Principal component analysis of sensory panel results for a reference and multiple prototypes

JC Castura[1], P Varela[2], T Næs[2]
[1]Compusense Inc., 412 Laird Road, Unit 104, Guelph, Ontario, N1G 3X7, Canada
[2]Nofima AS, Osloveien 1, P.O. Box 210, N-1431 Ås, Norway

Abstract
A panel of trained sensory assessors often evaluates samples by quantifying the intensities of sensory attributes. In some cases, samples are instances of in-market or prototype products. To explore results from the panel, it is conventional to obtain a products-by-attributes table of means, center and variance-standardize its columns, then conduct principal component analysis (PCA). The principal components that extract variance maximally from the products also extract variance maximally from all product paired comparisons. However, if there is a successful in-market reference product and multiple prototypes, then this PCA does not extract variance maximally from only the reference-prototype paired comparisons of primary interest. To investigate the relevant subset of paired comparisons, we create the input matrix for PCA using only the rows for paired comparisons of primary interest. Advantages of PCA of this modified input matrix are demonstrated visually and numerically using a case study involving cheddar cheeses. Data sets with related structures, including temporal sensory data sets and data sets in domains outside sensory evaluation can also be investigated using this approach.

Key Words: sensory evaluation; principal component analysis; paired comparisons; truncated total bootstrap; multivariate analysis

1. Investigating paired comparisons after principal component analysis (PCA)

1.1 Investigating row objects via PCA

A column-centered ($J \times M$) matrix $\mathbf{X}$ with rank R can be submitted to singular value decomposition (SVD), which we describe in a similar manner as Mardia, Kent and Bibby (1979). Ignoring the null space, SVD of $\mathbf{X}=\mathbf{UDP}^{\mathrm{T}}$ where the ($J \times R$) matrix $\mathbf{U}$ has left singular vectors in columns, the ($R \times R$) diagonal matrix $\mathbf{D}$ has singular values along its diagonal, and the ($M \times R$) matrix $\mathbf{P}$ has right singular vectors in columns. Singular vectors are orthonormal such that $\mathbf{U}^{\mathrm{T}}\mathbf{U}=\mathbf{P}^{\mathrm{T}}\mathbf{P}=\mathbf{I}_R$. The sum of the squared singular values in $\mathbf{D}$ equals the sum of squared elements in $\mathbf{X}$.

We multiply $\mathbf{U}$ and $\mathbf{D}$ to obtain the ($J \times R$) score matrix $\mathbf{T}$. We will call $\mathbf{P}$ the loading matrix. PCA of $\mathbf{X}=\mathbf{TP}^{\mathrm{T}}$. Usually dimension reduction is possible: the score and loading matrices are truncated to $A<R$ principal components (PCs), such that only the first A columns of $\mathbf{T}$ and the first A columns of $\mathbf{P}$ are retained. After dimension reduction,

$$\mathbf{X}=\mathbf{T}_A\mathbf{P}_A^{\mathrm{T}}+\mathbf{E} \qquad (1)$$

where only the ($J \times M$) matrix $\mathbf{T}_A\mathbf{P}_A^{\mathrm{T}}$ is interpreted. The ($J \times M$) matrix $\mathbf{E}$ contains residuals that in practice are small enough to ignore.

1.2 Investigating all paired comparisons after PCA

This section highlights some of the key results from Castura, Varela and Næs (2022) for investigating all paired comparisons of row objects in $\mathbf{X}$ after PCA.

The ($J^2 \times M$) matrix $\mathbf{X} \ominus \mathbf{X}$ is obtained by subtracting every row in $\mathbf{X}$ from each row in $\mathbf{X}$, a procedure called "crossdiff-unfolding". $\mathbf{X} \ominus \mathbf{X}$ contains J rows of only zeros that occur when a row is subtracted from itself. Each row in the remaining $J(J\text{-}1)$ rows has a twin with the opposite sign; for example, the paired comparison of rows $\mathbf{x}_1$ and $\mathbf{x}_2$ is represented by two paired differences, $\mathbf{x}_1\text{-}\mathbf{x}_2$ and $\mathbf{x}_2\text{-}\mathbf{x}_1$. For this reason, columns of $\mathbf{X} \ominus \mathbf{X}$ center naturally. The variances of the columns of $\mathbf{X} \ominus \mathbf{X}$ are all equal since columns of $\mathbf{X}$ were standardized to unit variance. The covariance matrices of $\mathbf{X}$ and $\mathbf{X} \ominus \mathbf{X}$ are related by a scalar that depends only on J, the number of rows in $\mathbf{X}$, not on the data. PCA of $\mathbf{X} \ominus \mathbf{X} = (\mathbf{T} \ominus \mathbf{T})\mathbf{P}^{\mathrm{T}}$ where $\mathbf{T}$ and $\mathbf{P}$ are the same as in PCA of $\mathbf{X}$. Since row objects and their paired comparisons share the same directions of maximum variability, they are explored optimally in the same directions. These directions are defined by the columns of $\mathbf{P}$. Any PC a extracts the same proportion of variance from $\mathbf{X}$ as it extracts from $\mathbf{X} \ominus \mathbf{X}$. Dimension reduction of the PCA of $\mathbf{X} \ominus \mathbf{X}$ to A PCs gives

