A sensory analysis on butter cookies -
An application of Generalized Procrustes Analysis

Hans Jørn Juhl
Project no 9
Executive Summary

1. A sensory analysis is one of the first steps in product development in the food industry. A thorough analysis of the results from such an analysis may give important input to the development process.

2. A sensory analysis on butter cookies is conducted in order to evaluate if some butter may be replaced by vegetable fat without a significant change in the sensory profile. The conclusion is that the replacement is possible without a considerable change in the sensory profile.

3. Generalized Procrustes Analysis is used to analyze the results. It is a relatively new technique adjusting for the fact that no two assessors are alike. The output is a representation of the products in a low-dimensional space and an evaluation of the differences between assessors. The technique may be used both when free-choice profiling and conventional profiling is applied. In this case we have used conventional profiling.
Table of Contents

1. Introduction ........................................................................................................ 1
2. Two different algorithms for Generalized Procrustes Analysis
   (GPA and GPPA) .................................................................................................... 1
2. A sensory analysis of butter cookies ................................................................. 3
4. Conclusion ........................................................................................................ 7
5. References ......................................................................................................... 7
1. Introduction

This paper presents the first empirical results from the QFooD project (Bech, Engelund, Juhl, Kristensen & Poulsen, 1994). QFooD is an acronym indicating that the purpose of the project is to apply QFD (Quality Function Deployment) in the food industry. A description of a food product in customer terms is obtained by a market analysis and a rich, more internal description of the food product is obtained from a sensory analysis. The intention is to apply these two data sets actively in the product development process in order to increase the rate of success of new food products.

The purpose of the study presented in this paper is to investigate the consequences of replacing butter with vegetable fat in cookie production. If the sensory profiles for a butter cookie and a cookie made of a mixture of vegetable fat and butter do not differ significantly, an inclusion of the mixture in the product portfolio may be considered, because cookies are cheaper to produce, if some of the butter is replaced by vegetable fat. Hence the focus in this article is on the results from a sensory analysis of cookies from one of the companies, Intergoods Bakery Ltd, cooperating with the QFood project. An eventual change in production has to be supported by results from a market research before any action is taken, but sensory analysis offers important input to the decision making process.

Generalized Procrustes Analysis is presented as an analytical tool especially well-suited for sensory analysis, but with a much broader possible field of application. The technique adjusts for the fact that no two assessors are alike. The adjustments regard 1) differences in the overall level of the scores, 2) differences in the interpretation of the attributes and 3) differences in their range of scoring. These three distortions are the ones most commonly mentioned in the literature.

2. Two different algorithms for Generalized Procrustes Analysis (GPA and GPPA)

We assume that there are K individuals each of whom provides information on the same N objects. Data consist of K matrices $X_k$ representing the scores for each of the products on the P sensory dimensions. Such a data matrix is called an individual configuration. The basic idea of GPA is to transform all K individual configurations at the same time so that each configuration matches all others as closely as possible. The internal relationships between the points of a configuration are not disturbed by any of the following operations, because relative distances are preserved: A translation of the whole configuration, a rotation or reflection of the whole configuration, a uniform dilation of the whole configuration. The idea of GPA in a least squares sense may be expressed in terms of a loss function (Borg & Lingoes, 1987):

$$L = \sum_{j<k}^{K} \text{tr}((\tilde{X}_j - \tilde{X}_k)'(\tilde{X}_j - \tilde{X}_k))$$
where $\tilde{X}_i = m_iX_iT_i - jt_i'$ and $T'_iT_i = I$. $L$ is to be minimized through choice of $K$ scale factors $m_i$, $K$ orthogonal matrices $T_i$ and $K$ translation vectors $t_i$. If $L = \text{min}$ then the most representative configuration can be defined to be the centroid configuration $Z$, whose coordinate matrix is the average of all $\tilde{X}_i$'s ie $Z = (1/K)\sum \tilde{x}_i$. Geometrically, each of $Z$'s points is the centroid of the corresponding points from the individual configurations. The matrix $Z$ is called the consensus configuration. No direct solution is known to this problem, but several iterative procedures based upon the criterion given in (1) are developed. The characteristic of these GPA methods is that they perform a symmetric analysis. This means that each judge must use the same number of attributes. Differentiating $L$ with respect to $t_i$, Gower (1975) shows that all configurations must be translated, so that their centroids correspond. The rotation/reflection problems and the dilation problem may be solved by an iterative procedure where the configurations are matched in a Procrustes sense to a common consensus configuration. From the consensus configuration two or three principal components are typically extracted.

