Food Quality and Preference 43 (2015) 34–46

Contents lists available at ScienceDirect

Food Quality and Preference

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

# A comparison of generalised procrustes analysis and multiple factor analysis for projective mapping data

O. Tomic [a], I. Berget [b], T. Næs [b,c,*]

[a] Norwegian Knowledge Centre for the Health Services, Pilestredet Park 7, P.O. Box 7004, St. Olavs plass, 0130 Oslo, Norway
[b] Nofima AS, Oslovegen 1, P.O. Box 210, N-1431 Ås, Norway
[c] University of Copenhagen, Faculty of Life Sciences, Department Food Science, Rolighedsvej 30, 1958 Fredriksberg, Copenhagen, Denmark

ARTICLE INFO

*Article history:*
Received 11 July 2014
Received in revised form 15 January 2015
Accepted 3 February 2015
Available online 9 February 2015

*Keywords:*
Projective mapping
Napping
Generalised procrustes analysis
GPA
Multiple factor analysis
MFA
Consumer test
Multiblock method
RV coefficient

ABSTRACT

Generalised procrustes analysis and multiple factor analysis are multivariate statistical methods that belong to the family of multiblock methods. Both methods are often used for analysis of data from projective mapping (a.k.a. Napping). In this study, generalised procrustes analysis and multiple factor analysis are compared for a number of simulated and real data sets. The type of data used in this study were (I) random data from Monte Carlo simulations; (II) constructed data that were manipulated according to some specific criteria; (III) real data from nine Napping experiments. Focus will be on similarities of the consensus solutions. In addition we considered interpretation of the RV coefficient and individual differences between assessors.



## 1. Introduction

In recent years, rapid sensory methods have gained a lot of interest in the field of sensory science (Dehlholm, Brockhoff, Meinert, Aaslyng, & Bredie, 2012; Valentin, Chollet, Lelièvre, & Abdi, 2012; Varela & Ares, 2012). These methods are generally simple to use, they can be applied with untrained assessors and the analysis can often be accomplished quickly. One of the best-known and most used methods in the category is projective mapping (Risvik, McEwan, Colwill, Rogers, & Lyon, 1994), also later known as Napping (Pagès, 2005). With this method, a number of individuals (typically between 10 and 100) are asked to place a number of products on a two-dimensional sheet according to how similar, or dissimilar, they consider the products to be, using their own criteria. Despite being documented to be less precise than descriptive sensory analysis (Valentin et al., 2012), projective mapping has gained much popularity especially within the food industry because of the advantages listed above. It should also be mentioned that the method sometimes can, due to its holistic character, provide additional information (see e.g. Varela & Ares, 2014) as compared to standard attribute based sensory methods.

By placing products on a sheet, each individual generates a two-dimensional data matrix representing the coordinates of all the placed products. These data need to be analysed with a suitable statistical method in order to extract information about the tested products, which can be utilised for further product development or product optimisation. The two most established methods for analysing projective mapping data are generalised procrustes analysis (GPA) (Gower, 1975) and multiple factor analysis (MFA) (Escofier & Pagès, 1994). Even though both GPA and MFA are conceptually very different, both belong to the family of the so-called multiblock methods (Abdi, Williams, & Valentin, 2013). They provide information about the "consensus" product configuration, which in practical terms represents the "mean" product configuration across all individuals and which gives important insight into the overall perception of the products. Although several other methods, for instance INDSCAL (Carroll & Chang, 1970), STATIS (Schlich, 1996) and the different Tucker methods, (Tucker, 1964) can be envisioned for handling this type of data, it is of interest to compare the two because of their frequent use.

To the authors' knowledge, there exists only one study that in some detail discusses the differences and similarities between

* Corresponding author at: Nofima AS, Oslovegen 1, P.O. Box 210, N-1431 Ås, Norway.
*E-mail address:* tormod.naes@nofima.no (T. Næs).

http://dx.doi.org/10.1016/j.foodqual.2015.02.004
0950-3293/

the two methods applied to the same set of projective mapping data. Nestrud and Lawless (2008) reported that both methods have been tested on the same data set and that results were very similar. In that study, GPA and MFA were applied to data that were generated from a single experiment where 13 citrus juices were evaluated by a group of experienced chefs and a group of untrained consumers.

The present study attempts to provide more insight into differences and similarities between results acquired with the methods GPA and MFA in the context of projective mapping. Focus will be on aspects related to standard practices in the area and a discussion will be given on how some of these standard practices may not be fully satisfactory. A secondary objective is to discuss the RV coefficient (Robert & Escoufier, 1976) which is used frequently in the area for comparing data sets and consensus solutions. For these purposes we will use both random data in Monte Carlo simulations, constructed data that were manipulated according some specific criteria as well as real data from nine Napping experiments. In particular, the following points will be addressed:

- The importance of proper validation of the consensus solution.
- The importance of using simple computer simulations in order to understand differences and similarities between methods better.
- The importance of looking at individual differences between assessors for obtaining information about validity and stability.
- The importance of accompanying the RV coefficient with graphical displays of the data.
- The importance of extending the focus to more than two principal components/dimensions.

