}} \quad \text{with} \quad \mathbf{P}^{\mathsf{T}}\mathbf{M}\mathbf{P} = \mathbf{I}. \tag{33}$$

Eqs (32) and (13), together indicate that the generalized eigenvectors of the compromise are the left generalized singular vectors of $\mathbf{X}$ (cf., Eq. (13)) and that the eigenvalues of $\mathbf{S}_{[+]}$ are the squares of the singular values of $\mathbf{X}$ (i.e., $\mathbf{\Lambda} = \mathbf{\Delta}^2$). The loadings can be computed by rewriting Eq. (13) as

$$\mathbf{Q} = \mathbf{X}^{\mathsf{T}}\mathbf{M}\mathbf{P}\mathbf{\Delta}^{-1}. \tag{34}$$

Similarly, the compromise factor scores can be computed from $\mathbf{S}_{[+]}$ (cf., Eq. (20)) as

$$\mathbf{F} = \mathbf{S}_{[+]}\mathbf{M}\mathbf{P}\mathbf{\Delta}^{-1}. \tag{35}$$

In this context, the loadings for the variables from table k are obtained from Eq. (34) as

$$\mathbf{Q}_{[k]} = \mathbf{X}_{[k]}^{\mathsf{T}} \mathbf{MP\Delta}^{-1}. \tag{36}$$

The factor scores for table k are obtained from Eqs (36) and (22) as

$$\mathbf{F}_{[k]} = K\alpha_k \mathbf{X}_{[k]}\mathbf{Q}_{[k]} = K\alpha_k \mathbf{X}_{[k]}\mathbf{X}_{[k]}^{\mathsf{T}}\mathbf{MP\Delta}^{-1}$$
$$= K\alpha_k \mathbf{S}_{[k]}\mathbf{MP\Delta}^{-1}. \tag{37}$$

MFA as Simple PCA

MFA can also be computed as the simple PCA of the set of the $\mathbf{X}_{[k]}$ matrices, each weighted by the square root of its respective α weight (this assumes, as it is the case in general for MFA, that matrix $\mathbf{M}$ is equal to $\frac{1}{I}\mathbf{I}$). Specifically, if we define the matrix

$$\widetilde{\mathbf{X}} = \left[\sqrt{\alpha_1}\mathbf{X}_{[1]}|\dots|\sqrt{\alpha_{[k]}}\mathbf{X}_{[k]}|\dots|\sqrt{\alpha_K}\mathbf{X}_{[K]}\right], \tag{38}$$

whose (simple) SVD is given by

$$\widetilde{\mathbf{X}} = \widetilde{\mathbf{P}}\widetilde{\mathbf{\Delta}}\widetilde{\mathbf{Q}}^{\mathsf{T}} \quad \text{with} \quad \widetilde{\mathbf{P}}^{\mathsf{T}}\widetilde{\mathbf{P}} = \widetilde{\mathbf{Q}}^{\mathsf{T}}\widetilde{\mathbf{Q}} = \mathbf{I}. \tag{39}$$

Then the factor scores for the observations can be obtained as

$$\mathbf{F} = \widetilde{\mathbf{P}}\widetilde{\mathbf{\Delta}}. \tag{40}$$

The loadings for the kth table are obtained as

$$\mathbf{Q}_{[k]} = \frac{1}{\sqrt{\alpha}}\widetilde{\mathbf{Q}}_{[k]}. \tag{41}$$

To prove this last identity, we use the relations between the simple SVD and the GSVD (see Refs 74,76,78,79 for more details). To do so, we first need to re-express $\widetilde{\mathbf{X}}$ as a function of $\mathbf{X}$ and $\mathbf{A}$ as:

$$\widetilde{\mathbf{X}} = \left[\sqrt{\alpha_1}\mathbf{X}_{[1]}|\cdots|\sqrt{\alpha_{[k]}}\mathbf{X}_{[k]}|\cdots|\sqrt{\alpha_K}\mathbf{X}_{[K]}\right] = \mathbf{X}\mathbf{A}^{\frac{1}{2}}. \tag{42}$$

To obtain this result we use here the facts (1) that $\mathbf{A}$ is defined as 'blocks' of K values $\alpha_{[k]}$ and (2) that $\mathbf{A}$ being a diagonal matrix with positive elements is positive definite (and therefore its square root is uniquely determined). Rewriting the GSVD of $\mathbf{X}$ in terms of $\widetilde{\mathbf{X}}$ shows that

$$\mathbf{X} = \widetilde{\mathbf{X}}\mathbf{A}^{-\frac{1}{2}} = \widetilde{\mathbf{P}}\widetilde{\mathbf{\Delta}}\widetilde{\mathbf{Q}}^{\mathsf{T}}\mathbf{A}^{-\frac{1}{2}} \tag{43}$$

and therefore that

$$\mathbf{Q} = \widetilde{\mathbf{Q}}\mathbf{A}^{-\frac{1}{2}}, \tag{44}$$

which is equivalent to Eq. (41) (and completes the proof).

MFA from the Table Factor Scores

The whole set factor scores derived from the PCA of one of the tables (i.e., from Eq. (8)) spans the same space as the column space of this table. Therefore, the MFA can be computed directly from these factor scores, and so the MFA factor scores (and loadings) can be obtained from the SVD of the table factor scores (as defined in Eq.(8)) instead than from the SVD of $\mathbf{X}$. This approach can be particularly useful when the number of variables in some tables is significantly larger than the number of observations (a configuration referred to as the '$N << P$ problem'), as can occur with problems such as, for example, brain imaging, genomics, or data mining.

To do so, suffice to replace in Eq. (1) each matrix $\mathbf{X}_{[k]}$ by the corresponding table of factor scores $\mathbf{G}_{[k]}$ and call the new grand matrix $\mathbf{G}$. The GSVD of matrix $\mathbf{G}$ with masses $\mathbf{M}$ and weights $\mathbf{A}$ (note that $\mathbf{A}$ needs to be made conformable with the new dimensions of $\mathbf{G}$) will provide observation factor scores which will be identical to those obtained from the analysis of $\mathbf{X}$. This GSVD will also provide loadings for the table factor scores. The variable loadings can be obtained by projecting these variables as supplementary columns (and will be identical to those obtained from the analysis of $\mathbf{X}$). Conversely, loadings for the table factor scores can be obtained by projecting the table factor scores as supplementary variables onto the analysis obtained from $\mathbf{X}$. All this shows that the analyses obtained from $\mathbf{X}$ or from $\mathbf{G}$ are identical.

The MFA results can also be obtained from the simple SVD of the matrix $\widetilde{\mathbf{G}}$ obtained by treating the $\mathbf{G}_{[k]}$ matrices as defined in Eq. (38) (i.e., by dividing all the elements of each matrix $\mathbf{G}_{[k]}$ by its first singular value).

More on Preprocessing

The preprocessing step is a crucial part of the analysis and can be performed on the columns or on the rows. Most of the time, each variable is centered (i.e., the mean of each variable is 0) and normalized (i.e., the sum of the squared elements of each column is equal to one, I, or even $I - 1$). In some cases, the normalization affects the rows of the matrix and in this case the sum of each row can be equal to 1 (e.g., as in correspondence analysis,[78] see also Ref 96 and 154 for an explicit integration of correspondence analysis and MFA) or the sum of squares of the elements of a given row can be equal to one (e.g., as in Hellinger/Bhattacharyya analysis[97–101]).

SUPPLEMENTARY ELEMENTS (A.K.A. OUT OF SAMPLE)

As in standard PCA, we can use the results of the analysis to compute approximate statistics (e.g., factor scores, loadings, or optimum weights) for *new* elements (i.e., elements that have not been used in the analysis). These new elements are called *supplementary*,[75,78,89–91,102–105] *illustrative*, or *out of sample* elements.[106] In contrast with the supplementary elements, the elements actually used in the analysis are called *active* elements. The statistics for the supplementary elements are obtained by projecting these elements onto the active space. In the MFA framework, we can have supplementary rows and columns (like in PCA) but also supplementary tables. Supplementary rows for which we have values for all J variables and supplementary variables for which we have measurements for all I observations are projected in the same way as for PCA (see Ref 75, pp. 436ff.). Computing statistics for supplementary tables, however, is specific to MFA.

Supplementary Rows and Columns

As MFA is a generalized PCA, we can add supplementary rows and columns as in standard PCA (see Ref 75 for details). Note incidentally, that this procedure assumes that the supplementary rows and columns are scaled in a manner comparable to the rows and columns of the original matrix $\mathbf{X}$. Specifically, from Eqs (22) and (13), we can compute the factor scores, denoted $\mathbf{f}_{sup}$ for a supplementary observation (i.e., a supplementary row of dimensions 1 by J recording measurements on the same variables as the whole matrix $\mathbf{X}$). This supplementary row is represented by a 1 by J vector denoted $\mathbf{r}_{sup}^{\mathsf{T}}$ (which has been preprocessed in the same way as $\mathbf{X}$), the supplementary factor scores are computed as

$$\mathbf{f}_{sup} = \mathbf{r}_{sup}^{\mathsf{T}}\mathbf{A}\mathbf{Q}. \qquad (45)$$

Loadings are denoted $\mathbf{q}_{sup}$ for a new column which is itself denoted by an I by 1 vector $\mathbf{o}_{sup}$ (note that $\mathbf{o}_{sup}$ needs to have been pre-processed in a way comparable with the tables). These loadings are obtained, in a way similar to Eq. (45) as

$$\mathbf{q}_{sup} = \mathbf{o}_{sup}\mathbf{M}\mathbf{P}\mathbf{\Delta}^{-1}. \qquad (46)$$

Supplementary Partial Observations

In some cases, we have supplementary observations for only one (or some) table(s). In this case, called a *supplementary partial observation*, we can obtain the supplementary *partial* factor scores for

this observation from Eq. (22). Specifically, let $\mathbf{x}_{sup[k]}^{\mathsf{T}}$ denote a 1 by $J_{[k]}$ vector of measurements collected on the $J_{[k]}$ variables of table k (note that $\mathbf{x}_{sup[k]}^{\mathsf{T}}$ should have been pre-processed in the same way as the whole matrix $\mathbf{X}_{[k]}$). The partial factor scores for this supplementary observation from table k are obtained as:

$$\mathbf{f}_{sup[k]} = K \times \mathbf{x}_{sup[k]}^{\mathsf{T}} \times \mathbf{Q}_{[k]}. \qquad (47)$$

Incidentally, the factor scores of a supplementary observation collected on all tables can also be obtained as the average of the supplementary partial factor scores (see Eqs (23) and (45)).

To compute the loadings for a supplementary variable for a specific table, it suffices to pre-process this variable like the variables of this table (and this includes dividing all the data of the table by the first singular value of this table) and then to use Eq. (46).

Supplementary Tables

As MFA involves tables, it is of interest to be able to project a whole table as a supplementary element. This table will include new variables measured on the same observations described by the active tables. Such a table is represented by the I by J_{sup} matrix $\mathbf{Y}_{sup}$. The matrix $\mathbf{Y}_{sup}$ is preprocessed in the same manner (e.g., centered, normalized) as the $\mathbf{Y}_{[k]}$ matrices to give the supplementary matrix $\mathbf{X}_{sup}$ which is also normalized by dividing all its elements by its first singular value (i.e., this normalization ensures that the first singular value of $\mathbf{X}_{sup}$ is equal to 1). This matrix will provide supplementary factor scores and loadings for the compromise solution, as described below.

Factor Scores

To obtain the factor scores for a new table, the first step is to obtain (from Eq. (46)) the supplementary loadings which are computed as

$$\mathbf{Q}_{sup} = \mathbf{X}_{sup}^{\mathsf{T}}\mathbf{M}\mathbf{P}\mathbf{\Delta}^{-1}. \qquad (48)$$

Then, using Eq. (22), (and taking into account that it first singular value is equal to one) we obtain the supplementary factor scores for the new table $\mathbf{X}_{sup}$ as

$$\mathbf{F}_{sup} = K\mathbf{X}_{sup}\mathbf{Q}_{sup} = K\mathbf{X}_{sup}\mathbf{X}_{sup}^{\mathsf{T}}\mathbf{M}\mathbf{P}\mathbf{\Delta}^{-1}$$
$$= K\mathbf{S}_{sup}\mathbf{M}\mathbf{P}\mathbf{\Delta}^{-1}. \qquad (49)$$

INFERENTIAL ASPECTS

MFA is a descriptive multivariate technique, but it is often important to be able to complement the descriptive conclusions of an analysis by assessing

if its results are reliable and replicable. For example, a standard question, in the PCA framework, is to find the number of reliable components and most of the approaches used in PCA will work also for MFA. For example, we can use of the informal 'scree' test (a.k.a 'elbow') and the more formal tests RESS, PRESS, and Q^2 statistics (see for details, Refs 75, p. 440ff. and also 107–109). We can also use techniques such as the jackknife[110] or the bootstrap[111,155] to identify important observations, variables, or tables. These approaches are implemented differently if we consider the observations, variables, or even tables as being a random or a fixed factor (see Ref 112).

Bootstrap for the Factor Scores

If we assume that the tables constitute a random factor (i.e., the tables are independent and identically distributed—i.i.d—and sampled from a potentially infinite set of tables) and if we consider the observations as a fixed factor, we may want to estimate the stability of the compromise factor scores. Such an evaluation could be done, for example, by using a jackknife or a bootstrap approach.