$$\mathbf{X} \ominus \mathbf{X} = (\mathbf{T}_A \ominus \mathbf{T}_A)\mathbf{P}_A^{\mathrm{T}} + (\mathbf{E} \ominus \mathbf{E}) \tag{2}$$

where $\mathbf{T}_A$, $\mathbf{P}_A$, and $\mathbf{E}$ are as defined in Section 1.1.

1.3 Investigating a subset of paired comparisons after PCA

In some cases, not all paired comparisons of row objects in $\mathbf{X}$ are of primary interest. This section highlights some of the key results from Castura, Varela and Næs (2023a) for investigating only a subset of the paired comparisons of row objects in $\mathbf{X}$ after PCA.

Suppose that C paired comparisons of row objects in $\mathbf{X}$ are relevant. Each of the C paired comparisons is associated with twinned paired differences (Section 1.2). The relevant subset of paired differences can be organized into the (2C$\times M$) matrix $\mathbf{\Delta}^*$. Each row in $\mathbf{\Delta}^*$ corresponds to a row in $\mathbf{X} \ominus \mathbf{X}$; i.e., there exists a ($2C \times M$) submatrix of relevant rows in $\mathbf{X} \ominus \mathbf{X}$ that is identical to $\mathbf{\Delta}^*$. Columns of $\mathbf{\Delta}^*$ center naturally. Since not all paired compared comparisons are relevant, the column variances of $\mathbf{\Delta}^*$ depend on which paired comparisons are relevant. Column variances will be larger for attributes in which the relevant paired differences are more dispersed than for attributes in which the relevant paired differences have low variability. The covariance matrix of $\mathbf{\Delta}^*$ is not related by a scalar to the covariance matrices of $\mathbf{X}$ and $\mathbf{X} \ominus \mathbf{X}$. PCA of $\mathbf{\Delta}^* = \mathbf{T}^*(\mathbf{P}^*)^{\mathrm{T}}$. In this case, $\mathbf{P} \neq \mathbf{P}^*$ almost always since only a subset of paired comparisons is relevant. Dimension reduction of the PCA of $\mathbf{\Delta}^*$ to A PCs yields

$$\mathbf{\Delta}^* = \mathbf{T}_A^*(\mathbf{P}_A^*)^{\mathrm{T}} + \mathbf{E}^* \tag{3}$$

where the ($J \times M$) matrix $\mathbf{T}_A^*(\mathbf{P}_A^*)^{\mathrm{T}}$ is interpreted and the ($J \times M$) residual matrix $\mathbf{E}^*$ is ignored.

Coordinates of the row objects in $\mathbf{X}$ can be obtained in the PCs of $\mathbf{\Delta}^*$ by

$$\mathbf{X}^* = \mathbf{X}\mathbf{P}_A^*. \tag{4}$$

Since the origin has the interpretation of no difference, the control product will not be located at the origin.

1.4 Gain of investigating a subset of paired comparisons after PCA

We summarize the approach used by Castura et al. (2023a) to quantify the benefit of PCA of $\mathbf{\Delta}^*$ over PCA of $\mathbf{X}\ominus\mathbf{X}$.

$SS(\mathbf{\Delta}^*)$ is sum of the squared elements in $\mathbf{\Delta}^*$, which is the relevant sum of squares. It is identical to the sum of the squared elements in the relevant submatrix of $\mathbf{X}\ominus\mathbf{X}$ containing the $2C$ relevant paired differences.