Peay (1988) developed the GPPA method based upon the consensus criterion, which is aimed at maximizing the variance of the low-dimensional consensus-space obtained. Hence it identifies the maximal common structure of a given dimension of the $N$ sets of configurations. It may be expressed in the following way:

$$C = \sum_{i<k} \tilde{x}_i'\tilde{x}_i$$

Apart from the difference in criterion compared to Gower (1975), the algorithm developed by Peay (1988) allows for matrices of different column order. In a full dimensional structure the different criteria are equivalent and in practice they visually produce almost identical results (Dijksterhuis & Punter, 1990). The PROCRUSTES-PC programme used in the empirical part of this paper applies the consensus criterion.

In order to evaluate the results of the GPPA analysis a partition of the total variation of the full set of matrices ($X_i$) is made. When $X_{pk}$ denotes the matrix consisting of the first $p$ columns in $X_k$ and $\bar{X}_p$ represents the mean of the set of matrices $X_{pk}$, according to Peay (1988):

$$\sum \text{tr}(X_kX'_k) = N\text{tr}(\bar{X}_p\bar{X}_p') + \sum \text{tr}((X_{pk} - \bar{X}_p)(X_{pk} - \bar{X}_p')') + \sum (\text{tr}(X_kX_k') - \text{tr}(X_{pk}X_{pk}')) = V_{\text{consensus}} + V_{\text{within}} + V_{\text{projection}}$$

The first term on the right hand side represents the contribution of the total variation of the consensus configuration within the $p$-dimensional subspace. This is the variance to be maximized. The other two are residual terms. The second gives the residual variation within the $p$-dimensional subspace because the individual configurations are averaged. The third term represents the variation "lost" in the projection of the configuration. In PROCRUSTES-PC the partition is presented as percentages of the total variation (Dijksterhuis & Burden, 1990).

Interpretation of the dimensions in the resulting low-dimensional consensus configuration is based on average correlations between the original attributes and the $p$ principal components.
3. A sensory analysis of butter cookies

10 butter cookies were selected for the sensory analysis. A description of the butter cookies is given in the table below.

*Table 1. Butter cookies included in the analysis*

<table>
<thead>
<tr>
<th>Product</th>
<th>Shape</th>
<th>Taste</th>
<th>Fat</th>
<th>Butter %</th>
<th>Vegetable fat %</th>
<th>Sugar %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>star</td>
<td>vanilla</td>
<td>normal</td>
<td>28.1</td>
<td>0</td>
<td>22.5</td>
</tr>
<tr>
<td>2</td>
<td>star</td>
<td>vanilla</td>
<td>normal</td>
<td>0</td>
<td>27.6</td>
<td>22.6</td>
</tr>
<tr>
<td>3</td>
<td>star</td>
<td>vanilla</td>
<td>normal</td>
<td>14.1</td>
<td>13.7</td>
<td>22.6</td>
</tr>
<tr>
<td>4</td>
<td>star</td>
<td>vanilla</td>
<td>low</td>
<td>22</td>
<td>0</td>
<td>23.4</td>
</tr>
<tr>
<td>5</td>
<td>star</td>
<td>vanilla</td>
<td>low</td>
<td>0</td>
<td>21.8</td>
<td>23.2</td>
</tr>
<tr>
<td>6</td>
<td>star</td>
<td>vanilla</td>
<td>low</td>
<td>10.9</td>
<td>11</td>
<td>23.2</td>
</tr>
<tr>
<td>7</td>
<td>round</td>
<td>chocolate</td>
<td>low</td>
<td>0</td>
<td>23.7</td>
<td>24.1</td>
</tr>
<tr>
<td>8</td>
<td>round</td>
<td>chocolate</td>
<td>low</td>
<td>0</td>
<td>20.6</td>
<td>30.3</td>
</tr>
<tr>
<td>9</td>
<td>rectangular</td>
<td>sugar</td>
<td>high</td>
<td>22.8</td>
<td>0</td>
<td>22.8</td>
</tr>
<tr>
<td>10</td>
<td>star</td>
<td>fruit</td>
<td>high</td>
<td>28.02</td>
<td>0</td>
<td>22.6</td>
</tr>
</tbody>
</table>