We briefly sketch here a possible bootstrap approach for the factor scores (see Ref 113 for the problem of bootstrapping in the context of the SVD, and Refs 114,115 for a review of the bootstrap in the PCA context, and Refs 116–119 for recent applications and developments to MFA). The main idea is to use the properties of Eq. (23) which indicate that the compromise factor scores are the average of the partial factor scores. Therefore, we can obtain bootstrap confidence intervals (CIs) by repeatedly sampling with replacement from the set of tables and compute new compromise factor scores (this approach corresponds to the *partial bootstrap* of Ref 114, see Ref 59 for an alternative approach using split-half resampling). From these estimates we can also compute *bootstrap ratios* for each dimension by dividing the mean of the bootstrap estimates by their standard deviation. These bootstrap ratios are akin to *t* statistics and can be used to detect observations that reliably contribute to a given dimension. So, for example, for a given dimension and a given observation a value of the bootstrap ratio larger than 2 will be considered reliable (by analogy with a *t* larger than 2 which would be 'significant' at $p < .05$). When evaluating bootstrap ratios, the multiple comparisons problem can be taken into account by using, for example, a Bonferroni-type correction (see Ref 120) and instead of using critical values corresponding to say $p < .05$ we would use values corresponding to $p < \frac{.05}{I}$.

More formally, in order to compute a bootstrap estimate, we need to generate a bootstrap sample.

To do so, we first take a sample of integers with replacement from the set of integers from 1 to K. Recall that, when sampling with replacement, any element from the set can be sampled zero, one, or more than one times. We call this set $\mathcal{B}$ (for *bootstrap*). For example with five elements, a possible bootstrap set could be $\mathcal{B} = \{1, 5, 1, 3, 3\}$. We then generate a new data set (i.e., a new $\mathbf{X}$ matrix comprising K tables) using matrices $\mathbf{X}_{[k]}$ with these indices. So with $K = 5$, this would give the following bootstrap set

$$\{\mathbf{X}_{[1]}, \mathbf{X}_{[5]}, \mathbf{X}_{[1]}, \mathbf{X}_{[3]}, \mathbf{X}_{[3]}\}. \tag{50}$$

From this set we would build a data matrix denoted $\mathbf{X}_1^*$ that would then be analyzed by MFA. This analysis would provide a set of *bootstrapped factor scores* (denoted $\mathbf{F}_1^*$) obtained by projecting the bootstrapped data table as a supplementary element (see Eqs (37) and (49)). Interestingly, as a consequence of the barycentric properties of the factor scores (see Eq. (23)), this last step can also be directly obtained by computing $\mathbf{F}_1^*$ as the weighted average of the corresponding partial factor scores. We then repeat the procedure a large number of times (e.g., $L = 1000$) and generate L bootstrapped matrices of factor scores $\mathbf{F}_\ell^*$. From these bootstrapped matrices of factor scores, we can derive CIs and estimate the mean factor scores as the mean of the bootstrapped factor scores. Formally, $\overline{\mathbf{F}}^*$ denotes the bootstrap estimated factor scores, and is computed as

$$\overline{\mathbf{F}}^* = \frac{1}{L} \sum_\ell^L \mathbf{F}_\ell^*. \tag{51}$$

In a similar way, the bootstrapped estimate of the variance is obtained as the variance of the $\mathbf{F}_\ell^*$ matrices. Formally, $\widehat{\sigma}_{\mathbf{F}^*}^2$ denotes the bootstrapped estimate of the variance and is computed as

$$\widehat{\sigma}_{\mathbf{F}^*}^2 = \frac{1}{L} \left(\sum_\ell^L \left(\mathbf{F}_\ell^* - \overline{\mathbf{F}}^* \right) \circ \left(\mathbf{F}_\ell^* - \overline{\mathbf{F}}^* \right) \right). \tag{52}$$

Bootstrap ratios (denoted $\mathbf{T}^*$) are computed by dividing the bootstrapped means by the corresponding bootstrapped standard deviations (denoted $\widehat{\sigma}_{\mathbf{F}^*}$ and is the square root of the bootstrapped estimate of the variance). These bootstrap ratios are often interpreted as Student's *t* statistics.

Bootstrapped Confidence Intervals

The bootstrap factor scores (i.e., the $\mathbf{F}_\ell^*$'s) can also be used to compute CIs for the observations. For a given dimension, the bootstrapped factor scores of an

observation can be ordered from the smallest to the largest and a CI for a given p value can be obtained by trimming the upper and lower $\frac{1}{p}$ proportion of the distribution. In general, for a given dimension, the bootstrap ratios, and the CIs will agree in detecting the relevant observations. In addition, CIs can also be plotted directly on the factor scores map as confidence ellipsoids or confidence convex-hulls which comprise a $1 - p$ proportion of the bootstrapped factor scores (see Ref 59 and our example illustration in Figure 8). When the ellipsoids or convex-hulls of two observations do not overlap, these two observations can be considered as reliably different. Like for the bootstrap ratios, in order to correct for the potential problem of multiple comparisons, a Bonferroni type of correction can also be implemented when plotting hulls or ellipsoids (see Ref 59 for details). Recent work[118] suggests that these CIs could be too liberal (i.e., too small) when the number of tables is large and that a better procedure would be to use instead a 'total bootstrap' (i.e., recomputing the factors scores from the whole bootstrap tables rather than from the partial factor scores cf., Ref 114).

AN EXAMPLE

To illustrate MFA, we selected a (fictitious) example previously used (see Ref 3) to illustrate the STATIS method which is a possible alternative to MFA. In this example, a set of wines was described by a group of expert tasters called assessors. This type of data could be analyzed using a standard PCA, but this approach obviously neglects the inter-assessor differences. MFA has the advantages of providing a common space for the products (i.e., the factor scores), as well as information about how each assessor relates to this common space (i.e., the partial factor scores).

Specifically, this example concerns 12 wines made from Sauvignon Blanc grapes coming from three wine regions (four wines from each region): New Zealand, France, and Canada. Ten expert assessors were asked to evaluate these wines. The assessors were asked (1) to evaluate the wines on 9-point rating scales, using four variables considered as standard for the evaluation of these wines (cat-pee, passion-fruit, green pepper, and mineral) and, (2) if they felt the need, to add some variables of their own (some assessors choose none, some choose one or two more variables). The raw data are presented in Table 1. The goals of the analysis were twofold:

1. to obtain a typology of the wines, and
2. to discover agreement (if any) between the assessors.

For the example, the data consist of $K = 10$ tables (one for each assessor) shown in Table 1. For example, the first table denoted $\mathbf{Y}_{[1]}$ is equal to

$$\mathbf{Y}_{[1]} = \begin{bmatrix} 8 & 6 & 7 & 4 & 1 & 6 \\ 7 & 5 & 8 & 1 & 2 & 8 \\ 6 & 5 & 6 & 5 & 3 & 4 \\ 9 & 6 & 8 & 4 & 3 & 5 \\ 2 & 2 & 2 & 8 & 7 & 3 \\ 3 & 4 & 4 & 9 & 6 & 1 \\ 5 & 3 & 5 & 4 & 8 & 3 \\ 5 & 2 & 4 & 8 & 7 & 4 \\ 8 & 6 & 8 & 4 & 4 & 7 \\ 4 & 6 & 2 & 5 & 3 & 4 \\ 8 & 4 & 8 & 1 & 3 & 3 \\ 5 & 3 & 6 & 4 & 4 & 2 \end{bmatrix}. \quad (53)$$

Each table was then preprocessed by first centering and normalizing each column such that its mean is equal to 0 and the sum of the square values of all its elements is equal to 1. For example, $\mathbf{X}_{[1]}$, the pre-processed matrix for Assessor 1 is equal to:

$$\mathbf{X}_{[1]} = \begin{bmatrix} 0.30 & 0.32 & 0.18 & -0.09 & -0.44 & 0.27 \\ 0.16 & 0.13 & 0.31 & -0.45 & -0.31 & 0.57 \\ 0.02 & 0.13 & 0.04 & 0.03 & -0.17 & -0.02 \\ 0.43 & 0.32 & 0.31 & -0.09 & -0.17 & 0.12 \\ -0.52 & -0.45 & -0.49 & 0.39 & 0.37 & -0.17 \\ -0.39 & -0.06 & -0.22 & 0.51 & 0.24 & -0.47 \\ -0.11 & -0.26 & -0.09 & -0.09 & 0.51 & -0.17 \\ -0.11 & -0.45 & -0.22 & 0.39 & 0.37 & -0.02 \\ 0.30 & 0.32 & 0.31 & -0.09 & -0.03 & 0.42 \\ -0.25 & 0.32 & -0.49 & 0.03 & -0.17 & -0.02 \\ 0.30 & -0.06 & 0.31 & -0.45 & -0.17 & -0.17 \\ -0.11 & -0.26 & 0.04 & -0.09 & -0.03 & -0.32 \end{bmatrix}. \quad (54)$$

PCA of the Data Tables

A PCA of each of the data table will then be performed. For example, the SVD of the first data table gives

$$\mathbf{X}_{[1]} = \mathbf{U}_{[1]}\mathbf{\Gamma}_{[1]}\mathbf{V}_{[1]}^{\mathsf{T}} \text{ with } \mathbf{U}_{[1]}^{\mathsf{T}}\mathbf{U}_{[1]} = \mathbf{V}_{[1]}^{\mathsf{T}}\mathbf{V}_{[1]} = \mathbf{I}, \quad (55)$$

with

$$\mathbf{U}_{[1]} = \begin{bmatrix} 0.32 & 0.26 & -0.03 & 0.09 & 0.32 & -0.22 \\ 0.38 & -0.06 & 0.35 & -0.49 & 0.17 & 0.41 \\ 0.06 & 0.14 & -0.15 & 0.07 & 0.14 & 0.10 \\ 0.30 & 0.00 & -0.06 & 0.45 & -0.13 & -0.23 \\ -0.49 & 0.06 & 0.24 & -0.23 & 0.19 & 0.01 \\ -0.38 & 0.20 & -0.28 & 0.41 & 0.00 & 0.52 \\ -0.21 & -0.40 & 0.12 & -0.08 & -0.65 & -0.01 \\ -0.31 & -0.16 & 0.42 & 0.19 & 0.32 & -0.46 \\ 0.29 & 0.05 & 0.39 & 0.29 & -0.26 & 0.25 \\ -0.07 & 0.63 & -0.18 & -0.39 & -0.37 & -0.30 \\ 0.21 & -0.45 & -0.44 & -0.14 & -0.01 & -0.23 \\ -0.10 & -0.27 & -0.39 & -0.17 & 0.27 & 0.16 \end{bmatrix}, \quad (56)$$

TABLE 1 | Raw Data (Tables $Y_{[1]}$ Through $Y_{[10]}$)

	Assessor 1						Assessor 2						Assessor 3						Assessor 4					Assessor 5					
	V1	V2	V3	V4	V5	V6	V1	V2	V3	V4	V7	V8	V1	V2	V3	V4	V9	V10	V1	V2	V3	V4	V8	V1	V2	V3	V4	V11	V12
NZ$_1$	8	6	7	4	1	6	8	6	8	3	7	5	8	6	8	3	7	2	9	5	8	2	6	9	6	9	3	8	2
NZ$_2$	7	5	8	1	2	8	6	5	6	3	7	7	8	7	7	2	8	2	8	7	7	3	5	7	7	7	1	9	2
NZ$_3$	6	5	6	5	3	4	6	6	6	5	8	7	8	7	7	6	9	1	8	8	9	2	7	7	7	7	1	7	2
NZ$_4$	9	6	8	4	3	5	8	6	8	4	6	6	8	2	8	3	9	3	8	8	9	4	7	8	9	7	5	6	1
FR$_1$	2	2	2	8	7	3	2	3	1	7	4	3	3	4	3	6	4	6	4	2	2	4	3	4	4	4	2	4	4
FR$_2$	3	4	4	9	6	1	4	3	4	9	3	5	4	3	4	8	3	9	3	2	2	6	2	4	5	5	6	1	5
FR$_3$	5	3	5	4	8	3	3	3	2	7	4	4	5	4	5	2	3	6	4	4	4	6	4	6	5	7	2	3	1
FR$_4$	5	2	4	8	7	4	4	3	5	5	3	3	6	3	7	7	1	7	5	2	2	9	4	6	6	5	8	4	5
CA$_1$	8	6	8	4	4	7	8	6	9	5	5	6	8	5	9	1	5	2	7	5	6	3	2	8	6	8	2	5	4
CA$_2$	4	6	2	5	3	4	5	5	5	6	5	8	5	5	4	6	5	1	5	6	6	4	4	6	6	6	4	6	3
CA$_3$	8	4	8	1	3	3	8	4	8	3	7	7	8	3	7	3	5	4	7	3	6	1	6	7	4	8	4	5	1
CA$_4$	5	3	6	4	4	2	5	3	7	4	8	5	5	4	4	5	4	3	5	2	2	6	6	5	5	5	6	1	

	Assessor 6					Assessor 7				Assessor 8						Assessor 9					Assessor 10			
	V1	V2	V3	V4	V13	V1	V2	V3	V4	V1	V2	V3	V4	V14	V5	V1	V2	V3	V4	V15	V1	V2	V3	V4
NZ$_1$	8	5	6	2	9	8	5	8	4	7	6	7	4	9	2	8	6	9	1	7	8	6	7	5
NZ$_2$	6	6	6	2	4	7	6	8	4	6	5	6	2	7	2	8	7	9	1	6	7	5	7	3
NZ$_3$	7	7	7	2	7	6	7	6	3	6	6	6	4	9	2	7	7	8	4	7	7	6	6	2
NZ$_4$	8	7	8	2	8	7	8	6	1	8	7	8	2	8	2	8	9	9	3	9	8	7	7	4
FR$_1$	3	2	2	7	2	4	2	3	6	3	3	4	4	4	4	3	4	4	5	4	2	3	1	7
FR$_2$	3	3	3	3	4	4	4	4	4	4	4	4	7	3	6	5	5	5	7	2	3	3	3	9
FR$_3$	4	2	3	3	3	4	3	4	4	5	3	5	3	3	5	5	5	5	6	3	4	2	5	8
FR$_4$	5	3	5	9	3	5	3	5	7	6	4	6	3	2	4	5	5	6	5	3	3	4	2	8
CA$_1$	7	7	7	1	4	8	4	9	4	8	6	5	4	5	4	8	7	8	4	7	8	6	7	4
CA$_2$	4	6	2	4	6	4	7	5	2	5	7	5	4	6	1	5	6	4	5	6	5	6	4	4
CA$_3$	7	4	8	2	3	8	5	7	3	7	4	8	2	6	2	8	4	7	4	5	7	4	8	5
CA$_4$	4	5	3	3	7	4	3	5	2	5	4	6	2	4	3	5	4	5	3	4	5	4	6	6