$SS(\mathbf{T}_A\ominus\mathbf{T}_A)$ is the relevant sum of squares extracted in the first A PCs of $\mathbf{X}\ominus\mathbf{X}$, which is obtained by summing the squared elements in the $2C$ rows in the relevant submatrix of $\mathbf{T}_A\ominus\mathbf{T}_A$. The proportion of the relevant sum of squares that is extracted in the first A PCs of $\mathbf{X}\ominus\mathbf{X}$ is $SS(\mathbf{T}_A\ominus\mathbf{T}_A)/SS(\mathbf{\Delta}^*)$.

The relevant sum of squares extracted in the first A PCs of $\mathbf{\Delta}^*$ is the sum of the squared elements of $\mathbf{T}_A^*$, which is denoted $SS(\mathbf{T}_A^*)$. The proportion of the relevant sum of squares that is extracted in the first the first A PCs of $\mathbf{\Delta}^*$ is $SS(\mathbf{T}_A^*)/SS(\mathbf{\Delta}^*)$.

The benefit of investigating the relevant paired comparisons in the first A PCs of $\mathbf{X}\ominus\mathbf{X}$ instead of the first A PCs of $\mathbf{\Delta}^*$ can be quantified by the percentage

$$\text{Gain}=100(SS(\mathbf{T}_A^*)/\ SS(\mathbf{T}_A\ominus\mathbf{T}_A)-1)\%. \qquad (5)$$

For the approaches given in Sections 1.2 and 1.3, Gain is non-negative since $SS(\mathbf{T}_A^*)\geq SS(\mathbf{T}_A\ominus\mathbf{T}_A)$. Gain might be (nearly) zero if $\mathbf{X}\ominus\mathbf{X}$ and $\mathbf{\Delta}^*$ have the same directions of maximum variability or if A is large; however, in most other cases Gain is positive.

1.5 Visual and numerical approaches for evaluating paired comparisons

The truncated total bootstrap (TTB; Cadoret & Husson, 2013) method can be used to investigate uncertainty in a PCA solution. TTB results can be used to obtain nonparametric uncertainty regions (Castura, Varela & Næs, 2022) that can be approximated by 95% confidence ellipses (Castura, Varela and Næs, 2023b) where exclusion of the origin indicates a significant difference. We also calculate P values that coincide in interpretation with these confidence ellipses (Castura et al., 2023b).

2. Materials and Methods

2.1 Cheddar cheese data set

A panel of 10 assessors trained in sensory descriptive analysis evaluated eight cheddar cheese products in duplicate in laboratory conditions. Details of these products are withheld for confidentiality reasons. The market-leading product was designated a benchmark, reference, or control product which, in some cases, has a sensory profile that other companies want to match. Although we do not know the business objectives in this study, we assume that the other products were prototypes or test products. In figures shown later, the control product is denoted “C” and the seven test products are denoted “T1”, …, “T7”.

Samples were presented in sequential monadic format using an experimental design to balance presentation order and carry-over effects. Each sample was evaluated on seven flavour/taste attributes (*maturity*, *rindy*, *creamy*, *sweet*, *bitter*, *acidic*, *salty*) and six texture attributes (*first-bite firmness*, *rubbery*, *pasty*, *bitty breakdown*, *dry*, *breakdown rate*). Each attributes had a precise definition; in training, assessors had practiced identifying and quantifying the perceived intensity of each attribute. Attributes were evaluated using line scales anchored at 0 and 100, where endpoints corresponded respectively to none and high, or for *breakdown rate* to slow and fast.

Before conducting PCA, attributes were screened using a two-way analysis of variance with factors assessor and product. Any attribute on which the panel failed to discriminate the products with 95% confidence was discarded (Næs, Tomic, Endrizzi & Varela, 2021). In the cheddar cheese data set, the product effect was significant for all attributes ($p<0.01$) except *creamy flavour* ($p=0.16$) which was dropped from further analysis. A products-by-attributes matrix of panel means was obtained for the eight products and 12 sensory attributes. Columns of the products-by-attributes matrix of means were centered and variance-standardized to produce the matrix $\mathbf{X}$.

2.2 Statistical software

All analyses were conducted in R version 4.3.1 (R Core Team, 2023). Two-way analysis of variance was conducted using the `Anova` function in the R package `car` (Fox & Weisberg, 2019) with contrasts that summed to zero and type III sum of squares. After PCA of all paired comparisons, the solution was truncated to retain as few PCs as were needed to account for at least 80% of the variance. The PCA of selected paired comparisons was truncated to the same number of PCs to facilitate comparisons.