Butter cookie numbers 1 to 6 are experimentally varied. Shape, taste and sugar are held constant. Fat percent and the mixture between vegetable fat and butter are varied. Products 7 and 8 are new products without any butter but with a different flavour than the other cookies. Products 9 and 10 are taken from the existing product line. 9 assessors were selected to take part in the sensory analysis carried out as a conventional profiling. Each product was evaluated 4 times by the same assessor on 25 attributes. Because of the training and discussion within the panel it is usual to assume that all assessors score the attributes in a similar manner. In order to test this assumption a MANOVA analysis was carried out using a nested design. For all 25 attributes the effect of Assessor (Repetition) was highly significant and for taste of vanilla, taste of sweetness and taste of bitterness the variation due to Assessor (Repetition) exceeded the variation due to Product. The results for vanilla and sweetness may be partially explained by the fact that the quantity of sugar and vanilla was held constant for 6 of the products in the study. Generally the results of the MANOVA analysis stress the problems caused by the use of human beings as reliable measuring instruments.
In the rest of the paper we use the mean scores for each assessor on each of the attributes. Some variation between mean scores may be due to 1) variation in the overall level of the scores the assessors give; 2) variations in the range of scoring and 3) use of different terms or combinations of terms to describe the same stimulus. The stages in GPPA can adjust for the above types of variation. Translation neutralizes effects due to 1). Dilation adjusts for 2) and rotation/reflection adjusts for 3). GPPA was applied to the individual configurations and as shown in Table 2 a two-dimensional configuration space explains 65.3% of the variance in individual mean scores.

Table 2. Percentage of explained total variance

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Consensus</th>
<th>Within</th>
<th>Total</th>
<th>Projection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41.367</td>
<td>1.748</td>
<td>43.115</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>23.934</td>
<td>1.090</td>
<td>25.023</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>65.300</td>
<td>2.838</td>
<td>68.138</td>
<td>31.862</td>
</tr>
</tbody>
</table>

The first principal component explains 41.4% of the total variation in mean scores and the within variance is only 2.8% indicating that the individually transformed configurations are very similar. The projection variance is equal to 31.8%. Hence 31.8% of the total variation in individual mean scores is due to the sources that GPPA adjusts for.

In order to investigate if the within variance can be traced back to a specific product, a partition of the within variance is made.

Table 3. Variance and products

<table>
<thead>
<tr>
<th>Product</th>
<th>Consensus</th>
<th>Within</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.678</td>
<td>0.230</td>
</tr>
<tr>
<td>2</td>
<td>1.398</td>
<td>0.170</td>
</tr>
<tr>
<td>3</td>
<td>5.845</td>
<td>0.535</td>
</tr>
<tr>
<td>4</td>
<td>4.093</td>
<td>0.115</td>
</tr>
<tr>
<td>5</td>
<td>12.684</td>
<td>0.253</td>
</tr>
<tr>
<td>6</td>
<td>6.312</td>
<td>0.352</td>
</tr>
<tr>
<td>7</td>
<td>10.811</td>
<td>0.253</td>
</tr>
<tr>
<td>8</td>
<td>10.288</td>
<td>0.230</td>
</tr>
<tr>
<td>9</td>
<td>0.011</td>
<td>0.306</td>
</tr>
<tr>
<td>10</td>
<td>3.179</td>
<td>0.394</td>
</tr>
</tbody>
</table>

Table 3 shows the percentages of the different variances associated with the 10 butter cookies. 44% of the total variance in the consensus space is represented by products 1, 5, 7 and 8. Product 9 has only 0.01% consensus space variance associated with it. The total variance in product
9 is lost as a consequence of projection. Hence product 9 will have coordinates in consensus space near the origin and have an interpretation as a reference product. In fact it is one of the most popular cookies in Denmark. The values in the column of within variances indicate that the reason for within variance can not be associated with any particular product. None of the products differ considerably in their coordinates from individual space to individual space.