V1, Cat Pee; V2, Passion Fruit; V3, Green Pepper; V4, Mineral; V5, Smoky; V6, Citrus; V7, Tropical; V8, Leafy; V9, Grassy; V10, Flinty; V11, Vegetal; V12, Hay; V13, Melon; V14, Grass; V15, Peach.

$$\Gamma_{[1]} = \mathrm{diag} \left\{ \begin{bmatrix} 2.037 \\ 0.886 \\ 0.701 \\ 0.589 \\ 0.402 \\ 0.255 \end{bmatrix} \right\}, \quad (57)$$

and

$$V_{[1]} = \begin{bmatrix} 0.45 & -0.25 & 0.02 & 0.44 & -0.03 & -0.73 \\ 0.38 & 0.62 & -0.17 & 0.26 & -0.58 & 0.20 \\ 0.42 & -0.47 & -0.04 & 0.39 & 0.19 & 0.65 \\ -0.40 & 0.39 & 0.20 & 0.69 & 0.41 & 0.02 \\ -0.41 & -0.38 & 0.41 & 0.25 & -0.67 & 0.06 \\ 0.37 & 0.19 & 0.87 & -0.22 & 0.12 & 0.05 \end{bmatrix}. \quad (58)$$

From this PCA we obtain the following factor scores for the first table:

$$G_{[1]} = U_{[1]}\Gamma_{[1]}$$

$$= \begin{bmatrix} 0.65 & 0.23 & -0.02 & 0.05 & 0.13 & -0.06 \\ 0.77 & -0.05 & 0.25 & -0.29 & 0.07 & 0.11 \\ 0.13 & 0.13 & -0.11 & 0.04 & 0.06 & 0.03 \\ 0.60 & 0.00 & -0.04 & 0.26 & -0.05 & -0.06 \\ -0.99 & 0.05 & 0.17 & -0.14 & 0.08 & 0.00 \\ -0.77 & 0.18 & -0.19 & 0.24 & 0.00 & 0.13 \\ -0.43 & -0.35 & 0.09 & -0.05 & -0.26 & -0.00 \\ -0.64 & -0.15 & 0.29 & 0.11 & 0.13 & -0.12 \\ 0.59 & 0.04 & 0.27 & 0.17 & -0.10 & 0.06 \\ -0.15 & 0.56 & -0.12 & -0.23 & -0.15 & -0.08 \\ 0.43 & -0.40 & -0.31 & -0.08 & -0.00 & -0.06 \\ -0.20 & -0.24 & -0.27 & -0.10 & 0.11 & 0.04 \end{bmatrix} \quad (59)$$

From Eq. (57), we have found that the first singular value of $\mathbf{X}_{[1]}$ had a value of 2.037, and therefore the α weight for the first table is obtained as (cf., Eq. (10)):

$$\alpha_1 = \frac{1}{\gamma_{1,1}^2} = \gamma_{1,1}^{-2} = \frac{1}{2.037^2} = 0.241. \quad (60)$$

Creating the α Weight Vector

We collect the value of the α weights of the K tables into a K by one weight vector denoted by $\boldsymbol{\alpha}$

$$\boldsymbol{\alpha} = \begin{bmatrix} 0.241 \\ 0.239 \\ 0.275 \\ 0.273 \\ 0.307 \\ 0.302 \\ 0.417 \\ 0.272 \\ 0.264 \\ 0.309 \end{bmatrix}. \quad (61)$$

There are $K = 10$ values in $\boldsymbol{\alpha}$ (see Eqs (2) and (61)).

These values are stored in the $J = 53 \times 1$ vector $\mathbf{a}$, which can itself be used to fill in the diagonal elements of the 53×53 diagonal matrix $\mathbf{A}$:

$$\mathbf{a} = \begin{bmatrix} \mathbf{1}_{[1]} \times 0.241 \\ \mathbf{1}_{[2]} \times 0.239 \\ \mathbf{1}_{[3]} \times 0.275 \\ \mathbf{1}_{[4]} \times 0.273 \\ \mathbf{1}_{[5]} \times 0.307 \\ \mathbf{1}_{[6]} \times 0.302 \\ \mathbf{1}_{[7]} \times 0.417 \\ \mathbf{1}_{[8]} \times 0.272 \\ \mathbf{1}_{[9]} \times 0.264 \\ \mathbf{1}_{[10]} \times 0.309 \end{bmatrix} \quad \text{and}$$

$$\mathbf{A} = \text{diag}\{\mathbf{a}\} = \text{diag}\left\{ \begin{bmatrix} \mathbf{1}_{[1]} \times 0.241 \\ \mathbf{1}_{[2]} \times 0.239 \\ \mathbf{1}_{[3]} \times 0.275 \\ \mathbf{1}_{[4]} \times 0.273 \\ \mathbf{1}_{[5]} \times 0.307 \\ \mathbf{1}_{[6]} \times 0.302 \\ \mathbf{1}_{[7]} \times 0.417 \\ \mathbf{1}_{[8]} \times 0.272 \\ \mathbf{1}_{[9]} \times 0.264 \\ \mathbf{1}_{[10]} \times 0.309 \end{bmatrix} \right\}, \quad (62)$$

where $\mathbf{1}_{[k]}$ is an $J_{[k]}$ by 1 vector of ones.

Generalized PCA of X

The 12×12 diagonal mass matrix $\mathbf{M}$ for the observations (with equal masses set to $m_i = \frac{1}{I} = .08$) is given as

$$\mathbf{M} = \text{diag}\{\mathbf{m}\}$$

$$= \text{diag}\left\{ \frac{1}{12} \times \mathbf{1} \right\} = \text{diag}\left\{ \begin{bmatrix} 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \\ 0.08 \end{bmatrix} \right\}. \quad (63)$$

The GSVD of $\mathbf{X}$ with matrices $\mathbf{M}$ (Eq. (63)) and $\mathbf{A}$ (Eq. (62)) gives $\mathbf{X} = \mathbf{P\Delta Q}^\mathsf{T}$. The eigenvalues (denoted λ) are equal to the squares of the singular values and are often used to informally decide upon the number of components to keep for further inspection. The eigenvalues and the percentage of inertia that they explain are given in Table 2

The pattern of the eigenvalue distribution suggests to keep two or three dimensions for future examination and we decided (somewhat arbitrarily) to keep only the first two dimensions for this example.

The matrix $\mathbf{Q}$ of the right singular vectors (loadings for the variables) is given in Table 3, the matrix $\mathbf{P}$ of the right singular vectors and the matrix $\mathbf{\Delta}$ of the singular values are given below:

$$\mathbf{P} = \begin{bmatrix} -1.117 & 0.466 \\ -0.922 & 0.093 \\ -0.867 & -1.295 \\ -1.270 & -0.473 \\ 1.564 & -0.366 \\ 1.440 & -0.308 \\ 0.921 & 0.584 \\ 1.054 & 1.163 \\ -0.762 & 1.051 \\ 0.083 & -2.158 \\ -0.542 & 1.463 \\ 0.418 & -0.217 \end{bmatrix} \quad \mathbf{\Delta} = \text{diag}\left\{ \begin{bmatrix} 0.878 \\ 0.351 \end{bmatrix} \right\}. \quad (64)$$

Factor Scores

The factor scores for $\mathbf{X}$ show the best two-dimensional representation of the compromise of the K tables. Using Eqs (18), (20), and (64), we obtain:

TABLE 2 | Eigenvalues and Percentage of Explained Inertia of the MFA of **X**

	Component										
	1	2	3	4	5	6	7	8	9	10	11
Singular value (δ)	0.878	0.351	0.301	0.276	0.244	0.198	0.176	0.158	0.137	0.116	0.106
Eigenvalue ($\lambda = \delta^2$)	0.770	0.123	0.091	0.076	0.060	0.039	0.031	0.025	0.019	0.013	0.011
cumulative	0.770	0.893	0.984	1.060	1.120	1.159	1.190	1.215	1.233	1.247	1.258
% Inertia (τ)	61	10	7	6	5	3	2	2	1	1	1
cumulative	61	71	78	84	89	92	94	96	97	98	100

$$\mathbf{F} = \mathbf{P\Delta} = \mathbf{XAQ} = \begin{bmatrix} -0.980 & 0.163 \\ -0.809 & 0.033 \\ -0.761 & -0.454 \\ -1.115 & -0.166 \\ 1.373 & -0.128 \\ 1.264 & -0.108 \\ 0.808 & 0.205 \\ 0.925 & 0.408 \\ -0.669 & 0.369 \\ 0.073 & -0.757 \\ -0.476 & 0.513 \\ 0.367 & -0.076 \end{bmatrix}. \quad (65)$$

In the **F** matrix, each row represents a wine and each column is a component. Figure 2a shows the wines in the space created by the first two components. The first component (with an eigenvalue equal to $\lambda_1 = 0.878^2 = 0.770$) explains 61% of the inertia, and contrasts the French and New Zealand wines. The second component (with an eigenvalue of $\lambda_2 = 0.351^2 = 0.123$) explains 10% of the inertia and is more delicate to interpret from the wines alone (its interpretation will becomes clearer after looking at the loadings).

Partial Factor Scores

The partial factor scores (which are the projections of the tables onto the compromise) are computed from Eq. (22). For example, for the first assessor, the partial factor scores of the 12 wines are obtained as:

$$\mathbf{F}_{[1]} = K\alpha_1 \mathbf{X}_{[1]} \mathbf{Q}_{[1]}$$

$$= 10 \times 0.241 \begin{bmatrix} -1.037 & 0.155 \\ -1.179 & 0.596 \\ -0.213 & -0.104 \\ -0.946 & 0.446 \\ 1.546 & -0.676 \\ 1.176 & -0.747 \\ 0.698 & 0.166 \\ 1.006 & -0.063 \\ -0.922 & 0.486 \\ 0.189 & -0.936 \\ -0.643 & 0.640 \\ 0.323 & 0.036 \end{bmatrix} \quad (66)$$

and are displayed in Figure 3 (i.e., first top left panel, see also Figure 2). From Figure 3, we can see that, for all the assessors, Component 1 separates the New Zealand from the French Sauvignon Blancs, a configuration replicating the pattern seen in the compromise (Figure 2a). However, the assessors show a large inter-individual differences in how they rated the Canadian wines.

The original variables are analyzed, as in standard PCA, by computing loadings which are given in Table 3 for the first two dimensions. The loadings are also plotted in a biplot fashion in Figure 3. Here we show the partial factor scores (of the wines) along with the loadings for each assessor (which we have re-scaled so that their variance, for each dimension, is equal to the singular value of the compromise). From these plots, we can see that for all the assessors, the New Zealand wines are rated as having a more cat-pee aroma, with some green pepper and passion fruit, while the French wines are rated as being more mineral, smoky, or hay-like.

Determining the Importance of the Tables in the Compromise

There are two ways to determine which tables play the largest role in the compromise: contributions and partial inertias. The contribution of a table reflects the proportion of the variance of a dimension that can be attributed to this table (see Eq. (28)). The larger the contribution of a table to a component, the more important this table is for this component. The contributions for the tables are as follows:

$$\text{ctr}_{k,\ell} = \begin{bmatrix} 0.101 & 0.095 \\ 0.100 & 0.068 \\ 0.101 & 0.152 \\ 0.096 & 0.049 \\ 0.098 & 0.063 \\ 0.101 & 0.104 \\ 0.102 & 0.224 \\ 0.096 & 0.134 \\ 0.100 & 0.053 \\ 0.105 & 0.057 \end{bmatrix}. \quad (67)$$

TABLE 3 | α Weights, Loadings, Squared Loadings, and Contributions for the MFA of **X**

	Assessor 1						Assessor 2						Assessor 3					
	V1	V2	V3	V4	V5	V6	V1	V2	V3	V4	V7	V8	V1	V2	V3	V4	V9	V10
α-Weights	0.241	0.241	0.241	0.241	0.241	0.241	0.239	0.239	0.239	0.239	0.239	0.239	0.275	0.275	0.275	0.275	0.275	0.275
Loadings (Q)																		
Dim 1	−0.294	−0.267	−0.260	0.241	0.286	−0.233	−0.297	−0.296	−0.267	0.256	−0.238	−0.222	−0.305	−0.136	−0.258	0.203	−0.277	0.267
Dim 2	0.318	−0.248	0.396	−0.184	0.161	0.129	0.183	−0.178	0.200	−0.240	−0.113	−0.333	0.234	−0.228	0.379	−0.365	−0.297	0.283
Squared loadings																		
Dim 1	0.087	0.071	0.068	0.058	0.082	0.054	0.088	0.088	0.071	0.066	0.057	0.049	0.093	0.018	0.067	0.041	0.077	0.071
Dim 2	0.101	0.062	0.157	0.034	0.026	0.017	0.034	0.032	0.040	0.058	0.013	0.111	0.055	0.052	0.144	0.133	0.088	0.080
Contribution×1000																		
Dim 1	21	17	16	14	20	13	21	21	17	16	14	12	26	5	18	11	21	20
Dim 2	24	15	38	8	6	4	8	8	10	14	3	27	15	14	39	37	24	22

