Multivariate Approaches for Relating Consumer Preference to Sensory Characteristics

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By
Rachel Esther Liggett, M.S.
Graduate Program in Food Science & Nutrition

The Ohio State University
2010

Dissertation Committee:
Jeannine F. Delwiche, Co-advisor
Michael Mangino, Co-advisor
W. James Harper
M. Monica Giusti
Abstract

Preference mapping refers to a category of statistical methods used to relate consumer acceptance to a product’s characteristics (often measured by sensory descriptive analysis). Several techniques for creating preference maps exist, and they vary in the manner by which these two types of data are related. Techniques are generally classified into three main categories: external, internal, and hybrid. For external preference mapping, the sensory perceptual space is set by descriptive sensory data and consumer preference information is subsequently overlaid onto this sensory space. For internal preference mapping, the consumers’ preference for the products (usually overall liking ratings) is used to create a liking space upon which sensory descriptive ratings are subsequently mapped. Hybrid techniques locate products on the map using both consumer and sensory descriptive data simultaneously onto a restricted liking space.

Preference mapping studies consume a considerable amount of time and resources. When such investigations are unsuccessful, it is highly desirable to determine the root cause and avoid the same issue in the future. Preference mapping studies are necessarily complex and error can be introduced at any of several steps along the way, such as product selection, descriptive analysis, consumer testing, data analysis and interpretation of the outputs. The purpose of this investigation was to examine these
issues and develop best practices for conducting preference mapping studies. One popular technique from each category was selected for investigation using a common dataset.

Fifteen commercially available Swiss-type cheeses (10 domestic Swiss cheeses, 4 Baby Swiss cheeses, and 1 imported Swiss Emmenthal) were evaluated by twelve trained panelists using the Spectrum method. Significant differences between the cheeses were exhibited for 15 flavor attributes. The same 15 cheeses were also evaluated for overall liking by 101 untrained consumers (53 female; ages 18-65). Significant differences in liking of the cheeses were also found.

External preference mapping (EPM) was able to fit some consumers but generally performed poorly. The resulting preference map explained only 42% of variability. The region of highest appreciation for products was quite broad and determination of an optimum profile was impossible. Landscape Segmentation Analysis® (LSA), a version of internal preference mapping, explained 90% of variability once two outlying products were removed. An optimal profile and 7 liking drivers were obtained. Partial Least Squares (PLS) Regression, a hybrid preference mapping technique, explained 93% of variability after 1 outlying product was removed and an optimal profile and 7 liking drivers were obtained similar to LSA.

PLS and LSA offered advantages over EPM in understanding consumer liking, seemingly because the product space was set at least in part using consumer information. Optimal profiles and liking drivers for PLS and LSA were strikingly similar, suggesting both are successful and valid preference mapping techniques. Even though PLS requires pre-treatment of the data to identify possible consumer segments, LSA does not. However, PLS offers a slight advantage over LSA because its output is easier to interpret.
Research was funded through financial and material gifts from the Swiss Cheese Consortium, Ohio Agricultural Research & Development Center, and USDA Cooperative State Research, Education, and Extension Service.
Dedication

To me
Acknowledgements

This project was initiated through financial and material gifts from the Swiss Cheese Consortium. Financial support was also obtained through an Ohio Agricultural Research & Development Center Small Industry Matching Grant (number 2003-201) and the United States Department of Agriculture Cooperative State Research, Education, and Extension Service Special Research Grant (number 25-6231-0085-003).

I owe a debt of gratitude...

To Jeannine Delwiche, my advisor and my friend, for guiding my research (and knocking this thing out in record time!)

To MaryAnne Drake, my co-author, for collecting the descriptive data

To my committee for helping me make this happen (specifically Dr. Mangino and Dr. Harper for telling me it probably wouldn’t happen after I went to industry, which really just made me try harder; and Dr. Giusti for joining late in the game)

To Gregg, my husband, for making me food while I was holed up in my “office” writing (and rewriting) this thing

To my FOCUS colleagues, Jason Parcon, Benoit Clavier, Guillaume Blancher, and Carine Egoroff, for assisting me in developing the concepts

To my lab-mates for spending hours cutting cheese

To my panelists for participating
Vita

August 31, 1973.................................Born – Marion, OH

May 1991...........................................North Union High School, Richwood, OH

1995..............................................A.A., The Ohio State University, Marion

1997.............................................B.S. Food Science & Technology, The Ohio State University, Columbus

1997-1998.................................Sensory Associate, Nestle Research & Development Center, Inc., Marysville, OH

1997-1998.................................Sensory Scientist, Kerry Inc., Beloit, WI

2000-2006.................................Sensory Program Coordinator, Sensory Science Group, Department of Food Science & Technology, The Ohio State University, Columbus