A partition of the within variance as regards assessors may be used to evaluate if one or a few of the individual configurations differ considerably compared to the consensus space.

Table 4. Variance and assessors

<table>
<thead>
<tr>
<th>Assessor</th>
<th>Within</th>
<th>Scaling factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.219</td>
<td>1.080</td>
</tr>
<tr>
<td>2</td>
<td>0.355</td>
<td>1.464</td>
</tr>
<tr>
<td>3</td>
<td>0.605</td>
<td>0.880</td>
</tr>
<tr>
<td>4</td>
<td>0.164</td>
<td>1.020</td>
</tr>
<tr>
<td>5</td>
<td>0.361</td>
<td>0.987</td>
</tr>
<tr>
<td>6</td>
<td>0.229</td>
<td>1.156</td>
</tr>
<tr>
<td>7</td>
<td>0.352</td>
<td>0.927</td>
</tr>
<tr>
<td>8</td>
<td>0.119</td>
<td>0.971</td>
</tr>
<tr>
<td>9</td>
<td>0.433</td>
<td>0.822</td>
</tr>
</tbody>
</table>

Table 4 shows that assessor 3 represents nearly 25% of the within variance, but because of the general level of within variance, the number does not indicate that the assessor is unqualified as a judge. The scaling factors should be approximately 1 if the assessors use the scale in the same way. Assessor 2 has a scaling factor equal to 1.46 indicating that this judge has applied a more narrow range of the scale than the other assessors.

The consensus space is given below and has a relatively straightforward interpretation.

Three different clusters of products are identified. The first cluster consisting of products 7 and 8 are products solely with vegetable fat and a flavour additive. Products 1, 3 and 10 have a high content of fat. Products 1 and 10 are butter cookies and product 3 is an interesting mixture of butter and vegetable fat. As long as the content of fat is high, the panel is not able to identify considerable sensory differences between 1, 3 and 10. The results indicate an interaction effect between the level of fat and the mixture of butter and vegetable fat. Products 2, 4, 5 and 6 form the last cluster. Products 4, 5 and 6 are all products with a low level of fat. Products 2 and 4 are close, stressing the possible interaction effect mentioned above. In general, the map shows that it is possible to develop products based upon mixtures of fat that resemble cookies made with butter on sensory attributes. From an economical point of view, this conclusion may have an important implication for product development. The cost price for product 3 is roughly 15% below the cost price for product 1 and 10.
The two dimensions in the consensus space are interpreted on the basis of average correlation coefficients between the 10 products’ scores on the principal components and their scores on the 25 attributes. Only correlations exceeding an absolute value of 0.65 are taken into account. Comparing the coordinates of the products with the average correlations leads to the fol-

**Table 5. Correlations and attributes**

<table>
<thead>
<tr>
<th>Sensory attribute</th>
<th>Dimension 1</th>
<th>Dimension 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height</td>
<td>-0.78</td>
<td>-0.17</td>
</tr>
<tr>
<td>Hard-bite</td>
<td>-0.84</td>
<td>0.10</td>
</tr>
<tr>
<td>Seed</td>
<td>-0.80</td>
<td>0.02</td>
</tr>
<tr>
<td>Crispness</td>
<td>0.90</td>
<td>-0.02</td>
</tr>
<tr>
<td>Melting time</td>
<td>-0.80</td>
<td>-0.16</td>
</tr>
<tr>
<td>Dry</td>
<td>-0.85</td>
<td>0.10</td>
</tr>
<tr>
<td>After-taste</td>
<td>0.87</td>
<td>-0.03</td>
</tr>
<tr>
<td>Yellow</td>
<td>0.36</td>
<td>0.68</td>
</tr>
<tr>
<td>Crumbling</td>
<td>-0.25</td>
<td>-0.73</td>
</tr>
<tr>
<td>Coconut flavour</td>
<td>0.40</td>
<td>-0.75</td>
</tr>
<tr>
<td>Chocolate flavour</td>
<td>0.41</td>
<td>-0.84</td>
</tr>
</tbody>
</table>
lowing conclusions. Fat content is in good accordance with dimension 1 and hence products with normal fat content seem to be more crispy and have more after-taste. They are smaller, easier to bite, have less melting time, are less dry and without a tendency to leave the inner part of the cookie unbaked (kernel). As regards the correlations with dimension 2 the flavour attributes are not surprising. The second implication is that cookies without butter are considered less yellow than cookies with butter. The cookies made solely with vegetable fat have a tendency to crumble more than cookies made out of butter.