	Assessor 4					Assessor 5						Assessor 6					Assessor 7			
	V1	V2	V3	V4	V8	V1	V2	V3	V4	V11	V12	V1	V2	V3	V4	V13	V1	V2	V3	V4
α-Weights	0.273	0.273	0.273	0.273	0.273	0.307	0.307	0.307	0.307	0.307	0.307	0.302	0.302	0.302	0.302	0.302	0.417	0.417	0.417	0.417
Loadings (Q)																				
Dim 1	−0.313	−0.261	−0.303	0.230	−0.205	−0.296	−0.213	−0.268	0.124	−0.259	.177	−0.302	−0.277	−0.265	0.231	−0.205	−0.275	−0.246	−0.277	0.180
Dim 2	0.082	−0.353	−0.169	0.066	−0.117	0.201	−0.249	0.258	0.132	−0.144	.019	0.215	−0.274	0.328	0.031	−0.340	0.380	−0.410	0.290	0.376
Squared Loadings																				
Dim 1	0.098	0.068	0.092	0.053	0.042	0.088	0.045	0.072	0.015	0.067	0.015	0.091	0.077	0.070	0.053	0.042	0.076	0.061	0.076	0.032
Dim 2	0.007	0.125	0.029	0.004	0.014	0.040	0.062	0.067	0.017	0.021	0.017	0.046	0.075	0.107	0.001	0.116	0.144	0.168	0.084	0.141
Contribution×1000																				
Dim 1	27	19	25	14	12	27	14	22	5	20	10	28	23	21	16	13	31	25	32	14
Dim 2	2	34	8	1	4	12	19	20	5	6	0	14	23	32	0	35	60	70	35	59

TABLE 3 | Continued

	Assessor 8						Assessor 9					Assessor 10			
	V1	V2	V3	V4	V14	V5	V1	V2	V3	V4	V15	V1	V2	V3	V4
α-Weights	0.272	0.272	0.272	0.272	0.272	0.272	0.264	0.264	0.264	0.264	0.264	0.309	0.309	0.309	0.309
Loadings (Q)															
Dim 1	−0.276	−0.247	−0.235	0.138	−0.286	.239	−0.303	−0.235	−0.287	0.251	−0.296	−0.323	−0.274	−0.286	0.282
Dim 2	0.309	−0.376	0.231	−0.219	−0.261	0.293	0.241	−0.221	0.226	−0.083	−0.188	0.080	−0.262	0.187	0.272
Squared Loadings															
Dim 1	0.076	0.061	0.055	0.019	0.082	0.057	0.092	0.055	0.082	0.063	0.088	0.104	0.075	0.082	0.080
Dim 2	0.096	0.142	0.054	0.048	0.068	0.086	0.058	0.049	0.051	0.007	0.035	0.006	0.069	0.035	0.074
Contribution ×1000															
Dim 1	21	17	15	5	22	16	24	14	22	17	23	32	23	25	25
Dim 2	26	39	15	13	19	23	15	13	13	2	9	2	21	11	23

V1, Cat Pee; V2, Passion Fruit; V3, Green Pepper; V4, Mineral; V5, Smoky; V6, Citrus; V7, Tropical; V8, Leafy; V9, Grassy; V10, Flinty; V11, Vegetal; V12, Hay; V13, Melon; V14, Grass; V15, Peach.

The contributions can also be plotted to obtain a visual representation of the importance of the studies. Figure 4 shows the relative contributions of each of the tables to Components 1 and 2. From this figure, we can see that Assessor 10 contributes the most to the first component of the compromise, while Assessor 7 contributes most to the second component. Assessor 5, by contrast, contributes the least to both Components 1 and 2.

In addition to using contributions, we can determine a table's importance with partial inertia that gives the proportion of the compromise variance (i.e., inertia) explained by the table. This is obtained by multiplying, for each dimension, the table contribution by the eigenvalue of the dimension. For our example, the partial inertias denoted $\mathcal{I}_{partial}$ are as follows:

$$\mathcal{I}_{partial} = \begin{bmatrix} 0.0779 & 0.0117 \\ 0.0771 & 0.0084 \\ 0.0778 & 0.0186 \\ 0.0743 & 0.0060 \\ 0.0751 & 0.0078 \\ 0.0776 & 0.0128 \\ 0.0787 & 0.0275 \\ 0.0736 & 0.0165 \\ 0.0771 & 0.0065 \\ 0.0810 & 0.0070 \end{bmatrix}. \quad (68)$$

Like the contributions, the partial inertia can be plotted to get a visual representation. From Figure 5, we see that Assessor 10 accounts for the most inertia on the first dimension, while Assessor 5 accounts for the lowest proportion of the total variance.

Supplementary Table

As we would like to know what qualities of the wines are associated with the assessors' ratings, we wanted to include some chemical components of the wines as variables, namely, titratable acidity, pH, alcohol, and residual sugar. The values for these variables are shown in Table 4. However, because these properties are qualitatively different from the assessors' ratings (i.e., did not come from the same population), we did not want to include them as active elements in the analysis and therefore projected them as a supplementary table.

To obtain the factor scores, the first step is to center and normalize each variable of the supplementary data table. Then after computing its first singular value (equal to 1.3867), each element of the centered and normalized table is divided by the
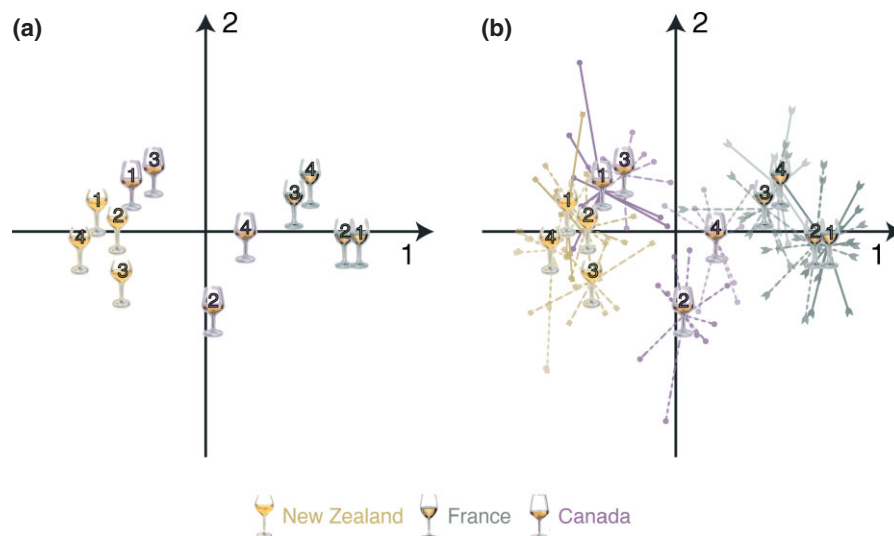
FIGURE 2 | Compromise of the 10 tables. (a) Factor scores (wines). (b) Assessors' partial factor scores projected into the compromise as supplementary elements. Each assessor is represented by a dot, and for each wine a line connects the wine factor scores to the partial factors scores of a given assessor for this wine. ($\lambda_1 = 0.770$, $\tau_1 = 61\%$; $\lambda_2 = 0.123$, $\tau_2 = 10\%$).

first singular value. This gives the matrix $\mathbf{X}_{\text{sup}}$ (whose first singular value is now equal to 1):

$$
\mathbf{X}_{\text{sup}} = \begin{bmatrix}
-0.094 & 0.081 & 0.315 & 0.139 \\
-0.152 & 0.171 & 0.143 & 0.288 \\
0.023 & 0.015 & 0.315 & 0.139 \\
0.470 & -0.032 & 0.143 & 0.362 \\
-0.210 & 0.213 & -0.200 & -0.234 \\
-0.039 & -0.146 & -0.200 & -0.110 \\
-0.307 & 0.051 & -0.029 & -0.408 \\
-0.094 & 0.093 & -0.372 & -0.085 \\
0.295 & 0.033 & -0.029 & 0.089 \\
-0.074 & 0.111 & 0.143 & -0.085 \\
0.023 & 0.033 & -0.200 & 0.014 \\
0.159 & -0.625 & -0.029 & -0.110
\end{bmatrix}.
\tag{69}
$$

We then compute the supplementary loadings, which are obtained as (cf., Eq. (48)):

$$
\mathbf{Q}_{\text{sup}} = \mathbf{X}_{\text{sup}}^{\mathsf{T}}\mathbf{MP\Delta}^{-1} = \begin{bmatrix}
-0.125 & -0.009 \\
-0.024 & 0.032 \\
-0.173 & -0.298 \\
-0.201 & -0.037
\end{bmatrix}.
\tag{70}
$$

From these loadings, we can see that, on Component 1, the New Zealand Sauvignon Blancs are more acidic, have a greater alcohol content, and more residual sugar than the French wines.

Next, we compute the supplementary factor scores for the first 2 components as (cf., Eq. 49):

$$
\mathbf{F}_{\text{sup}} = K\mathbf{X}_{\text{sup}}\mathbf{Q}_{\text{sup}} = K\mathbf{X}_{\text{sup}}\mathbf{X}_{\text{sup}}^{\mathsf{T}}\mathbf{MP\Delta}^{-1}
$$

$$
= \begin{bmatrix}
-0.727 & -0.954 \\
-0.677 & -0.463 \\
-0.857 & -0.986 \\
-1.556 & -0.615 \\
1.030 & 0.771 \\
0.651 & 0.594 \\
1.241 & 0.281 \\
0.910 & 1.178 \\
-0.506 & 0.035 \\
-0.011 & -0.353 \\
0.281 & 0.600 \\
0.219 & -0.089
\end{bmatrix}.
\tag{71}
$$

In a manner similar to the biplot approach we used for the assessors (see Figure 3) we plotted together, in a biplot way, the supplementary partial factor scores and loadings for the chemical properties of the wines (see Figure 6). This figure confirms the interpretation that we reached from the numerical values.

Bootstrap

To estimate the stability of the compromise factor scores, we used a bootstrap approach. We generated 1000 bootstrap samples that gave 1000 estimated bootstrapped factor scores. For example, for the first
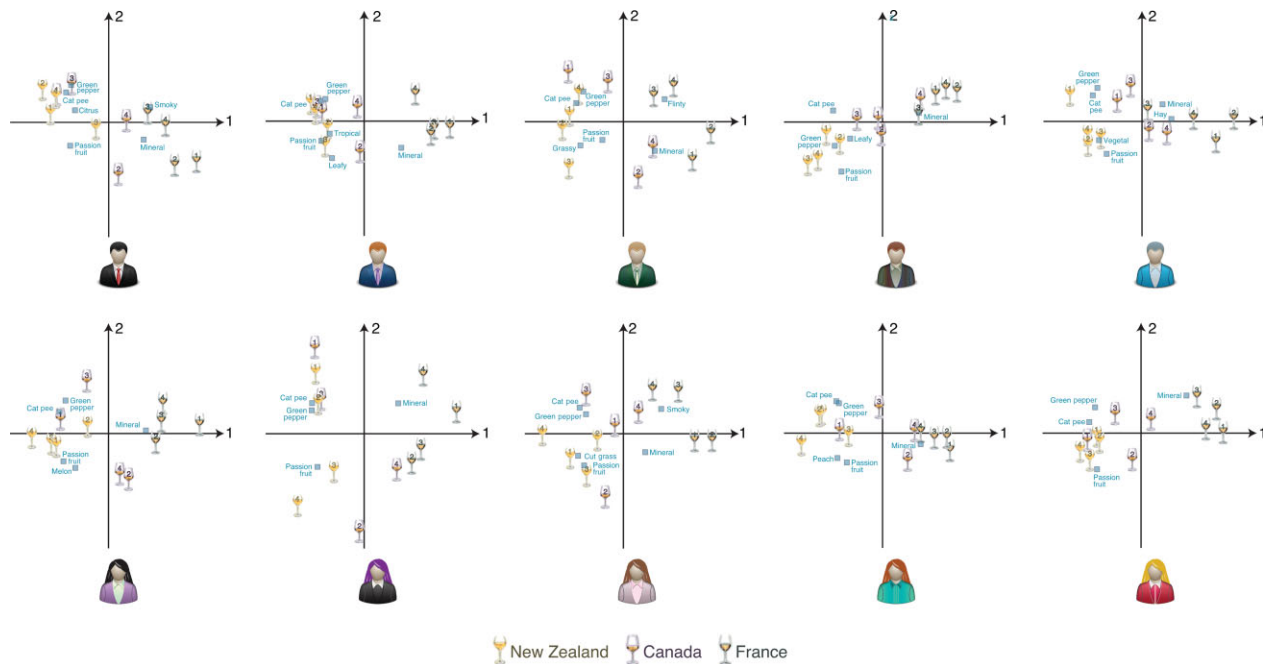
FIGURE 3 | Partial factor scores and variable loadings for the first two dimensions of the compromise space. The loadings have been re-scaled to have a variance equal the singular values of the compromise analysis.
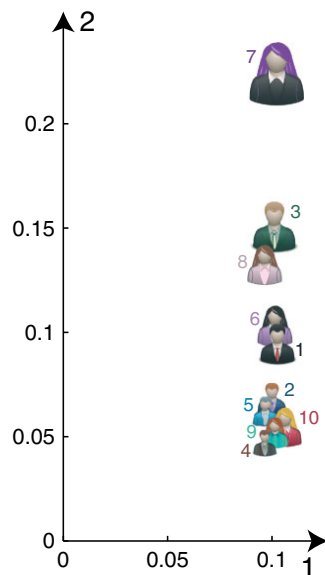

FIGURE 4 | Contributions of the tables to the compromise. The sizes of the assessors' icons are proportional to their contribution to Components 1 and 2.

FIGURE 5 | Partial Inertias of the tables. The sizes of the assessors' icons are proportional to their explained inertia for Components 1 and 2.