## 2. Methods

### 2.1. Projective mapping

Projective mapping is a method where individuals evaluate the overall perception of a number of products and place them on a sheet according to the products' similarities or dissimilarities (Pagès, 2005; Risvik et al., 1994; Risvik, McEwan, & Rødbotten, 1997). Placement can be done either by putting products directly on a sheet of paper or by indicating their position on a computer screen. Individuals are instructed to place similar products close to each other using their own criteria or criteria given by the instructor. Other than that, individuals are generally not given further directions. If the placement of the products needs to be refined, the individuals may taste the products again until placement is considered to be satisfactory.

Optionally, individuals may be asked to write down sensory descriptors that best describe each group of products. By doing so, the projective map is turned into an Ultra Flash Profile method as described previously (Perrin et al., 2008, see also Williams & Arnold, 1985 for other situations where free assignment of words is relevant). In this study, however, focus will be only on the product coordinates derived from the positions of the products on the sheet or on the computer screen (two-dimensional data blocks in form of *x*- and *y*-coordinates).

A well know critique regarding projective mapping which is worth mentioning, is that complex multidimensional products may be difficult to place on the two-dimensional sheet since the two dimensions of the sheet may not be enough to distinguish the products properly and may then leave the user with a non-satisfying placement of the products. Recent research (Nestrud & Lawless, 2011), however, refutes this criticism by claiming that important components and configurations could be recovered using MFA and multidimensional scaling. Since the two first calculated components are the dominating ones, and also those that are given main attention in the literature, main focus will here be on these two components (see also scope indicated in the introduction). We will, however, also discuss briefly the importance of interpreting more than two components and indicate some paths of further development.

### 2.2. General structure of projective mapping data

Every individual taking part in the projective mapping trial, is supposed to place a number of products on a projective mapping sheet resulting in individual data blocks $\boldsymbol{Z}_k$ that are of dimension $(I \times J)$ with $J = 2$. Here $i = 1, \ldots, I$ represents the number of objects or products tested by the $k = 1, \ldots, K$ individuals.

### 2.3. Generalised procrustes analysis (GPA)

GPA (Dijksterhuis, 1996; Gower, 1975; Gower & Dijksterhuis, 2004) is a multivariate statistical method that is applied for multiple data blocks. The main goal is to acquire a consensus from the blocks after they have undergone Procrustes transformations that reduce individual differences by means of translation, rotation and reflection as well as isotropic scaling. GPA is therefore well suited for analysis of projective mapping data given our goal to find a consensus product configuration across all individuals. In most cases, one will use principal component analysis on the consensus to improve interpretation (optional). Since in our case the consensus is two-dimensional, the PCA only represents a rotation of the original axes found by the Procrustes transformation.

Clearly, there will always be variations in how the individuals place the products on the sheet. The variation between the data blocks $\boldsymbol{Z}_k$ comes from different perception of products, and because of the more or less arbitrary ways of using the directions on the mapping sheet. The former aspect represents the sensory differences between the products. One would, however, usually like to eliminate the latter since this is generally not related to differences between the products.

In more detail, the Procrustes transformation itself consists of three steps that can be summarised in the following way: (A.1) translation, meaning that all individual configurations are moved to the middle of the mapping sheet; (A.2) rotation and reflection of individual configurations until they are in best possible agreement with one another (Eq. (2) below); (A.3) isotropic scaling, i.e. shrinking or stretching of individual configurations until they are as alike as possible, but without changing the relative distances between the products in each configuration (Eq. (2)). Since the mean, scaling and rotation are related to individual differences of minor value for the interpretation of the product differences, the Procrustes method is very well suited for the situation. It preserves relative distances between objects (see criterion below), which may be seen as an advantage.

Mathematically, the three steps of the Procrustes transformation $\tau(\boldsymbol{Z}_k)$ may be summarised in the following way:

$$\tau(\boldsymbol{Z}_k) = \rho_k \boldsymbol{Z}_k \boldsymbol{H}_k + \boldsymbol{T}_k \qquad (1)$$

The $\boldsymbol{T}_k$ is the matrix of translation constants (step (A.1)), the $\boldsymbol{H}_k$ represents the rotation matrix (step (A.2)) and $\rho_k$ represents the isotropic scaling constant (step (A.3)). Note that $\boldsymbol{H}_k$ is an orthogonal matrix; $\boldsymbol{H}^T\boldsymbol{H} = \boldsymbol{H}\boldsymbol{H}^T = \boldsymbol{I}$.

Translation can be removed from Eq. (1) by centring of each variable first. The $\boldsymbol{H}_k$ and $\rho_k$ of each data block are then obtained by minimising (under a constraint on the total variability after isotropic scaling):

$$\sum_{k=1}^{K} \|\rho_k \boldsymbol{Z}_k \boldsymbol{H}_k - \boldsymbol{Y}_{GPA}\|^2 \qquad (2)$$