3. Results

3.1 PCA of all paired comparisons

The matrix $\mathbf{X}$ from Section 2.1 was crossdiff-unfolded to obtain $\mathbf{X}\ominus\mathbf{X}$. PCA of $\mathbf{X}$ and PCA of $\mathbf{X}\ominus\mathbf{X}$ as in (1) and (2) respectively yielded solutions in which the first four PCs extracted 53.6%, 31.0%, 10.4%, and 2.5% of the variance in $\mathbf{X}$ and in $\mathbf{X}\ominus\mathbf{X}$. We truncated the solution to two PCs (i.e., A=2) that together extracted 84.7% of the variance both from $\mathbf{X}$ and from $\mathbf{X}\ominus\mathbf{X}$.

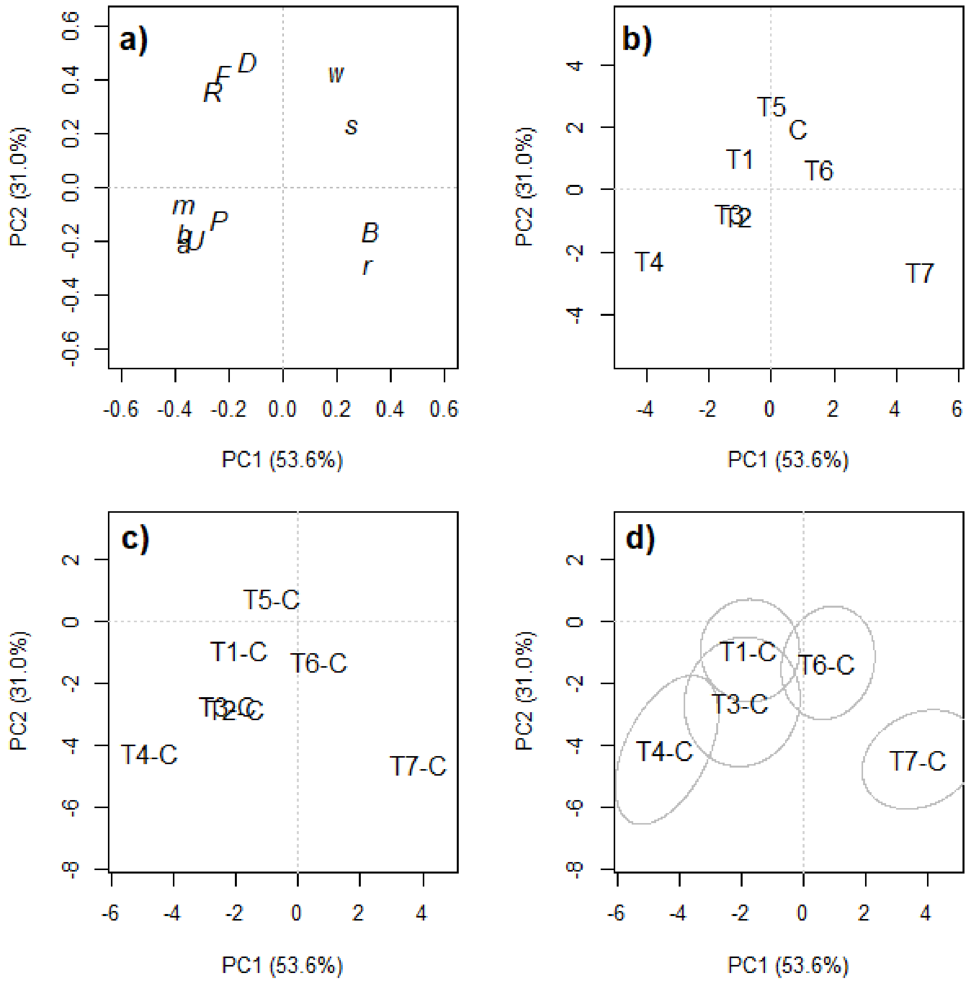
The loading coefficients ($\mathbf{P}_A$), which are shown in Figure 1a, were the same in both PCA solutions. Mature cheddar tanginess tended to coincide with a more pasty texture, more bitter and acidic tastes, and lower sweetness and saltiness. High firmness and dryness tended to coincide with a faster breakdown rate, less breakdown of samples into bits under compression, and a lower rindy flavour, which refers to an earthy note in mature cheddar rind.