4. Conclusion

The empirical part of the paper has demonstrated that a sensory analysis may give valuable information to the product development process. In this case it has been shown that the sensory profile of a butter cookie does not differ considerably from the one obtained by a cookie made out of a mixture of butter and vegetable fat. If preference analysis based upon consumer data give similar results, including mixtures in the production portfolio may be considered.

Generalized Procrustes Analysis has been presented and it is evident that this technique gives a thorough analysis of sensory data. A market analysis consisting of consumer evaluations of products on different attributes may be analyzed in a similar way. The procrustes analysis may also be a good choice in an attempt to identify the structure in consumer evaluations. The main difference compared to sensory analysis is that consumers deviating from the consensus space will not be seen as a problem but as an opportunity to identify segments in the market.

5. References

MAPP publications

MAPP working papers


No. 9: Bonke, J. *Choice of foods - allocation of time and money, household production and market services*, PART II, September 1993.


MAPP conference papers


MAPP reprints


Furthermore there are a number of project papers, which are not available to the public.
The Mapp programme consists of the following 15 projects

1. Strategic Planning and Innovation Capability in the Danish Food Sector
   Morten Kvistgaard & Kirsten Plichta, Copenhagen Business School; Lone Rossen, Biotechnological Institute

2. Innovation Capability as a Key Success Factor
   Klaus G. Grunert & Hanne Harmsen, The Aarhus School of Business

3. Quality Certification as a Key Success Factor in International Marketing of Food Products
   Niels Jørgensen & Erik Lund, Business University of South Jutland

4. Definition of the Sales Potential for a New Food Product to be Launched on Home or Foreign Markets
   Anne Martensen & Kenneth Kæregaard, Copenhagen Business School

5. Primary Producers and Product Innovation in the Food Industry
   Villy Søgaard, University Centre of South Jutland

6. Controlling Processes of Production to Guarantee Process Characteristics Demanded by Consumers of Food Products: Paradigms and Danish Experiences
   Esben Sloth-Andersen, Aalborg University Centre

7. The Role of the Distribution System in Product Innovation
   Hanne Hartvig Larsen & Nick Norman Jensen, Copenhagen Business School

8. Prototyping in the Danish Food Industry
   Preben Sander Kristensen & Elsebeth Holmen, Aalborg University Centre

9. Product Quality and Consumer Preferences: Assessing the Optimum Design of Food Products
   Kai Kristensen, Hans Jørn Juhl, Anne Bech & Erling Engelund, The Aarhus School of Business; Carsten Stig Poulsen, Aalborg University Centre

10. Product Innovation and Packaging in the Food Industry - Environmental Consequences and Consumer Reactions
    John Thøgersen & Tino Bech-Larsen, The Aarhus School of Business

11. The Consumer as Agent in Relation to Research and Development in Food Technology
    Erling Jelsøe, Birgit Land & Jesper Lassen, Roskilde University Centre

12. Households’ Choice of Foodstuffs with Different Kinds of Preparation
    Jens Bonke, University of Copenhagen

13. The Cultural Dimensions of Food Consumption and the Implications for Strategy Formation and Implementation in Small and Medium-sized Danish Companies
    Dominique Bouchet, Josette Andersen, Søren Askegaard, Tage Koed Madsen & Per Østergaard, Odense University

14. Market Surveillance Systems for the Food Sector
    Klaus G. Grunert & Karen Brunsø, The Aarhus School of Business

15. Identification of Key Success Factors
    Klaus G. Grunert & Elin Sørensen, The Aarhus School of Business