2001...........................................Assistant Coach, Sensory Evaluation of Dairy Product Judging Team, Department of Food Science & Technology, The Ohio State University, Columbus
2003………………………………………………First place winner of the Rose Marie Pangborn Memorial Graduate Student Research Paper Competition sponsored by the Sensory Evaluation Division at the Institute of Food Technologists Annual Meeting, Chicago, IL

2003-2006………………………………………………Co-leader, “Tasting Science” Extension Program, Sensory Science Group, Department of Food Science & Technology, The Ohio State University, Columbus

2004…………………………………………………………M.S. Food Science & Nutrition, The Ohio State University, Columbus

2004-2006………………………………………………………Supporting Advisor, Food Science & Technology Club, Department of Food Science & Technology, The Ohio State University, Columbus

2006-2008………………………………………………………Project Manager, Sensory Science, Givaudan Flavors Corporation, Cincinnati, OH

2008-present…………………………………………………Senior Project Manager, Global Sensory, Givaudan Flavors Corporation, Cincinnati, OH

2009-present…………………………………………………Adjunct Professor, Chemical Technology Department, University of Cincinnati, OH
Publications


Field of Study

Major Field: Food Science & Nutrition
# Table of Contents

Abstract .......................................................................................................................... ii

Dedication......................................................................................................................... v

Acknowledgements........................................................................................................ vi

Vita ................................................................................................................................... vii

Table of Contents............................................................................................................ xi

List of Tables .................................................................................................................. xiii

List of Figures ................................................................................................................ xiv

Chapter 1: Introduction................................................................................................. 1

  What is preference mapping? ...................................................................................... 1

  Common guidelines for preference mapping studies................................................... 1

  How is preference mapping applied? ........................................................................... 4

  Major categories of preference models to investigate ................................................. 7

  Investigations comparing preference mapping methods............................................. 17

  Pre-processing data before preference mapping ........................................................ 20

  Considerations for choosing the best model............................................................... 23

  The purpose of this research investigation................................................................. 25

Chapter 2: Materials and Methods ............................................................................. 27

  Overview....................................................................................................................... 27

  Materials....................................................................................................................... 27

  Descriptive Analysis.................................................................................................... 28
List of Tables

Table 2.1. Lexicon used to profile Swiss cheeses.........................................................29
Table 4.1. Mean descriptive analysis intensity ratings for Swiss cheeses.....................37
Table 4.2. Mean consumer liking ratings for Swiss cheeses.........................................41
Table 4.3. Mean consumer liking ratings for two segments of consumers....................57
Table 4.4. Comparison of drivers determined by EPM, LSA and PLS...........................65
Table 5.1. Recommendations toward successful preference mapping studies.............71
# List of Figures

Figure 4.1. PCA of descriptive analysis for Swiss cheeses, PC2 vs PC1 ......................39
Figure 4.2. PCA of descriptive Analysis for Swiss cheeses, PC3 vs PC1 .....................40
Figure 4.3. EPM contour map of Swiss cheeses, PC2 vs PC1 ......................................43
Figure 4.4. EPM contour map of Swiss cheeses, PC3 vs PC1 ......................................44
Figure 4.5. EPM trace plots of each attribute vs. liking for Swiss cheeses .................45-47
Figure 4.6. LSA contour map of Swiss cheeses ............................................................50
Figure 4.7. LSA contour map of Swiss cheeses, excluding S09 & S07 .........................52
Figure 4.8. Optimal profiles for two potential consumer segments identified by LSA ...53
Figure 4.9. HCA of consumers clustered by individual liking of Swiss cheeses ...........55
Figure 4.10. LSA contour map of Swiss cheeses (excluding S09 & S07) for differentiating consumers ...........................................................................................58
Figure 4.11. PLS map of Swiss cheese for PLS2 vs. PLS1 ...........................................59
Figure 4.12. Optimal profile of Swiss cheeses identified by PLS model derived with all products ..................................................................................................................60
Figure 4.13. PLS of Swiss cheeses (excluding B1) for differentiating consumers, PLS2 vs. PLS1 .....................................................................................................................62
Figure 4.14. Optimal profile of Swiss cheeses for differentiating consumers identified by PLS model, excluding cheese B1 and attributes Bitter and Metallic .................63
Figure 4.15. Comparison of optimal profiles determined by LSA and PLS ..................66
Chapter 1: Introduction

What is preference mapping?

Preference mapping refers to a category of statistical methods used to relate consumer acceptance to a product’s characteristics (Meilgaard, Civille, & Carr, 2007). The output is generally displayed graphically in order to visually assess the relationship between the product space and patterns of preference (Lawless & Heymann, 1999). Commonly, a product’s characteristics are measured by instrumental analysis and/or sensory descriptive analysis by a trained panel in order to gain insight into key attributes which affect liking. Consumers are not questioned directly to find these attributes because it is a challenge for consumers to describe these characteristics themselves. The information can be used to guide product development by revealing where a new product would fit in the space or how to reformulate existing products to increase consumer appeal (MacFie, 2007; van Kleef, van Trijp, & Luning, 2006). Using the model behind a preference map, it is also possible to predict an optimal profile of a product (Meullenet, Xiong, & Findlay, 2007).