With this sample, we then computed the bootstrapped estimate of the factor scores as the projection as supplementary element of the whole bootstrapped matrix. Interestingly (for computational efficiency) these factor scores can directly be obtained as the average of the partial factor scores. Specifically, the bootstrapped estimate of the factor scores from the first bootstrap sample, is denoted $\mathbf{F}_1^*$ and computed as:

bootstrapped sample, we sampled with replacement in the set of the integers from 1 to 10 and obtained the following bootstrap set

$$\mathcal{B} = \left\{ \begin{array}{cccccccccc} 6 & 4 & 4 & 2 & 9 & 3 & 1 & 1 & 2 & 8 \end{array} \right\}. \quad (72)$$

$$\mathbf{F}_1^* = \frac{1}{K} \sum_{k \in \mathcal{B}} \mathbf{F}_{[k]}$$

$$= \mathbf{F}_{[6]} + \mathbf{F}_{[4]} + \mathbf{F}_{[4]} + \mathbf{F}_{[2]} + \mathbf{F}_{[9]} + \mathbf{F}_{[3]}$$

$$+ \mathbf{F}_{[1]} + \mathbf{F}_{[1]} + \mathbf{F}_{[2]} + \mathbf{F}_{[8]}$$

TABLE 4 | SupplementaryTable: Chemical Properties of the Wines

	Chemical Properties			
	Titratable Acidity	pH	Alcohol	Residual Sugar
NZ_1	5.60	3.38	14.00	3.00
NZ_2	5.30	3.53	13.50	3.60
NZ_3	6.20	3.27	14.00	3.00
NZ_4	8.50	3.19	13.50	3.90
F_1	5.00	3.60	12.50	1.50
F_2	5.88	3.00	12.50	2.00
F_3	4.50	3.33	13.00	0.80
F_4	5.60	3.40	12.00	2.10
CA_1	7.60	3.30	13.00	2.80
CA_2	5.70	3.43	13.50	2.10
CA_3	6.20	3.30	12.50	2.50
CA_4	6.90	2.20	13.00	2.00

FIGURE 6 | Supplementary table: chemical components of the wines. Supplementary partial scores and loadings. (cf., Figure 2a).

$$
= \begin{bmatrix}
-0.986 & 0.040 \\
-0.817 & 0.048 \\
-0.765 & -0.432 \\
-1.088 & 0.004 \\
1.335 & -0.157 \\
1.252 & -0.143 \\
0.856 & 0.164 \\
0.919 & 0.393 \\
-0.619 & 0.265 \\
0.096 & -0.692 \\
-0.524 & 0.449 \\
0.340 & 0.059
\end{bmatrix} . \tag{73}
$$

$$
\widehat{\sigma}_{\mathbf{F}^*} = \begin{bmatrix}
0.038 & 0.127 \\
0.077 & 0.105 \\
0.093 & 0.086 \\
0.070 & 0.156 \\
0.068 & 0.120 \\
0.082 & 0.117 \\
0.094 & 0.097 \\
0.060 & 0.110 \\
0.099 & 0.149 \\
0.073 & 0.139 \\
0.077 & 0.077 \\
0.072 & 0.124
\end{bmatrix} . \tag{75}
$$

From the bootstrap estimates, we can also compute bootstrap ratios, which, like t statistics, can be used to find the observations that reliably contribute to a given component. To get the bootstrap ratios, we first computed the mean of the bootstrap samples which are equal to (see Eq. (51))

$$
\overline{\mathbf{F}}^* = \frac{1}{L} \sum_{\ell}^{L} \mathbf{F}_\ell^* = \begin{bmatrix}
-0.979 & 0.163 \\
-0.810 & 0.037 \\
-0.762 & -0.456 \\
-1.115 & -0.167 \\
1.374 & -0.128 \\
1.261 & -0.110 \\
0.809 & 0.205 \\
0.927 & 0.410 \\
-0.672 & 0.374 \\
0.075 & -0.763 \\
-0.476 & 0.515 \\
0.369 & -0.081
\end{bmatrix} , \tag{74}
$$

and the standard deviations as (see Eq. (52)):

Then, a bootstrap ratio is computed by dividing each bootstrap mean by its standard deviation. This gives the following bootstrap ratios:

$$
\mathbf{T}^* = \begin{bmatrix}
-25.787 & 1.286 \\
-10.456 & 0.351 \\
-8.212 & -5.270 \\
-15.907 & -1.068 \\
20.056 & -1.061 \\
15.401 & -0.940 \\
8.628 & 2.122 \\
15.454 & 3.733 \\
-6.768 & 2.515 \\
1.024 & -5.497 \\
-6.224 & 6.687 \\
5.091 & -0.654
\end{bmatrix} . \tag{76}
$$

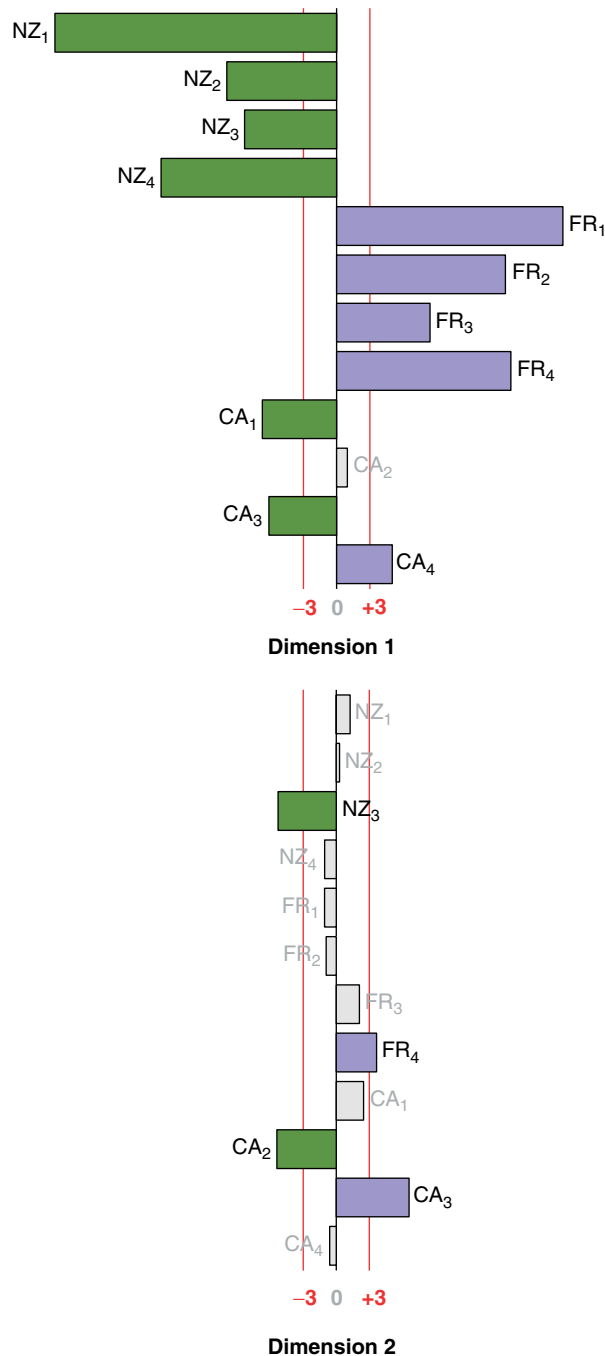These bootstrap ratios can be plotted as a bar chart to give a visual representation of which observations

Dimension 1

Dimension 2

From the bootstrap ratios shown in Figure 7 (see also Figure 8), we can see that all of the wines—except the second Canadian Sauvignon Blanc—contribute reliably to the first component, with the New Zealand wines separated from the French wines. This confirms our previous interpretation of Figure 2a. However, only 4 wines contribute reliably to the second component, with the third New Zealand Sauvignon Blanc grouping with the second Canadian wine and the fourth French wine grouping together with the third Canadian wine.

We also used the set of bootstrapped factors scores to obtain the 95% CIs around the factor scores (see Figure 8). Here, around each wine, we fitted an ellipsoid that comprises 95% of the bootstrapped factor scores. This ellipsoid represents the possible positions of a wine for replications of the analysis (assuming that the assessors were randomly sampled from a population of assessors). When the ellipses of two wines do not overlap, these two wines can be considered as reliably differentiated by the assessors.

RECENT DEVELOPMENTS

HMFA: Hierarchical Multiple Factor Analysis

In standard MFA, each variable belongs to one and only one table. (e.g., a variable comes from one and only one assessor). In an analysis of variance framework, this structure is defined as a *nesting* relation stating that the experimental factor corresponding to the variables is *nested* in the table factor (see Ref 112). Formally we say that one factor is nested in another factor if every level of the first factor occurs in conjunction with one and only level of the second factor (i.e., a given variable represents the rating of one and only one assessor). In some cases, the variables (i.e., the 'variable factor')

most reliably contribute to a component (see Figure 7). To be conservative and to take into account the multiple comparison problem, we chose a bootstrap ratio critical value of ± 3. This value correspond roughly to a Bonferroni corrected p-value for $J = 53$ comparisons (i.e., p-corrected is equal to $\frac{0.05}{53} \approx 0.001$ which approximatively corresponds to a t value of 3).

can be nested in a first factor that is itself nested within another factor. For example, the first five assessors (see Figure 2) are males and the rest of the assessors are females and therefore the assessor factor is nested in the gender factor (i.e., one assessor is either a male or a female). Note that, by transitivity, the variables are also nested in the gender factor. The analysis of data sets with such a nested structure can be performed with the recently developed HMFA (see Refs 121,122, Ref 2 for a dual version, and Refs 26,29,30,54 for applications), that can be seen as an integration of MFA with the analysis of variance tradition.[112]

In HMFA, the MFA normalization is applied hierarchically. First all the tables at the first level are normalized by dividing them by their first singular value, then a subsequent SVD is performed on each group of (normalized) tables and the data of all the tables of a given group are then divided by the first singular value obtained from this last PCA. The procedure will then continue till the highest level of the hierarchy is reached. The grand matrix $\mathbf{X}$ is then analyzed as indicated in Eq. (38). Alternatively, the α weights can be computed hierarchically, stored in the $\mathbf{A}$ matrix (see Eq. (12)), and then analyzed as indicated in Eq. (13).

For example, the first singular value of the first table was equal to 2.037 (cf., Eq. (57)) when the first five tables (each normalized by division by their first singular value) are analyzed by a new PCA, we now find that the first singular value of this 'men-table' is equal to 2.154. Therefore the HMFA-α weight for the first table will be equal to $(2.037 \times 2.154)^{-2} = 0.0519$. Taking into account that the first singular value of the PCA of the last five tables is equal to 2.169 and the values from Eq. (61), we can compute the HMFA α weights as:

$$
\boldsymbol{\alpha}_{\mathrm{HMFA}} =
\begin{bmatrix}
0.241 \\
0.239 \\
0.275 \\
0.273 \\
0.307 \\
0.302 \\
0.417 \\
0.272 \\
0.264 \\
0.309
\end{bmatrix}
\circ
\begin{bmatrix}
2.154^{-2} \\
2.154^{-2} \\
2.154^{-2} \\
2.154^{-2} \\
2.154^{-2} \\
2.169^{-2} \\
2.169^{-2} \\
2.169^{-2} \\
2.169^{-2} \\
2.169^{-2}
\end{bmatrix}
=
\begin{bmatrix}
0.0519 \\
0.0514 \\
0.0592 \\
0.0588 \\
0.0661 \\
0.0643 \\
0.0886 \\
0.0579 \\
0.0560 \\
0.0656
\end{bmatrix}.
\tag{77}
$$

From $\boldsymbol{\alpha}_{\mathrm{HMFA}}$, we can build the $\mathbf{a}_{\mathrm{HMFA}}$ vector and the $\mathbf{A}_{\mathrm{HMFA}}$ weight matrix (see Eqs (11) and (12)). A GSVD with $\mathbf{M}$ and $\mathbf{A}_{\mathrm{HMFA}}$ (see Eq. (13)) will then

provide the following factors scores (see Eq. (14)):

$$
\mathbf{F}_{\mathrm{HMFA}} =
\begin{bmatrix}
-0.454 & 0.075 \\
-0.375 & 0.014 \\
-0.352 & -0.211 \\
-0.515 & -0.076 \\
0.635 & -0.060 \\
0.585 & -0.050 \\
0.373 & 0.095 \\
0.428 & 0.189 \\
-0.309 & 0.170 \\
0.034 & -0.350 \\
-0.220 & 0.237 \\
0.170 & -0.035
\end{bmatrix}.
\tag{78}
$$

As with standard MFA, these factors scores could have been obtained from a simple SVD of the grand matrix storing the (twice) normalized data matrices.

In our example, the singular value of the table of the male assessors (i.e., 2.154) is almost identical to the singular value of the table of the female assessors (i.e., 2.169), and, therefore, the effect of the normalization due to gender is very small. In fact, the factor scores of the HMFA (given in Eq. 78) are almost perfectly correlated with their respective factor scores of the plain MFA (given in Eq. 65).

DUAL-MFA

In dual-MFA, the data consist of K sets of observations measured on the same set of variables (see Ref 2, see also Ref 123 for a review of related methods). In this case, the pattern of Eq. (1) is replaced by

$$
\mathbf{X} =
\begin{bmatrix}
\mathbf{X}_{[1]} \\
\vdots \\
\mathbf{X}_{[k]} \\
\vdots \\
\mathbf{X}_{[K]}
\end{bmatrix},
\tag{79}
$$

with $\mathbf{X}_{[k]}$ being an $I_{[K]}$ by J matrix (measuring the J variables for the kth sample of observations).

Here, instead of computing K cross-product matrices between the observations, we compute K cross-product covariance matrices between the variables (one per set of observations). The dual-MFA approach then follows the same steps as standard MFA and will provide a compromise map for the variables (instead of the observations in MFA), and partial loadings for each table.