The configuration of products was visualized after PCA of $\mathbf{X}$ by plotting their respective scores ($\mathbf{T}_A$), as shown in Figure 1b. C was sweeter and saltier than most other products. Compared to C, T5 and T1 tended to be drier, firmer, and break down faster. Products T2, T2, and especially T4 tended have more mature flavours, higher bitterness and acidity, and a more rubbery and pasty texture. T7 had relatively rindy flavours and tended not to break down into bits under compression.

All paired differences can be obtained either by subtraction of the respective product row vectors in $\mathbf{T}_A$ or directly from the relevant rows in $\mathbf{T}_A\ominus\mathbf{T}_A$. Since the interpretations for the control-test paired differences mirror the interpretations for the test-control paired differences, we plot only the latter in Figure 1c. T5 and C had similar levels of sweetness and saltiness; these products differed mainly in that T5 was drier, firmer, and had a faster breakdown rate than C. T6 and C also had similar levels of sweetness and saltiness; they differed mainly in that T6 tended to have more rindy flavours and break into bits more than C. Differences between the T7-C were directionally similar but more pronounced than the T6-C differences. Relative to C, the test products T1 to T5 tended to have more mature flavours, higher bitterness and acidity, and a more rubbery and pasty textures.

Figure 1d shows selected paired differences with their TTB-derived 95% confidence ellipses. Ellipses for the T5-C paired difference (not shown; P=0.419) and the T6-C paired difference (P=0.063) overlap the origin. We are more confident in concluding that T1, T2, T3, T4, and T7 are perceived significantly differently than the control product since these ellipses were well separated from the origin.

Figure 1. Visualization of PCA of $\mathbf{X}$ and PCA of $\mathbf{X}\ominus\mathbf{X}$ results in the first A=2 PCs: (a) the loading plot for $\mathbf{P}_A$; (b) the score plot for $\mathbf{T}_A$; (c) the score plot for $\mathbf{T}_A\ominus\mathbf{T}_A$ showing only the test-reference paired differences; and (d) 95% confidence ellipses for selected paired differences. (Attribute codes: *maturity* [*m*], *rindy* [*r*], *sweet* [*w*], *bitter* [b], *acidic* [a], *salty* [s], *first-bite firmness* [*F*], *rubbery* [*U*], *pasty* [*P*], *bitty breakdown* [*B*], *dry* [*D*], *breakdown rate* [*R*].)