Common guidelines for preference mapping studies

For the most effective preference mapping, certain guidelines should be followed. First, since drivers of consumer acceptance are determined based on the product set under investigation, it is very important to build a stable sensory space by selection of products in the set. This can be done with marketed consumer products as well as
prototypes. Without a representative product space (i.e., one that encompasses the entire product space), results will be difficult both to interpret and utilize in product development. Ideally, twelve or more products should be tested, but in practice it is generally more feasible to use eight. Six carefully selected products is the absolute minimum number of products that can be used to construct a stable sensory space (MacFie, 2007).

Sales, marketing and other sources of information are useful inputs for choosing products to go into a consumer test. However, descriptive data are the most important criteria to consider since products must be chosen strategically based upon their sensory profiles to ensure the stability of the sensory space. A certain amount of variation in profiles is needed because if products are too similar, results will not provide any insight. Ideally, a sensory map is developed with as many relevant products as possible. From this information, a subset of the products can recreate the sensory space for consumer testing (MacFie, 2007). In other words, the sensory space should not be restricted by the number of samples it is realistic to use in consumer testing, nor should it be limited by market interests or existing products.

As with all statistical techniques, certain assumptions must be met in order for their conclusions to be valid. For some preference mapping techniques a repeated measures design is necessary, in which every respondent evaluates every product. Regardless of the technique used, a repeated measures design is preferable (MacFie, 2007). In cases where such a design is not feasible, evidence indicates respondents should evaluate at least 75% of the products (Hedderly & Wakeling, 1995). Specifically, this recommendation was determined through imputation of missing values wherein various techniques were employed to generate values to substitute for missing data. Techniques ranged from simply substituting the missing value with the sample mean to
complex iterative calculations. Using each set of values, corresponding preference maps were assessed to determine which was acceptably close to the preference map generated from a complete dataset. Results indicated not only that respondents should evaluate at least 75% of the products but that imputation of missing values by mean substitution was the most successful (Hedderly & Wakeling, 1995).

The number of consumers to test is another major consideration when designing a preference mapping study. Values found in the literature vary widely with studies as low as 40 subjects (Bovell-Benjamin, Allen, & Guinard, 1999) and as high as 480 subjects (Font I Furnols, Gispert, Diestre, & Oliver, 2003) or more. However, these are the extremes and reasons for testing the given number of consumers in an acceptability test are often not reported. Parameters such as alpha-level, beta-level, standard error of experiment and sought after difference in means have commonly been used in power calculations for estimating sample size (Kraemer & Thiemann, 1987). In a research study evaluating 108 consumer studies executed in 5 countries, standard error of experiment was found to be similar across all studies. Thus using common levels of these parameters, a power calculation suggests a minimum sample size of 112. However, this minimum would not be large enough to study population sub-groups (Hough et al., 2006), as it is desirable to maintain sub-groups of at least 50 consumers for analysis (MacFie, 2007).

In central location testing, the impact of carry-over effects must also be considered. The common tendency is for (untrained) consumers to react differently to the first sample because they are unsure of how to use the scale (MacFie, 2007). While counterbalanced designs such as Williams Latin squares (Williams, 1949), are meant to minimize first order carry-over effects by even distribution, the effect is still present. One possible solution to reduce or remove this first order effect would be to present
consumers with a dummy sample (often inaccurately referred to as a warm up sample). Specifically, a sample is presented in the first position that consumers believe to be part of the study and is later discarded prior to data analysis. In studies where consumers participate in multiple sessions, it is recommended there be a dummy sample for each session (MacFie, 2007). Challenges to this solution include the cost associated with testing additional samples or collecting less information by “giving up” slots for test samples to include dummy samples. In many cases, researchers are not willing to make this trade-off. When dummy samples are not included and serving order effects are seen, mathematical adjustments can be made to counteract this effect, particularly in studies with six or fewer samples (Hottenstein, Taylor, & Carr, 2008).

**How is preference mapping applied?**

Many studies on preference mapping have been published and undoubtedly many more have gone unpublished as many such investigations are proprietary. While preference mapping can be done with any consumer product, the focus of this investigation will be on food. Preference mapping has been applied to a wide variety of food products, ranging from natural (e.g., apples and kiwifruit) to processed (e.g., ice cream and sausage), and marketed (e.g., cereal bars and beer) to experimental (e.g., pudding and ranch dressing) (see Bower & Baxter, 2003; Bower & Whitten, 2000; Elmore, Heymann, Johnson, & Hewett, 1999; Guinard, Uotani, & Schlich, 2001; Harker, Kupferman, Marin, Gunson, & Triggs, 2008; Helgesen, Solheim, & Naes, 1997; Hough & Sanchez, 1998; Jaeger, Rossiter, Wismer, & Harker, 2003; Yackinous, Wee, & Guinard, 1999).