An issue in dual-MFA is how to center the variables. If the variables are centered for the whole table (i.e., the mean of all the observations of a

variable is subtracted from each observation) then the analysis will be sensitive to differences in means between the tables (this will correspond to a main effect of the tables). If the data are considered as *ipsative* (i.e., the differences are meaningful only within a table, see Ref 124) then the variables should be centered separately for each table.

Procrustes MFA

Procrustes MFA (PMFA) (see Ref 36) can be used to analyze several Euclidean distance matrices describing the same observations. To do so, PMFA combines ideas from metric multidimensional scaling (MDS see Refs 101,125–127,153 for details) and MFA. Recall that, in MDS, a (Euclidean) distance matrix between some observations is transformed into a set of factor scores, ordered by their variance (i.e., eigenvalue), such that the distances between observations computed from their factor scores match (as well as possible) the original distances. These factor scores are then used in lieu of the variables of a plain MFA. Alternatively, in PMFA, each distance matrix, is analyzed via MDS and then the factor scores of each matrix are normalized *à la* MFA such that the first eigenvalue of each set of factor scores is equal to one, then all the normalized factor scores are collected into a grand matrix that is then analyzed with a plain PCA (see Eq. (38)).

MFA for Qualitative Data

MFA was originally developed as an extension of PCA and as such it requires quantitative variables. Just like PCA, MFA can be extended to handle qualitative variables. The equivalent of PCA for qualitative data is correspondence analysis or multiple correspondence analysis (see Refs 78,90,128,129). When dealing with multiple qualitative data tables, MFA becomes MFACT (for multiple factor analysis of contingency tables, see Refs 24,130–132, see also Refs 72,133 for an alternative approach called simultaneous analysis). The main idea of MFACT is the same as MFA namely to normalize each table by its first singular value. When all the rows in each table have the same sum (as in multiple correspondence analysis), this approach boils down to implementing the MFA normalization and then to proceed to the analysis of the grand data table (with correspondence analysis instead of PCA). When the rows of each table do not have the same sum then these sums need to be taken into account and each table needs to be expressed as the deviations to its own centroid (see Refs 130,131 for details). This approach can also be extended to the analysis of mixture of quantitative and qualitative data (see Refs 134,154).

MFA Barycentric Discriminant Analysis: MUFABADA

The idea of MFA can also be incorporated in the multiblock barycentric discriminant analysis framework (MUBADA, see Refs 4,5,10,135,136). Recall that MUBADA and related techniques, generalize discriminant analysis to the multitable case when the goal is to assign observations to a prior defined groups when these observations are described according to several tables. As in illustration, for our example, the goal of the analysis would be to use the description provided by the ten assessors to assign each of the twelve wines to one of the three wine groups (i.e., France, Canada, or New Zealand). To do so, a multitable PCA is performed on the table of the barycenters (i.e., the mean vector of each category) of the groups and each of the observations is projected as supplementary element. The factor scores of the observations and the groups are then used to compute the Euclidean distances between each observation and all the groups. Then each observation is assigned to its nearest group. In the multiple factor analysis version of MUBADA (called multiple factor barycentric discriminant analysis or MUFABADA), each data table is first centered and normalized and the table of the group barycenters is computed. For each group barycenter table, the first singular value is computed and this singular value is then used to normalize (like in MFA) the group barycenters and the observation tables. A standard MFA is then performed on the group barycenters grand data table and the observations are projected as supplementary elements.

In order to assess the quality of the discrimination, confusion matrices can be computed (as fixed or random effect models), and bootstrap confidence ellipsoid can be computed (in our example, the bootstrap can be performed in the factor space and will treat both wines and assessors as random factors). As an illustration, the result of MUFABADA performed on the wines is shown in Figure 9. Interestingly, the result is almost identical to the more sophisticated canonical STATIS analysis of the same data (cf., Ref 3, for the original analysis), and in fact the same figure can illustrate both techniques.

RELATED METHODS

MFA is part of the multitable family and also of the Procrustes family. A thorough evaluation of the complex relationships between all these techniques is beyond the scope of this article, but some directions for comparisons could be of interest. MFA is also closely related to generalized canonical correlation analysis GCCA. We will look first at the relationship between
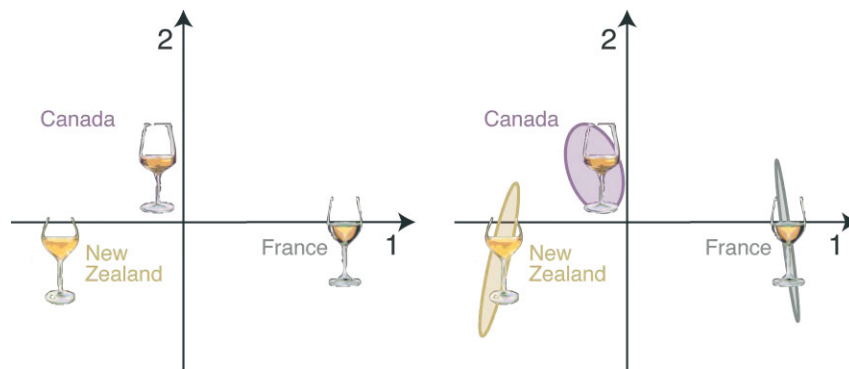
FIGURE 9 | MUFUBADA. Left : Discriminant factor scores for the three wine groups (regions). Right: Discriminant factor scores for the three wine groups with bootstrapped 95% confidence intervals.

MFA and GCCA and then at the relationships between MFA and the other techniques.

MFA and Generalized Canonical Correlation Analysis

An important multitable technique is generalized canonical correlation analysis (GCCA). In GCCA,[137–139] the goal is to maximize the inner-product of the factor scores of the compromise under a constraint that makes it equivalent to maximizing the correlations between the compromise factor scores and the partial factor scores. Specifically, in GCCA, we seek a matrix of J weights by L dimensions (with L being the number of dimensions for the solution) $\mathbf{W}_{\text{gcca}}$ such that

$$\mathbf{F}_{\text{gcca}} = \mathbf{X}\mathbf{W}_{\text{gcca}} \text{ with } \text{trace}\left\{\mathbf{F}_{\text{gcca}}^{\mathsf{T}}\mathbf{F}_{\text{gcca}}\right\} = \max. \quad (80)$$

under the constraints that

$$\mathbf{W}_{\text{gcca}}^{\mathsf{T}}\widetilde{\mathbf{\Sigma}}\mathbf{W}_{\text{gcca}} = \mathbf{I} \quad (81)$$

where $\widetilde{\mathbf{\Sigma}}$ is a J by J block diagonal matrix with each $J_{[k]}$ by $J_{[k]}$ diagonal block equal to $\frac{1}{I}\mathbf{X}_{[k]}^{\mathsf{T}}\mathbf{X}_{[k]}$ (note that, here, as in several equivalent techniques, the inverse can be replaced by a pseudo-inverse[138,139]). The set of $\mathbf{W}_{\text{gcca}}$ weights is obtained from the following GSVD of $\mathbf{X}$:

$$\mathbf{X} = \mathbf{P}_{\text{gcca}}\mathbf{\Delta}_{\text{gcca}}\mathbf{Q}_{\text{gcca}}^{\mathsf{T}} \text{ with } \mathbf{P}_{\text{gcca}}^{\mathsf{T}}\mathbf{M}\mathbf{P}_{\text{gcca}}$$
$$= \mathbf{Q}_{\text{gcca}}^{\mathsf{T}}\widetilde{\mathbf{\Sigma}}^{-1}\mathbf{Q}_{\text{gcca}} = \mathbf{I}. \quad (82)$$

The factors scores $\mathbf{F}_{\text{gcca}}$ are obtained as

$$\mathbf{F}_{\text{gcca}} = \mathbf{P}_{\text{gcca}}\mathbf{\Delta}_{\text{gcca}} = \mathbf{X}\widetilde{\mathbf{\Sigma}}^{-1}\mathbf{Q}_{\text{gcca}}$$
$$= \mathbf{X}\mathbf{W}_{\text{gcca}} \text{ with } \mathbf{W}_{\text{gcca}} = \widetilde{\mathbf{\Sigma}}^{-1}\mathbf{Q}_{\text{gcca}}. \quad (83)$$

So, in brief, GCCA and MFA are both multiblock technique. By contrast to MFA, GCCA is an anisotropic technique (i.e., it gives a different weight to each variable, whereas MFA gives the same weight to all variables of a table). Also GCCA minimizes a well defined criterion. Further work is needed to develop the comparison between GCCA, MFA, and their variants.

MFA and Procrustes Methods

Interestingly, MFA was originally[98,140] developed as a possible alternative to general Procrustes analysis (GPA)—a technique whose goal is to find the common factor solution from a set of factor scores obtained on the same observations. In fact, as mentioned earlier, PMFA—the specific Procrustes version of MFA—was recently developed.[36] Compared to GPA, MFA has the obvious advantage of being an eigendecomposition technique and therefore it does not require multiple iterations to reach a consensus and is also guaranteed to converge. A main point of difference between the two techniques is the space they try to fit: MFA considers the whole space (i.e., the whole cross product matrices) in which the K tables are, whereas GPA considers the specific dimensions (i.e., the factor scores) and is, in general, restricted to a subset of these dimensions. However for the set of dimensions kept, GPA, by contrast with MFA, optimizes a well defined criterion (see Refs 36,140, for a discussion of this point).

Procrustes multiple factor analysis is also obviously related to DISTATIS which is a technique that combines MDS and STATIS and could be obtained by adding a STATIS like optimization step after the MDS solution for each table has been computed. This step ensures that the compromise will optimally represents the original data matrices by adding

a optimum weight to each of the data table (for more details (see Refs 59,141–143,153,156)). Actually, it can be shown that PMFA corresponds to a 'degenerated' DISTATIS model in which the optimization step is dropped (and when the MFA normalization is applied as a pre-processing step).

MFA and INDSCAL

MFA can also be seen as a constrained simplified version of INDSCAL (see Refs 18,144). INDSCAL having more free parameters than MFA is likely to provide a better fit to the data than MFA. Also MFA being an eigenvector base technique is guaranteed to converge whereas INDSCAL is an iterative technique whose convergence is not always guaranteed.

MFA and Other Multiblock Analyses

MFA is also part of the multitable or multiblock family of PCA extensions.[7–20,145–147] The well-known members of this family—such as STATIS, SUM-PCA, consensus PCA, and multiblock correspondence analysis—all reduce to the PCA of a matrix X in which each $X_{[k]}$ has been normalized in a specific way for each technique. Therefore, the results obtained by these different methods are often similar.

Alternative To the MFA Table Weighting Schemes

There are several alternative (or complementary) ways to the MFA approach of normalizing data tables. The simplest one is to divide all entries of the data table by its number of columns or, better, by the square root of the number of columns (as done in the so-called 'Tucker-1' model and in 'consensus PCA'[148–150]). Another straightforward transformation divides all entries of the data table by the square root of the total sum of squares of its elements. This sets the total variance of each table to one and will guarantee that all tables participate equally in the analysis. This normalization is used, for example, in the SUM-PCA technique which originated in the chemometrics tradition (see Refs 19,20). A closely related procedure is to normalize $X_{[k]}$ by the square root of the norm of the matrix $Y_{[k]}Y_{[k]}^T$. Specifically, here $X_{[k]}$

is obtained as

$$X_{[k]} = Y_{[k]} \times \|Y_{[k]}Y_{[k]}^T\|^{-\frac{1}{2}}. \qquad (84)$$

This normalization ensures that the cross-product matrices $S_{[k]}$ all have norm equal to one. This normalization is particularly relevant for multitable analysis such as STATIS (see Ref 3). Another weighting scheme is obtained in the STATIS method by computing weights obtained in order to make the observation factor scores most similar to the partial factor scores (see Ref 3 for more details).

COMPUTER PACKAGES

Several statistical packages implement MFA and some of their variants. The package `FactoMineR`,[151] is a comprehensive R package written by the group that originally developed MFA and can be considered as the reference for MFA and its extensions. Also, R-programs—implementing the techniques described in this article—written by Derek Beaton, Cherise Chin, and Hervé Abdi are also available from the R package `MExPosition` (available from CRAN and from the first author's address at www.utdallas.edu/~herve). The R package `ade4` incorporates MFA as part of its multitable methods.[152,153] Simultaneous analysis (an extension of MFA to contingency tables) is available from the R package `SimultAnR` developed by Amaya Zarraga and Beatriz Goitisolo (package available from CRAN). In addition, the techniques described in this article have been programmed in MATLAB and are available from the first author's home page (at (www.utdallas.edu/~herve). Among commercial programs, XLSTAT has a comprehensive implementation of MFA (equivalent to the implementation of `FactoMineR`).

CONCLUSION

MFA is part of the multitable family of PCA related techniques. As such it is a simple, elegant, versatile, and robust technique that can be used to integrate multiple data tables collected on the same set of observations. Its simplicity (both theoretical and computational) makes it an ideal tool for the very large data sets of modern science.

REFERENCES

1. Thurstone LL. *Multiple Factor Analysis*. Chicago, IL: University of Chicago Press; 1947.

2. Lê S., Pagès J. DMFA: dual hierarchical multiple factor analysis. *Commun Stat: Theory Methods* 2010, 39:483–492.

3. Abdi H, Williams LJ, Valentin D, Bennani-Dosse M. STATIS and DISTATIS: optimum multi-table principal component analysis and three way metric multidimensional scaling. *Wiley Interdiscip Rev: Comput Stat* 2012, 4:124–167.

4. Abdi H, Williams LJ, Connolly AC, Gobbini MI, Dunlop JP, Haxby JV. Multiple Subject Barycentric Discriminant Analysis (MUSUBADA): how to assign scans to categories without using spatial normalization. *Comput Math Methods Med* 2012, 2012:1–15. doi:10.1155/2012/634165.