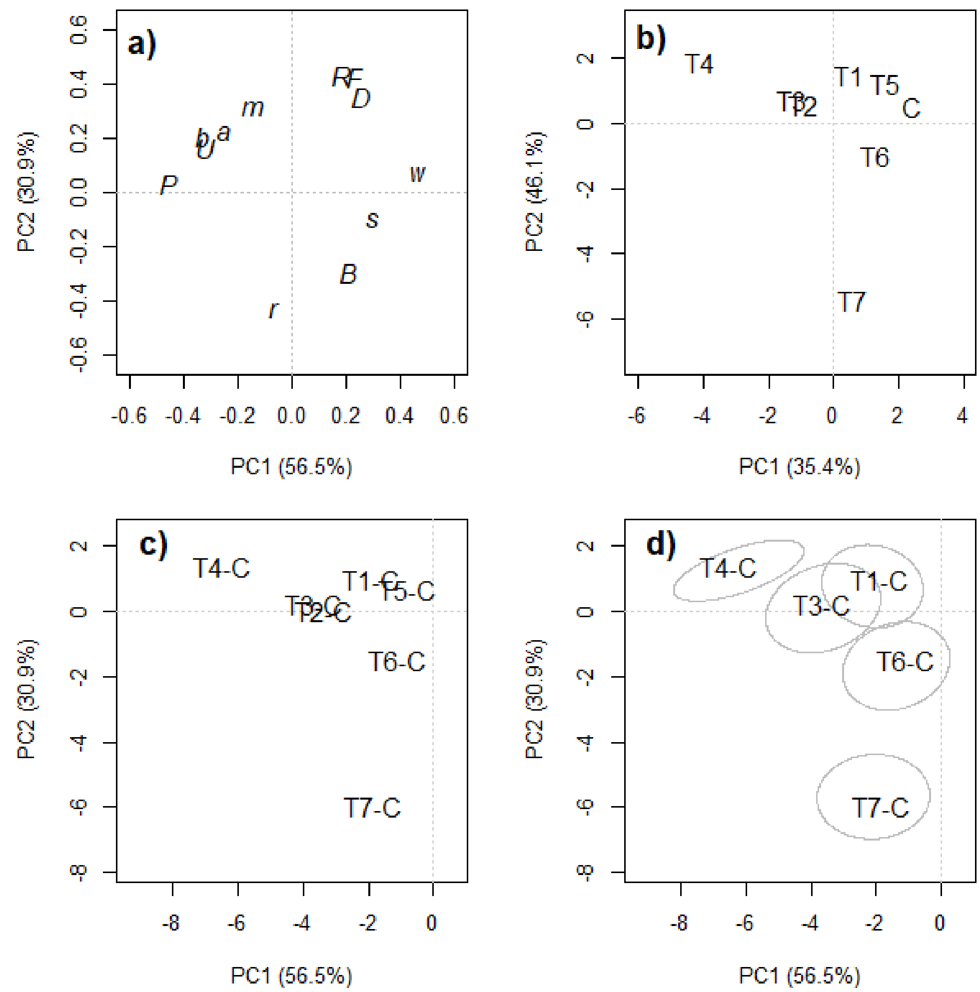


3.2 PCA of selected paired comparisons

This analysis focused on the C=7 test-reference paired comparisons that were of primary interest. Since each paired comparison was associated with twinned paired differences, the matrix $\mathbf{\Delta}^*$ (Section 1.3) contained 2C=14 rows. PCA of $\mathbf{\Delta}^*$ as in (3) extracted 56.5%, 30.9%, 8.4%, and 0.2% of the variance from $\mathbf{\Delta}^*$ in the first four PCs. To facilitate comparison with the solution in Section 3.1, we retained two PCs that together extracted 87.4% of the variance from $\mathbf{\Delta}^*$.

The loading coefficients ($\mathbf{P}_A$*) from PCA of $\mathbf{\Delta}^*$ are shown in Figure 2a. Compared to the loading coefficients in $\mathbf{P}_A$ (Figure 1a), loading coefficients in $\mathbf{P}_A$* (Figure 2a) appear to be rotated clockwise approximately 45 degrees and are more spread out.

Figure 2. Visualization of PCA of $\boldsymbol{\Delta}^*$ results in the first A=2 PCs: (a) the loading plot for $\mathbf{P}_A^*$; (b) the coordinates of products based on $\mathbf{XP}_A^*$ (axis labels indicate the percentage of variance extracted from $\mathbf{X}$ in the PCs of $\boldsymbol{\Delta}^*$); (c) the plot of $\mathbf{T}_A^*$ showing only the test-reference paired differences; and (d) 95% confidence ellipses for selected paired differences. (Attribute codes: *maturity* [*m*], *rindy* [*r*], *sweet* [*w*], *bitter* [b], *acidic* [a], *salty* [s], *first-bite firmness* [*F*], *rubbery* [*U*], *pasty* [*P*], *bitty breakdown* [*B*], *dry* [*D*], *breakdown rate* [*R*].)



Coordinates of the row objects in $\mathbf{X}$ in the PCs of $\boldsymbol{\Delta}^*$ were obtained by multiplying $\mathbf{X}$ by the loading matrix $\mathbf{P}_A^*$ as in (4). The configuration of products is shown in Figure 2b. As expected, the first two PCs of $\boldsymbol{\Delta}^*$ extracted a smaller percentage of variance (81.6%) from the row objects in $\mathbf{X}$ than was extracted in the first two PCs of $\mathbf{X}$ (84.7%). Here, we found that relative to the product configuration, C is in a more extreme position in Figure 2b than in Figure 1b. We advise against comparing interpreting test-test paired differences which are not included in $\boldsymbol{\Delta}^*$. Figure 2b is provided for completeness to avoid potential misinterpretations that might occur when inspecting the test-control paired differences, which are discussed next.