An experimenter can exert considerable control over the products under investigation. While in most instances one selects products that encompass the sensory
space, some investigations focus on specific characteristics of products, the sensory space is narrowed, and instead these characteristics are manipulated. For example, in a study on puddings, experimenters formulated samples to vary on three dimensions (thickness, fat content and smoothness) by varying the ingredients while attempting to keep other characteristics (such as color) constant (Elmore et al., 1999). Such sample manipulation can systematically follow a pre-selected experimental design, as in Yackinous (1999), where experimenters followed a 3 x 3 factorial design, systematically manipulating ranch salad dressings with three levels of fat and three levels of garlic and ultimately testing nine systematically varied samples.

Most commonly, overall liking is linked with sensory descriptive information (for example, see Meullenet et al., 2001; Pagliarni, Monteleone, & Wakeling, 1997; N. D. Young, Sanders, Drake, Osborne, & Civille, 2005). In terms of modality, flavor was considered in nearly all studies but appearance (e.g., see Geel, Kinnear, & de Kock, 2005; Hough & Sanchez, 1998), aroma (e.g., Martinez, Santa Cruz, Hough, & Vega, 2002; Sveinsdottir et al., 2009), texture (e.g., Meullenet et al., 2003; Pagliarni et al., 1997), and mouthfeel (e.g., Elmore et al., 1999; Richardson-Harman et al., 2000) have often been reported as well. Several studies have also linked overall liking to instrumental measures (e.g., Alves et al., 2008; Ares, Giminez, & Gambaro, 2006; Berna, Lammertyn, Buysens, Di Natale, & Nicolai, 2005; Bower & Whitten, 2000; Geel et al., 2005; Harker et al., 2008; Hough & Sanchez, 1998; Pham et al., 2008). It is also possible to extend preference mapping beyond overall liking to liking of a specific characteristic such as texture (e.g., Ares et al., 2006; Elmore et al., 1999).

Demographic and behavioral data has also been linked to preference by mapping. Thybo, Kuhn & Martens (2003) investigated children’s preferences for apples. They modeled not only sensory, chemical and instrumental measures, but also
demographic information (gender, age, etc.), behavioral data (for example, frequency of apple consumption), and appearance and taste preferences (5-pt facial hedonic scale). By mapping appearance preference along with appearance-related behavioral data and objective measures, inter-disciplinary relationships could be identified and utilized to provide insights explaining varying preferences among the children.

It is also possible to link liking to an attribute’s temporal profile. When investigating strawberry jams, Alves et al. (2008) had judges assess strawberry flavor, sourness and sweetness by time-intensity. Maximum intensity, time to maximum intensity, total duration, and area under the curve showed significant differences between jams and these measures were related to consumer liking. In this instance intermediate values for all measures were associated with the most liked samples (Alves et al., 2008).

Product acceptance is dependent on more than just a product’s flavor characteristics. A product’s packaging is important to purchase and can also be linked to liking. In a study of Cheddar cheese, descriptive panelists were trained to evaluate twenty package characteristics, including shape and performance, in addition to the sensory characteristics of appearance, aroma, flavor and texture. Consumers were asked to rate liking by both taste and packaging. Resulting preference models revealed sensory and packaging characteristics that were most appealing to consumers (Delahunty & Murray, 2000).

Although instrumental measures alone can be used for preference mapping, such approaches are at risk for missing relevant and measurable sensory attributes. For example, to predict consumer liking of texture, the texture of dulce de leche, both sensory and instrumental measures were needed (Ares et al., 2006). In this study, consumers assessed overall liking of texture, a trained sensory panel evaluated texture
attributes, and instrumental measurements were made. While certain instrumental measures were correlated to overall liking, they did not tell the whole story. For example, instrumental measures and sensory ratings for hardness were highly correlated and certain instrumental measures for hardness could be used to predict liking as well as the sensory rating of hardness. On the other hand, ropiness was not correlated to any of the instrumental measures (Ares et al., 2006).

However, it is important to note that even when consumers are asked to focus on a particular set of characteristics, other unrelated attributes can impact their assessments. For example, Elmore et al. (1999) asked consumers to rate their liking of creamy texture for pudding samples of varying thickness, fat content and smoothness. Descriptive analysis of appearance, texture, mouthfeel and flavor was also conducted for these samples. Even though consumers were directed to rate their liking of texture, preference mapping indicated that flavor attributes impacted the preferences (Elmore et al., 1999).