5. Abdi H, Williams LJ, Beaton D, Posamentier M, Harris TS, Krishnan A, Devous MD. Analysis of regional cerebral blood flow data to discriminate among Alzheimer's disease, fronto-temporal dementia, and elderly controls: a multi-block barycentric discriminant analysis (MUBADA) methodology. *J Alzheimer's Dis* 2012, 31:s189–s201.

6. Abdi H. Multivariate analysis. In: Lewis-Beck M, Bryman A, Futing T, eds. *Encyclopedia for Research Methods for the Social Sciences*. Thousand Oaks, CA: Sage; 2003, 699–702.

7. Acar E, Yener B. Unsupervised multiway data analyis: a literature survey. *IEEE Trans Knowl Data Eng* 2009, 21:6–19.

8. Arcidiacono C, Sarnacchiaro P, Velleman R. Testing fidelity to a new psychological intervention for family members of substance misuers during implementation in Italy. *J Subst Abuse* 2008, 13:361–381.

9. Stanimirova I, Boucon C, Walczak B. Relating gas chromatographic profiles to sensory measurements describing the end products of the Maillard reaction. *Talanta* 2011, 83:1239–1246.

10. Williams LJ, Abdi H, French R, Orange JB. A tutorial on multi-block discriminant correspondence analysis (MUDICA): a new method for analyzing discourse data from clinical populations. *J Speech Lang Hear Res* 2010, 53:1372–1393.

11. Daszykowski M, Walczak B. Methods for the exploratory analysis of two-dimensional chromatographics data. *Talanta* 2011, 83:1088–1097.

12. Carlier A, Lavit C, Pagès M, Pernin M, Turlot J. A comparative review of methods which handle a set of indexed data tables. In: Coppi R, Bollasco S, eds. *Multiway Data Analysis*. Amsterdam, the Netherlands: North Holland; 1989, 85–101.

13. Derks EPPA, Westerhuis JA, Smilde AK, King BM. An introduction to multi-blocks component analysis by means of a flavor language case study. *Food Qual Prefer* 2003, 14:497–506.

14. Escofier B, Pagès J. Multiple factor analysis. *Comput Stat Data Anal* 1990, 18:121–140.

15. Guebel DV, Cánovas M, Torres N. Model Identification in presence of incomplete information by generalized principal component analysis: application to the common and differential responses of *Escherichia coli* to multiple pulse perturbation in continuous high-biomass density culture. *Biotechnol Bioeng* 2009, 104:785–795.

16. Hassani S, Martens M, Qannari EM, Hanafi M. Analysis of -omic data: graphical interpretation and validation tools in multi-block methods. *Chemometr Intell Lab Syst* 2010, 104:140–153.

17. Meyners M, Kunert J, Qannari EM. Comparing generalized Procrustes analysis and STATIS. *Food Qual Prefer* 2000, 11:77–83.

18. Qannari EM, Wakeling I, MacFie JH. A hierarchy of models for analyzing sensory data. *Food Qual Prefer* 1995, 6:309–314.

19. Smilde AK, Westerhuis JA, de Jong S. A framework for sequential multiblock component methods. *J Chemometr* 2003, 17:323–337.

20. Van Deun K, Smilde AK, van der Werf MJ, Kiers HAL, Van Mechelen IV. A structured overview of simultaneous component based data integration. *BMC-Bioinf* 2009, 10:246–261.

21. Escofier B, Pagès J. Méthode pour l'analyse de plusieurs groupes de variables: application à la caractérisation des vins rouges du Val de Loire. *Revue de Statistique Appliquée* 1983, 31:43–59.

22. Escofier B, Pagès J. Analyses Factorielles Simples et Multiples: Objectifs, Méthodes, Interprétation. Dunod, Paris: 1988.

23. Abdi H, Valentin D. Multiple factor analysis. In: Salkind NJ, ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007, 657–663.

24. Pagès J. *Mulitple Factor Analysis by Example Using R*. Boca Raton, IL: CRC Press, 2014. To appear.

25. Ares G, Giménez A, Barreiro C, Gámbaro A. Use of an open-ended question to identify drivers of liking of milk desserts. Comparison with preference mapping techniques. *Food Qual Prefer* 2010b, 21:286–294.

26. Bécue-Bertaut M, Lê S. Analysis of multilingual labeled sorting tasks: application to a cross-cultural study in wine industry. *J Sens Stud* 2011, 26:299–310.

27. Bécue-Bertaut M, Álvarez-Esteban R, Pagès J. Rating of products through scores and free-text assertions: comparing and combining both. *Food Qual Prefer* 2008, 19:122–134.

28. Causse M, Friguet C, Coiret C, Lépicier M, Navez B, Lee M, Holthuysen N, Sinesio F, Moneta E, Grandillo S. Consumer preferences for fresh tomato at the European scale: a common segmentation on taste and firmness. *J Food Sci* 2010, 75: S531–S541.

29. Cadoret M, Lê S, Pagès J. A factorial approach for sorting task data (FAST). *Food Qual Prefer* 2009, 20:410–417.

30. Cadoret M, Lê S, Pagès J. Statistical analysis of hierarchical sorting data (FAST). *J Sens Stud* 2009, 26:96–105.

31. Dehlholm C, Brockhoff PB, Meinert L, Wender A, Bredie WLP. Rapid descriptive sensory methods: comparison of free multiple sorting, partial napping, napping, flash profiling and conventional profiling. *Food Qual Prefer* 2012, 26:267–277.

32. Dooley L, Lee Y, Meullenet JF. The application of check-all-that-apply (CATA) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping. *Food Qual Prefer* 2010, 21:394–401.

33. Gómez AT, Hernandez Cervantes M, Lopez Velazquez J, Cabrera RS, Ramon Canul LG, Juarez Barrientos JM, Ramirez Rivera EdJ. Caracterización sensorial del queso fresco "cuajada" en tres localidades de Oaxaca, México: diferencias en la percepción sensorial[Sensory characterization of "cuajada style" fresh cheese in three locations in Oaxaca, Mexico: differences in sensory perception]. *Revista Venezolana de Ciencia y Tecnología de Alimentos* 2010, 1:127–140.

34. Hopfer H, Heymann H. A summary of projective mapping observations: the effect of replicates and shape, and individual performance measurements. *Food Qual Preference* 2013, 26:164–181.

35. Mestres C, Ribeyre F, Pons B, Fallet V, Matencio F. Sensory texture of cooked rice is rather linked to chemical than to physical characteristics of raw grain. *J Cereal Sci* 2011, 53:81–89.

36. Morand E, Pagès J. Procrustes multiple factor analysis to analyse the overall perception of food products. *Food Qual Prefer* 2006, 17:36–42.

37. Pagès J. 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 Qual Prefer* 2005, 16:642–649.

38. Le Quéré J-M, Husson F, Renard CMGC, Primault J. French cider characterization by sensory, technological and chemical evaluations. *LWT - Food Sci Technol* 2006, 39:1033–1044.

39. Pagès J, Tenenhaus M. Multiple factor analysis combined with PLS path modeling. Application to the analysis of relationship between physicochemical variables, sensory rpofiles and hedonic judgements. *Chemometr Intell Lab Syst* 2001, 58:261–273.

40. Tu VP, Valentin D, Husson F, Dacremont C. Cultural differences in food description and preference: contrasting Vietnamese and French panellists on soy yogurts. *Food Qual Prefer* 2010, 21:602–610.

41. Kvalheim OM. Interpretation of direct latent-variable projection methods and their aims and use in the analysis of multicomponent spectroscopic and chromatographic data. *Chemometr Intell Lab Syst* 1988, 4:11–25.

42. Stanimirova I, Daszykowski M, Van Gyseghem E, Bensaid FF, Lees M, Smeyers-Verbeke J, Massart DL, Vander Heyden Y. Chemometrical exploration of an isotopic ratio data set of acetylsalicylic acid. *Analytica Chimica Acta* 2005, 552:1–12.

43. Stanimirova I, Walczak B, Massart DL. Multiple factor analysis in environmental chemistry. *Analytica Chimica Acta* 2005, 545:1–12.

44. Bertrand F, Maumy M, Fussler L, Kobes N, Savary S, Grossman J. Using factor analyses to explore data generated by the National Grapevine Wood Diseases survey. *Case Stud Bus Ind Gov Stat* 2007, 1:183–202.

45. Bocard DF, Gillet F, Legendre P. *Numerical Ecology with R*. New York: Springer; 2011.

46. Carlson ML, Flagstad LA, Gillet F, Mitchell EAD. Community development along a proglacial chronosequence: are above-ground and below-ground community structure controlled more by biotic than abiotic factors? *J Ecol* 2010, 98:1084–1095.

47. Daufresne M, Bady P, Fruget J-F. Impacts of global changes and extreme hydroclimatic events on macroinvertebrate community structures in the French Rhône River. *Oecologia* 2007, 151:544–559.

48. Gailhard I, Gros P, Durbec J, Beliaeff B, Belin C, Nezan E, Lassus P. Variability patterns of microphytoplankton communities along the French coasts. *Marine Ecol Progress Ser* 2002, 242:39–50.

49. Jassey VEJ, Chiapusio G, Mitchell EAD, Binet P, Toussaint ML, Gilbert D. Fine-scale horizontal and vertical micro-distribution patterns of testate amoebae along a narrow Fen/Bog gradient. *Microb Ecol* 2010, 61:374–385.

50. Lamentowicz M, Lamentowicz U, Knaap WO, Gabka M, Mitchell EAD. Contrasting species: environment relationships in communities of testate amoebae, bryophytes and vascular plants along the Fen-Bog gradient. *Microb Ecol* 2009, 59:499–510.

51. Legendre P, Legendre L. *Numerical Ecology*. New York: Elsevier; 2012.

52. Mazouni N, Gaertner J, Deslous-Paoli J-M. Composition of biofouling communities on suspended oyster cultures: an in situ study of their interactions with the water column. *Mar Ecol Prog Ser* 2001, 214:93–102.

53. Poulard J-C, Léauté J-P. Interaction between marine populations and fishing activities: temporal patterns of landings of La Rochelle trawlers in the Bay of Biscay. *Aquat Living Resour* 2002, 15:197–210.

54. Franco J, Crossa J, Desphande S. Hierarchical multiple-factor analysis for classifying genotypes based on phenotypic and genetic data. *Crop Sci* 2010, 50:105.

55. Ruget F, Moreau JC, Ferrand M, Poisson S, Gate P, Lacroix B, Lorgeou J, Cloppet E, Souverain F.

Describing the possible climate changes in France and some examples of their effects on main crops used in livestock systems. *Adv Sci Res* 2010, 4:99–104.

56. Strohmeier D, Jumisko-Pyykkö S, Kunze K, Bici MO. The extended-OPQ method for user-centered quality of experience evaluation: a study for mobile 3D video broadcasting over DVB-H. *EURASIP J Image Video Process* 2011:1–24.

57. Duvert C, Gratiot N, Evrard O, Navratil O, Némery J, Prat C, Esteves M. Drivers of erosion and suspended sediment transport in three headwater catchments of the Mexican Central Highlands. *Geomorphology* 2010, 123:243–256.

58. Kumagai T, Utsugi A. Removal of artifacts and fluctuations from MEG data by clustering methods. *Neurocomputing* 2004, 62:153–160.

59. Abdi H, Dunlop JP, Williams LJ. How to compute reliability estimates and display confidence and tolerance intervals for pattern classifiers using the Bootstrap and 3-way multidimensional scaling (DISTATIS). *NeuroImage* 2009, 45:89–95.

60. Buchsbaum BR, Lemire-Rodger S, Fang C, Abdi H. The neural basis of vivid memory is patterned on perception. *J Cogn Neurosci* 2012, 24:1867–1883.

61. Bécue-Bertaut M, Kern J., Hernández-Maldonado M-L, Juresa V, Vuletic S. Health-risk behaviour in Croatia. *Public Health* 2008, 122:140–150.

62. Damiani G, Farelli V, Anselmi A, Sicuro L, Solipaca A, Burgio A, Iezzi DF, Ricciardi W. Patterns of Long Term Care in 29 European countries: evidence from an exploratory study. *BMC Health Serv Res* 2011, 11:316

63. Joesch JM, Spiess K. European mothers' time spent looking after children-differences and similarities across nine countries. *Electron Int J Time Use Res* 2006, 3:1–27.

64. Wong S, Gauvrit H, Cheaib N, Carré F, Carrault G. Multiple factor analysis as a tool for studying the effect of physical training on the autonomic nervous system. *Comput Cardiol* 2002, 2002:437–440.

65. Pavoine S, Bailly X. New analysis for consistency among markers in the study of genetic diversity: development and application to the description of bacterial diversity. *BMC Evol Biol* 2007, 7:156–172.

66. de Tayrac M, Lê S, Aubry M, Mosser J, Husson F. Simultaneous analysis of distinct Omics data sets with integration of biological knowledge: multiple factor analysis approach. *BMC Genomics* 2009, 10:32.

67. Ares G, Deliza R, Barreiro C, Giménez A, Gámbaro A. Comparison of two sensory profiling techniques based on consumer perception. *Food Qual Prefer* 2010, 21:417–426.

68. Bécue-Bertaut M, Fernández-Aguirre K, Modrono-Herrán J. Analysis of a mixture of closed and open-ended questions in the case of a multilingual survey. In: CH Skiadas, eds. *Advances in Data Analysis: Theory and Applications to Reliability and Inference, Data Mining, Bioinformatics, Lifetime Data, and Neural Networks*. Boston: Birkhauser; 2008, 21–31.

69. Costard S, Porphyre V, Messad S, Rakotondrahanta S, Vidon H, Roger F, Pfeiffer, DU. Multivariate analysis of management and biosecurity practices in small-holder pig farms in Madagascar. *Preventive Veterinary Med* 2009, 92:199–209.