PCA of $\boldsymbol{\Delta}^*$ yields a score matrix of paired differences that is then truncated ($\mathbf{T}_A^*$). For the same reason as in Section 3.1, we plotted only the test-control paired differences in Figure 2c. If the test-control difference scores in $\mathbf{T}_A^*$ (Figure 2c) are rotated clockwise approximately 45 degrees, their configuration would be somewhat similar to the test-control difference scores in $\mathbf{T}_A \ominus \mathbf{T}_A$ (Figure

1c), as was the case for their respective loading matrices. The test-control paired differences and control-test paired differences intermingle in the first two PCs of $\mathbf{T}_A \ominus \mathbf{T}_A$ (Figure 1c), but there is better separation of the test-control and the control-test paired differences in $\mathbf{T}_A$* which respectively have negative and positive coordinates in PC1.

Figure 2d shows selected paired differences with their TTB-derived 95% confidence ellipses overlaid. Only the ellipse for the T5-C paired difference (not shown; P=0.392) overlaps the origin. Since ellipses from other paired differences were well separated from the origin, we conclude that other test products are perceived significantly differently than the control product. T6-C, which was not clearly separated from the origin in PCA of $\mathbf{X} \ominus \mathbf{X}$ (Figure 1d), is clearly separated from the origin in PCA of $\mathbf{\Delta}$* (Figure 2d).

Gain calculated via (5) quantifies the benefit of conducting PCA of $\mathbf{\Delta}$* instead of PCA of $\mathbf{X} \ominus \mathbf{X}$ when only the subset of paired comparisons in $\mathbf{\Delta}$* is of interest. Gain was 46.3% in one PC and 7.0% in the first two PCs.

4. Conclusions

Castura et al. (2023a) proposed an approach for investigating a relevant subset of paired comparisons after PCA. The procedure was applied to the cheddar cheese results. Results from PCA of all paired comparisons and results from PCA of only the relevant paired comparisons were shown. The Gain from using the proposed instead of the conventional approach was substantial (46.3%) if interpretation was limited to only one principal component. The Gain was more modest but still meaningfully large (7.0%) if interpretation focused on the plane of the first two principal components.

In the cheddar cheese data set, variability in all paired comparisons and in the relevant paired comparison were due to similar combinations of attributes. Consequently, Gain was not as large in the cheddar cheese data set as in data sets discussed by Castura et al. (2023a). A contributing factor was that the structure of the results matrix in both cases was relatively simple since most (>80%) of the variance was extracted in the first two PCs.

This manuscript focuses on an approach for investigating a relevant subset of paired comparisons after PCA. It is possible to miss important results if interpretation is based solely on visual inspection of plots, especially when there are many paired comparisons evaluated in more than two components. In such cases, P values can be used for screening purposes to draw attention to significant differences that might otherwise be missed.

The approach for investigating a subset of paired comparison after PCA could be further applied to explore multivariate results in other data sets in which only a subset of paired comparisons is of primary interest, including temporal sensory data sets and data sets in domains outside sensory evaluation.

Acknowledgements

Thanks to Marleen Chambault at Campden BRI (UK) for providing the cheddar cheese data set. Authors TN and PV acknowledge financial support from the Research Council of Norway and the Norwegian Fund for Research Fees for Agricultural Products (FFL) through the project FoodForFuture (Project number 314318; 2021-2024).

References

Cadoret, M., & Husson, F. (2013). Construction and evaluation of confidence ellipses applied at sensory data. *Food Quality and Preference*, 28, 106-115. https://doi.org/10.1016/j.foodqual.2012.09.005

Castura, J.C., Varela, P., & Næs, T. (2022). Investigating paired comparisons after principal component analysis. *Food Quality and Preference*, 106, 104814. https://doi.org/10.1016/j.foodqual.2023.104814

Castura, J.C., Varela, P., & Næs, T. (2023a). Investigating only a subset of paired comparisons after principal component analysis. *Food Quality and Preference*, 110, 104941. https://doi.org/10.1016/j.foodqual.2023.104941

Castura, J.C., Varela, P., & Næs, T. (2023b). Evaluation of complementary numerical and visual approaches for investigating pairwise comparisons after principal component analysis. *Food Quality and Preference*, 107, 104843. https://doi.org/10.1016/j.foodqual.2023.104843

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression*, 3rd ed. Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Mardia, K.V., Kent, J.T., & Bibby, J.M. (1979). *Multivariate Analysis*. Academic Press.

Næs, T., Tomic, O., Endrizzi, I., & Varela, P. (2021). Principal components analysis of descriptive sensory data: Reflections, challenges, and suggestions. *Journal of Sensory Studies*, 36, e12692. https://doi.org/10.1111/joss.12692

R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/