**Major categories of preference models to investigate**

Two preference mapping categories were originally proposed by Carroll (1972) delineating two possible treatments of the same data to create preference maps. These two main approaches are known as internal and external. In preference mapping, internal and external refer to how the products are located on the map. For internal preference mapping, the consumers’ preference for the products (usually overall liking ratings) is used to set the space, that is to say, consumers are internal to the map. While sensory descriptive or instrumental data is not needed to create the map, it may be projected onto the map in a later step to aid in interpretation. For external preference mapping, the sensory perceptual space is set by sensory descriptive or instrumental
data. Consumer preference information is overlaid on the map in a later step, that is to say, consumers are external to the map (MacFie, 2007; van Kleef et al., 2006).

In either analysis, product position on the map is driven by levels of the attributes that show the greatest variation. In the case of internal preference mapping, distances between products are due to variation in liking. The maximum amount of variation in liking is expressed on the first dimension. In contrast, external preference analyses set distances between products based on similarity in perception, that is to say, the objective sensory characteristics by which the products were measured. The maximum amount of variation in perception is expressed on the first dimension. Since the basis for constructing the maps is different, the highlighted information may differ and this has lead to debate on which method is “better.” A more useful distinction is under what circumstances each is more appropriate for the experimental goals (van Kleef et al., 2006).

In addition to these two main categories, a hybrid technique, partial least squares regression, locates products on the map using both consumer and sensory data (Ares et al., 2006; Meilgaard et al., 2007; Tenenhaus, Pages, Ambroisine, & Guinot, 2005). Some researchers suggest that partial least squares regression is an external preference mapping technique. The rationale is that linear combinations of the sensory data are extracted in order to explain the consumer data (Thybo et al., 2003). In this review, however, partial least squares regression will be treated as a hybrid technique because location of products in the resulting preference maps is based on variability of sensory data that directly explains the variability in consumer data. In other words, the only space considered is that space which is relevant to consumer liking - the consumer
assessments set the borders of this sensory space. Thus consumers are not purely external to the map (Helgesen et al., 1997; Meilgaard et al., 2007; Tenenhaus et al., 2005).

**Internal Preference Mapping**

Internal preference mapping (IPM) is a technique which utilizes mainly consumer data to investigate differences in preference between consumers (Carroll, 1972; Helgesen et al., 1997; MacFie & Thomson, 1988; Schiffman, Reynolds, & Young, 1981; Schlich, 1995). The basis of IPM is principal component analysis (PCA) using individual consumer ratings of products to extract new variables which correspond to the major underlying dimensions of preference. Overall liking is almost invariably correlated to the first principal component (D.M. Ennis, Seeing the Market through the Eyes of the Consumer, Institute for Perception Short Course, November 4, 2009). To conduct this analysis, no analytical information on the products is needed as the product space is constructed purely based on consumer hedonic scores (Schlich, 1995; van Kleef et al., 2006).

In perhaps its earliest form, IPM was conducted using a program called MDPREF (Chang & Carroll, 1969) but it is now widely available in various software packages (Meullenet et al., 2007). In practice, consumers are treated as variables in a PCA. For data pretreatment, one recommendation is mean centering and scaling to unit variance for each consumer (Greenhoff & MacFie, 1994). This effectively gives equal influence to all consumers even if some have low variance and do not seem to express preference for any of the products. This is the case when performing the analysis using the correlation matrix. However, one could argue that since consumers with low variance are not expressing preference for any of the products, they should not be given equal
influence in the analysis (Meullenet et al., 2007). In this case, PCA by covariance matrix
would be suitable since individuals are not scaled to unit variance (Borgognone, Bussi, &
Hough, 2001). In either case, the resulting PCA map contains a vector for each
consumer pointed in the direction of preference (Meullenet et al., 2007).

There are advantages associated with IPM. Since analyses are conducted on
individual ratings of consumers, heterogeneity of preferences may be seen. It is possible
to visually inspect resulting preference maps and observe segmentation of consumers
by examining the concentration of consumers across regions of the map (Guinard et al.,
2001; Helgesen et al., 1997; Meullenet et al., 2007; van Kleef et al., 2006). Another
advantage of constructing the preference map from consumer liking information alone is
that it can reveal the underlying basis by which consumers assign liking instead of
forcing consumers to fit the perceptual space (Ennis & Anderson, 2003). However,
mapping on consumer liking alone can confound preferential differences with perceptual
differences (Green & Rao, 1972). In other words, it is possible for two perceptually
distinct products to overlap in the liking space, which would make it difficult to map
apparently conflicting descriptive attributes onto the liking space.