70. Pagès J, Husson F. Inter-laboratory comparison of sensory profiles. *Food Qual Prefer* 2001, 12:297–309.

71. Worch T, Lê S, Punter P. How reliable are the consumers? Comparison of sensory profiles from consumers and experts. *Food Qual Prefer* 2010, 21:309–318.

72. Zárraga A, Goitisolo B. Simultaneous analysis and multiple factor analysis for contingency tables: two methods for the joint study of contingency tables. *Comput Stat Data Anal* 2009, 53:3171–3182.

73. Lautre IG, Fernández EA. A methodology for measuring latent variables based on multiple factor analysis. *Comput Stat Data Anal* 2004, 45:505–517.

74. Yanai H, Takeuchi K, Takane Y. *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition*. New York: Springer Verlag; 2011.

75. Abdi H, Williams LJ. Principal component analysis. *Wiley Interdiscip Rev: Comput Stat* 2010, 2:433–459.

76. Abdi H. Singular value decomposition (SVD) and generalized singular value decomposition (GSVD). In: Salkind N, ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007, 907–912. de Leeuw J. Derivatives of generalized eigensystems with applications. UCLA Department of Statistics Papers, 2007, 1–28.

77. Greenacre M. *Theory and Applications of Correspondence Analysis*. London, UK: Academic Press; 1984.

78. Takane Y. Relationships among various kinds of eigenvalue and singular value decompositions. In: Yanai H, Okada A, Shigemasu K, Kano Y, Meulman J, eds. *New Developments in Psychometrics*. Tokyo: Springer Verlag; 2002, 45–46.

79. Eckart C, Young G. The approximation of one matrix by another of lower rank. *Psychometrika* 1936, 1:211–218.

80. Cailliez F, Pagès JP. *Introduction à l'Analyse des Données*. Paris: SMASH; 1976.

81. Escoufier Y. Operators related to a data matrix: a survey. COMPSTAT: Proceedings in Computational Statistics; 17th Symposium Held in Rome, Italy, 2006. New York, Physica Verlag; 2007, 285–297.

82. Holmes S. Multivariate analysis: the French way. In: Nolan D, Speed T, eds. *Festschrift for David Freedman*. Beachwood: IMS; 2006, 1–14.

83. de la Cruz O, Holmes S. The duality diagram in data analysis: examples of modern applications. *Ann Appl Stat* 2011, 5:2266–2277.

84. Thioulouse J. Simultaneous analysis of a sequence of paired ecological tables: A comparison of several methods. *Ann Appl Stat* 2011, 5:2300–2325.

85. Abdi H. Centroid. *Wiley Interdiscip Rev: Comput Stat* 2009, 1:259–260.

86. Greenacre M. *Biplots in Practice*. Barcelona: Fundación BBVA; 2010.

87. Gower JC, Lubbe S., le Roux N. *Understanding Biplots*. New York: Wiley; 2011.

88. Husson F, Lê S, Pagès J. *Exploratory Multivariate Analysis by Example Using R*. Boca Raton, IL: CRC Press; 2011.

89. Lebart L, Morineau A, Warwick KM. *Multivariate Descriptive Statistical Analysis*. New York: Wiley; 1984.

90. Greenacre M. *Correspondence Analysis in Practice*. London: Chapman & Hall; 2007.

91. Escoufier Y. Le traitement des variables vectorielles. *Biometrics* 1973, 29:751–760.

92. Robert P, Escoufier Y. A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Appl Stat* 1976, 25:257–265.

93. Abdi H. RV coefficient and congruence coefficient. In: Salkind NJ, ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007, 849–853.

94. Abdi H. Congruence: congruence coefficient, RV coefficient, & Mantel coefficient. In: Salkind NJ, ed. *Encyclopedia of Research Design*. Thousand Oaks, CA: Sage; 2010, 222–229.

95. Gaertner JC, Chessel D, Bertrand J. Stability of spatial structures of dermersal assemblages: a multitable approach. *Aquat Living Resour* 1998, 11:75–85.

96. Bhattacharyya A. On a measure of divergence between two multinomial populations. *Sankhya* 1941; 7:401–406.

97. Escofier B. Analyse factorielle et distances répondant au principe d'équivalence distributionnelle. *Revue de Statistique Appliquée* 1978, 26:29–37.

98. Domenges D, Volle M. Analyse factorielle sphérique: une exploration. *Annales de l'INSEE* 1979, 35:3–83.

99. Rao CR. Use of Hellinger distance in graphical displays. In: Tiit EM, Kollo T, Niemi H., eds. *Multivariate Statistics and Matrices in Statistics*. Leiden, Netherlands: Brill Academic Publisher; 1995, 143–161.

100. Abdi H. Distance. In: Salkind N, ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007, 280–284.

101. Lebart L, Piron M, Morineau A. *Statistique Exploratoire Multidimensionnelle: Visualisations et Inférences en Fouille de Données*. Paris: Dunod; 2006.

102. Saporta G, Niang N. Principal component analysis: application to statistical process control. In: Govaert G, ed. *Data Analysis*. London: Wiley; 2009.

103. Pelé J, Abdi H, Moreau M, Thybert D, Chabbert M. Multidimensional scaling reveals the main evolutionary pathways of class A G-protein-coupled receptors. *PLoS One* 2011, 6: 1-10, e19094. doi:10.1371/journal.pone.0019094.

104. Pelé J, Bécu JM, Abdi H, Chabbert M. Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling. *BMC Bioinf* 2012, 13:133–140.

105. Gower JC Adding a point to vector diagrams in multivariate analysis. *Biometrika* 1968, 55:582–585.

106. Josse J, Husson F. Selecting the number of components in PCA using cross-validation approximations. *Comput Stat Data Anal* 2012, 56:1869–1879.

107. Peres-Neto PR, Jackson DA, Somers KM. How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Comput Stat Data Anal* 2005, 49:974–997.

108. Timmerman ME, Kiers HAL, Smilde AK. Estimating confidence intervals for principal component loadings: a comparison between the bootstrap and asymptotic results. *Br J Math Stat Psychol* 2007, 60: 295–314.

109. Abdi H, Williams LJ. Jackknife. In: Salkind NJ, ed. *Encyclopedia of Research Design*. Thousand Oaks, CA: Sage; 2010, 655–660.

110. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. New York: Chapman and Hall; 1993.

111. Abdi H, Edelman B, Valentin D, Dowling WJ. *Experimental Design and Analysis for Psychology*. Oxford: Oxford University Press; 2009.

112. Milan M. Applications of the parametric bootstrap to models that incorporate a singular value decomposition. *Appl Stat* 1995, 44:31–49.

113. Lebart L. Which Bootstrap for principal axes methods? In: Brito P, Bertrand P, Cucumel G, de Carvalho F, ed. *Selected Contributions in Data Analysis and Classification. Studies in Classification, Data Analysis, and Knowledge Organization*. Berlin: Springer Verlag; 2007, 581–588.

114. Husson F, Le Dien S, Pagès J. Confidence ellipse for the sensory profiles obtained by Principal Component Analysis. *Food Qual Prefer* 2005, 16:245–250.

115. Pagès J, Husson F. Multiple Factor Analysis with confidence ellipses: a methodology to study the relationships between sensory and instrumental data. *J Chemometr* 2005, 19:138–144.

116. Lê S, Husson F, Pagès J. Confidence ellipses applied to the comparison of sensory profiles. *J Sens Stud* 2006, 21:241–248.

117. Cadoret M, Husson F. Construction and evaluation of confidence ellipses applied at sensory data. *Food Qual Prefer* 2013, 28:106–115.

118. Dehlholm C, Brockhoff PB, Bredie WLP. Confidence ellipses: a variation based on parametric bootstrapping applicable on multiple factor analysis results for

rapid graphical evaluation. *Food Qual Prefer* 2012, 26:278–280.

119. Abdi H. The Bonferonni and Sidák corrections for multiple comparisons. In: Salkind N, ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007, 104–107.

120. Le Dien S, Pagès J. Hierarchical multiple factor analysis: application to the comparison of sensory profiles. *Food Qual Prefer* 2003, 14:397–403.

121. Le Dien S, Pagès J. Analyse Factorielle Multiple Hiérarchique. *Revue de Statistique Appliquée* 2003, 51:47–73.

122. Eslami A, Qannari EM, Kohler A, Bougeard S. General overview of methods of analysis of multi-group datasets. *Review of New Information Technologies* 2013. In press.

123. Horst P. *Factor Analysis of Data Matrices*. New York: Holt; 1965.

124. Young G, Householder AS. Discussion of a set of points in terms of their mutual distances. *Psychometrika* 1938, 3:19–22.

125. Torgerson W. *Theory and Methods of Scaling*. New York: Wiley; 1958.

126. Abdi H. Metric multidimensional scaling. In: Salkind, N, ed. *Encyclopedia of Measurement and Statistics*.Thousand Oaks, CA: Sage; 2007, 598–605.

127. Abdi H, Valentin D. Multiple correspondence analysis. In: Salkind NJ, ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007c, 651–657.

128. Abdi H, Williams LJ. Correspondence analysis. In: Salkind NJ, ed. *Encyclopedia of Research Design*. Thousand Oaks, CA: Sage; 2010d, 267–278.

129. Bécue-Bertaut M, Pagès J. A principal axis method for comparing multiple contingency tables: MFCAT. *Comput Stat Data Anal* 2004, 45:481–503.

130. Bécue-Bertaut M, Pagès J. Multiple factor analysis for contingency tables. In: Greenacre M, Blasius J, ed. *Multiple Correspondence Analysis and Related Methods*. Boca Raton, IL: Chapman & Hall; 2006, 300–326.

131. Kostov B, Bécue-Bertaut M, Husson F. Multiple Factor Analysis for contingency tables in the *FactoMineR* package. *R J*. Submitted for publication.

132. Zarraga A, Goitisolo B. Simultaneous analysis: a joint study of several contingency tables with different margins. In: Greenacre M, Blasius J, ed. *Multiple Correspondence Analysis and Related Methods*. Boca Raton, IL: Chapman and Hall/CRC; 2006, 327–350.

133. Pagès J. Analyse factorielle de données mixtes. *Revue de Statistique Appliquée* 2004, 52:93–111.

134. Abdi H, Williams LJ. Barycentric discriminant analysis (BADIA). In: Salkind NJ, ed. *Encyclopedia of Research Design*. Thousand Oaks, CA: Sage; 2010, 64–75.

135. Abdi H. Discriminant correspondence analysis. In Salkind NJ, ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007, 270–272.

136. Horst P. Generalized canonical correlations and their applications to experimental data. *J Clin Psychol* 1961, 17:331–347.

137. Takane Y, Yanai M, Hwang H. An improved method for generalized constrained canonical correlation analysis. *Comput Stat Data Anal* 2006, 50:221–241.

138. Takane Y, Hwang H, Abdi H. Regularized multiple set canonical correlation analysis. *Psychometrika* 2008, 73: 753–775.

139. Gower JC, Dijksterhuis GB. *Procrustes Problems*. Oxford: O.U.P; 2004.

140. Abdi H, Valentin D. STATIS. In: Salkind NJ, ed. *Encyclopedia of Measurement and Statistics*. Thousand Oaks, CA: Sage; 2007, 955–962.

141. Abdi H, Valentin D. Some new and easy ways to describe, compare, and evaluate products and assessors. In: Valentin D, Nguyen DZ, Pelletier L, ed. *New Trends in Sensory Evaluation of Food and Non-Food Products*. Ho Chi Minh, Vietnam: Vietnam National University-Ho chi Minh City Publishing House; 2007, 5–18.

142. Abdi H, Valentin D, Chollet S, Chrea C. Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. *Food Qual Prefer* 2007, 18:627–640.

143. Husson F, Pagès J. INDSCAL model: geometrical interpretation and methodology. *Comput Stat Data Anal* 2006, 50:358–378.

144. Areia A, Oliveira MM, Mexia JT. Models for a series of studies based on geometrical representation. *Stat Methodol* 2008, 5:277–288.

145. Abdi H. Partial least square regression, projection on latent structure regression, PLS-Regression. *Wiley Interdiscip Rev: Comput Stat* 2010, 2:97–106.

146. Krishnan A, Williams LJ, McIntosh AR, Abdi H. Partial Least Squares (PLS) methods for neuroimaging: a tutorial and review. *NeuroImage* 2011, 56: 455–475.

147. Tucker LR. The extension of factor analysis to three-dimensional matrices. In: Frederiksen N, Gulliken H, ed. *Contributions to Mathematical Psychology*. New York: Holt; 1964, 110–182.

148. Westerhuis JA, Kourti T, MacGregor JF. Analysis of multiblock and hierarchical PCA and PLS models. *J Chemometr* 1998, 12:301–321.

149. Næs T, Brockhoff PB, Tomic O. *Statistics for Sensory and Consumer Science*. London: Wiley; 2010.

150. Lê S, Josse J, Husson F. FactoMineR: an R package for multivariate analysis. *J Stat Softw* 2008, 24:1–18.

151. ThioulouseJ, Chessel D, Dolédec S, Olivier JM. ADE-4: a multivariate analysis and graphical display software. *Stat Comput* 1997, 7:75–83.

152. Dray S, Dufour AB, Chessel D. The ade4 package-II: two-table and K-table methods. *R News* 2007, 7:47–52.

153. Abdi H, Valentin D, O'Toole AJ, Edelman B. DISTATIS: the analysis of multiple distance matrices. In: Proceedings of the IEEE Computer Society: International Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 2005, 42–47.

154. Bécue-Bertaut M, Pagès J. Multiple factor analysis and clustering of a mixture of quantitative categorical and frequency data. *Comput Stat Data Anal* 2008, 52:3255–3268.

155. Chernick MR. *Bootstrap Methods: A Guide for Practitioners and Researchers*. New York: Wiley; 2008.

156. Pagès J. Eléments de comparaison entre l'analyse factorielle multiple et la méthode STATIS. *Revue de Statistique Appliquée* 1996, 44:81–95.