Additional disadvantages for IPM exist. When perceptual (that is to say, sensory)
information on the products of interest is unavailable, it can be very difficult to interpret
the dimensions underlying preference. However, in practice, this information is typically
available (van Kleef et al., 2006). A bigger problem associated with IPM is that it is
difficult to determine optimal levels for the identified key attributes in product profiles.
While IPM determines the attributes which drive liking in a given product space (when
perceptual information is available), the upper and lower bounds or optimal levels of the
attributes are much less evident (Meullenet et al., 2007).
**Landscape Segmentation Analysis.** A newer IPM technique is Landscape Segmentation Analysis® (LSA). In comparison to traditional preference mapping techniques, little has been published on LSA in peer-reviewed journals. Only a limited amount of information has been published by the Institute for Perception (Richmond, VA, USA) in their periodic newsletter and the manual for their proprietary software, IFPrograms™. While LSA constructs the preference map solely from consumer liking data, the main difference is in its probabilistic approach. The IPM technique described above (as well as external and hybrid techniques to be discussed later) is classified as a deterministic approach. A deterministic model assumes liking is absolute and a consumer’s liking of a given product is constant over time. Probabilistic approaches do not make these assumptions and instead assume that liking ratings are relative to an individual ideal point (Meullenet et al., 2007). That is to say, each consumer imagines what his or her ideal product would be and then rates liking of the product under evaluation in comparison to that ideal. In general, a high liking score given to a product indicates the product is closer to the imagined ideal and a low liking score indicates the product is further from the imagined ideal. Probabilistic approaches also work under the assumption that a product’s distance from the ideal varies over time due to a multitude of reasons which may or may not be related to the product itself (Ennis & Bi, 1998; Meullenet et al., 2007). Using this probabilistic approach, variability associated with preferential choice can explain inconsistent consumer behavior (Ennis, 1999).

LSA works by converting hedonic scores into “similarity to ideal” scores (Meullenet et al., 2007). These scores are then “unfolded” using a special case of multidimensional scaling where evaluated products and each consumer’s ideal product locations are estimated on a map (Ennis & Bi, 1998; MacKay, 2001). This is done through an iterative process where product locations and variances are varied until the
best fit is achieved (Meullenet et al., 2007). Unlike a deterministic internal preference map, this map is considered to be a sensory or perceptual space as it is created by similarities rather than hedonic scores (D.M. Ennis, Seeing the Market through the Eyes of the Consumer, Institute for Perception Short Course, November 4, 2009). In this space, large distances between evaluated and ideal products would indicate minimal liking of the evaluated products and vice versa. Where a deterministic map would indicate direction of an ideal product, a probabilistic map physically locates ideal products in the same space as real products (Meullenet et al., 2007). By examining the density of these individual ideal points, segments of consumers can be identified. Descriptive information can then be projected on this probabilistic liking space in order to interpret its underlying dimensions and understand the product characteristics appreciated by the different segments (Ennis, 2001). Looking at the contours on the map it is quite easy to see where an optimal product would be located and obtain its sensory profile (Meullenet et al., 2007).

**External Preference Mapping**

External preference mapping (EPM) is a technique which utilizes mainly consumer and descriptive data to investigate the sensory characteristics underlying consumer preference (Carroll, 1972; Helgesen et al., 1997; MacFie & Thomson, 1988; Schiffman et al., 1981; Schlich, 1995). The basis of EPM is principal component analysis using sensory descriptive data (or some other analytical measure of the products) to extract components which describe the underlying dimensions of the product attributes. Consumer hedonic scores are then regressed onto the perceptual map to help understand which attributes drive consumer preference (Helgesen et al., 1997; MacFie & Thomson, 1988; Schlich, 1995; Yackinous et al., 1999).
In perhaps its earliest form, EPM was conducted using a computer program called PREFMAP (Chang & Carroll, 1972). The product space was generally set by PCA where mean descriptive scores were used to generate a representation of the products’ sensory characteristics (Meullenet et al., 2007). In some instances, other analyses such as Canonical Variates Analysis were used when PCA did not explain a sufficient amount of variability in the data (B. Clavier, personal communication, February 5, 2007). Under some circumstances, it is more appropriate to use centered (covariance matrix) or standardized (correlation matrix) data. Centering (not standardizing) the data is preferable if the sensory characteristics were measured on the same scale as is typical in descriptive analysis (Borgognone et al., 2001; Meullenet et al., 2007). This is because when data is standardized using the correlation matrix, variance on the attributes is treated as equal such that small differences between samples can be exaggerated and large differences can be suppressed, potentially leading to incorrect orientation of samples within the sensory space.

In a second step, multiple linear regression was used to fit consumers into the descriptive sensory space. As the name suggests, the vector (or linear) model maps consumers onto the descriptive sensory space as vectors pointing in the direction of increasing preference. The limitation of this model is the point at which liking plateaus or begins to decrease is not indicated. As with IPM, data analysis with EPM is based on individual ratings of consumers (in the second step), which has advantages over analyses using averaged ratings. This makes it possible to investigate heterogeneity of preferences. Inspection of resulting preference maps can show segmentation of consumers (Guinard et al., 2001; Helgesen et al., 1997; van Kleef et al., 2006).

Another advantage of EPM is that it can begin by mapping a large number of products representing the entire product category. This map can be utilized to
strategically choose products from this space for consumer judgments (MacFie, 2007). Although a PCA can be used in the same way before IPM, for the uninitiated the apparent change in product placement between the PCA map of the descriptive sensory space and the IPM map of the liking space can be confusing.

Disadvantages also exist for EPM. There are many factors which drive consumer preference other than the products’ sensory characteristics (for example, brand, packaging, price, and nutritional content to name a few). An assumption of EPM is that similar profiles will be similarly preferred and in some cases this is not true (van Kleef et al., 2006). For example, descriptive analysis may focus specifically on one modality such as flavor but consumers also use appearance, texture, etc. in their assessment of liking. In creating the preference map, there is no prioritization of attributes based on importance to consumers. Because of this, in extreme cases, the map may have no relevance to consumer acceptance of products (Meullenet et al., 2007). For example, in evaluating vanilla ice cream, appearance, flavor and texture are considered equal in setting the sensory space but the visual cue of bean specks has a much greater impact on liking than flavor does.

Another disadvantage is that since EPM usually reduces data to fewer dimensions than IPM (Greenhoff & MacFie, 1994), the number of consumers who may be accounted for can be relatively low. In such cases, developers may miss opportunities to develop products for consumers who have not been sufficiently represented in the model (Guinard et al., 2001). Finally, as with IPM, EPM fails to indicate optimal levels of important attributes, making determination of optimum product profiles difficult (Meullenet et al., 2007).
**Response Surface Approach (RSA).** External preference mapping can be further refined using the response surface approach. As with more traditional EPM, it regresses individual consumer responses onto the sensory space. However, four possible models may be used in this process: vector, circular, elliptical and quadratic (in order of increasing complexity). The circular and elliptical models would take into account the point of optimum intensity of a sensory characteristic (Ennis & Bi, 1998). The elliptical model is more complex since the coefficients are not constrained to be equal across PCs. In some cases, an anti-ideal or U-shaped function is found. Finally, the quadratic model is the most complex as not only are these maximum points of liking considered, but also the interactions between sensory characteristics (Arditti, 1997; Ares et al., 2006; *XLSTAT Manual*, 2009). However, it can be difficult to translate map coordinates back into actual sensory attributes since these maximum liking points are regressed on principal components, not individual sensory attributes (Guinard et al., 2001). Furthermore, it is not always clear what product alterations will result in the desired sensory attribute scores.

Next, the “acceptable” region of the space is determined for each consumer and compiled onto one map. Various contours in the map then correspond to the percentage of consumers who would theoretically appreciate a virtual product located in that space. One advantage of the RSA to EPM over the original EPM is that because of the inclusion of higher order polynomials, it is easier to determine optimum levels of liking drivers. Another advantage is that the contour map can be utilized to reveal unmet consumer desires or gaps in the current product space. A disadvantage of this technique is that as a consequence of this map’s construction method, the optimal profile of the virtual product is not known and is typically derived based on the profiles of the three closest real products in the space. Additional disadvantages may arise based on location
of products in the space. When there are no real products located in the region of a preferred virtual product, the profile of the virtual product can only be roughly estimated and its predicted overall liking is often overestimated, reaching beyond the upper limit of the hedonic scale used (Meullenet et al., 2007).

**Partial Least Squares Regression**

Partial least squares (PLS) regression is another method commonly used to relate consumer and descriptive data to investigate the sensory characteristics underlying consumer preference. While the methods previously discussed have extracted principal components from either consumer data alone or descriptive data alone, PLS extracts factors that explain the variability in liking as a function of the variability in descriptive data simultaneously. The PLS product space is generally set where mean liking and descriptive scores are used to generate a representation of the products’ sensory characteristics that are most important to a products’ liking or acceptance (Meilgaard et al., 2007; Meullenet et al., 2007; Tang, Heymann, & Hsieh, 2000).

Some researchers are accustomed to normalizing data prior to PLS and it has become a common feature of statistical packages. This is done in order that attributes with low intensities or low variance have equal chance to explain liking in the model since there is no *a priori* prediction of this. This also prevents attributes with high intensities or high variance from masking the impact of other attributes. Statistical packages generally also include a validation step, such as jack-knifing or bootstrapping, in order to obtain a measure of the effect of each descriptive variable (Meullenet et al., 2007) and the predictive quality of the regression model (*XLSTAT Manual*, 2009).
One main advantage of PLS is that consumer and descriptive data are considered together to derive the product space where the most salient characteristics related to consumer liking are emphasized. Compared to other regression techniques, specifically Principal Component Regression used in IPM and EPM, PLS is desirable as it decomposes the data into fewer principal components thereby overcoming issues with multi-collinearity (correlated attributes) which can be typical of descriptive data (Tang et al., 2000; *XLSTAT Manual*, 2009). In addition, the principal components obtained are sure to be relevant to liking (Helgesen et al., 1997), yielding simpler models that are easier to interpret (Martens & Martens, 2001). Other advantages include the ability to predict outcomes once the model is developed, and determine and compare importance of drivers (Meullenet et al., 2007).

One main disadvantage of PLS is the use of average data (Meullenet et al., 2007; Tang et al., 2000). Averaging of data is sufficient when data are homogeneous but in consumer data this is rare (Moskowitz & Krieger, 1995) and average data may not represent any individual consumer’s opinion. When heterogeneous data are averaged, only high level consumer patterns may be realized and important patterns in consumer liking may be missed (Tang et al., 2000). This may be overcome by segmenting consumers into smaller homogeneous groups by alternate means, such as cluster analysis, prior to conducting PLS (Ares et al., 2006). An additional disadvantage of PLS is its deterministic nature and assumption that data are fixed points with no accounting for variance (Meullenet et al., 2007).

**Investigations comparing preference mapping methods**

As detailed above, there are three main types of analyses for relating sensory and consumer data, each of which derives the product space differently. First, IPM
derives the product space from consumer data alone; second, EPM derives the product space from sensory information alone; and third, PLS derives the product space by considering sensory and consumer information together. While these analyses are clearly different, it is not clear which is best at relating sensory and consumer data for the purpose of product development. Fortunately, a few published investigations have compared various combinations of these methods.

To compare IPM and EPM of beer, Guinard et al. (2001) compared the utility of the outputs. Interactions in the hedonic data between samples and demographic classifications indicated a need to apply preference mapping techniques which account for individuals, not just averages. Since both IPM and EPM have this capability, both methods gave rise to similar conclusions in terms of consumer segments. However, EPM was able to identify the location of an ideal beer on the map and characterize the profile based on that location, making it the more actionable method (Guinard et al., 2001).

Helgesen et al. (1997) compared preference directions for sausage from IPM, EPM and PLS. Overall, the three methods gave similar results with IPM and EPM differing the most. Qualitatively, IPM and EPM indicated the same product characteristics drove preference, but quantitatively the importance of the characteristics were weighted a bit differently. EPM and PLS created nearly identical product spaces. In this particular instance, the main characteristics used by consumers to determine preference were closely related to the main characteristics used by the trained descriptive panel to describe the perceptual differences between the samples (Helgesen et al., 1997). It is important to note that this is not always the case. In fact, the mindset for assessing a product’s perceptual/sensory characteristics may be very different from that used in assigning preference (van Kleef et al., 2006). Furthermore, it is well known
that trained assessors (for example, descriptive panelists) may detect smaller
differences between samples than the normal population (Ishii, Chang, & O’Mahony,
2007).

Ares et al (2006) also compared IPM, EPM and PLS while investigating dulce de
leche texture. The three methods gave similar results when the relationship between
attribute intensity and liking was linear. When the relationship was not linear, EPM was
considered superior since optimal levels of important attributes could be predicted.
Additionally, the results demonstrated how averaging consumer data could lead to
incorrect conclusions. Specifically, when all consumers were modeled together, it
appeared that an intermediate level of hardness was optimal. However, further
investigation indicated two sub-groups of consumers, one that desired a hard texture
and one that desired a soft texture (Ares et al., 2006).

While most research has focused on either developing the preference mapping
techniques or comparing the methods, one study explored how end-users interpret and
apply the information. That is to say, is the information actionable and easily linked to the
task at hand? Researchers hypothesized and confirmed that IPM was considered more
actionable by marketing (particularly for consumer segmentation purposes) and EPM
was considered more actionable for food technology. However, it is important to note
that the respondents in the study were consumer scientists, market researchers and
sensory analysts (van Kleef et al., 2006). In fact, the true actionability by product
developers was not measured since actionability for food technology was assessed by
interviewing sensory analysts. In practice, sensory analysts typically prepare such
models for product developers and the opinions of this latter group of end-users could
provide another perspective. Anecdotally, product developers seem more in tune with
EPM since they typically focus on the products themselves and their common practice is to study the product space prior to testing consumers. Nevertheless, formal investigation of this group could provide additional insights.

**Pre-processing data before preference mapping**

When a researcher chooses to conduct a preference mapping study, a decision must be made on how to treat the consumer data. If a consumer researcher is interested in studying the population as a whole, then it could be sufficient to average the data across consumers and submit the means to preference mapping. One major assumption for this type of analysis is that all consumers in the population behave in the same manner and are adequately represented by the averaged data (MacFie & Thomson, 1988). However, the experienced consumer researcher knows this is rarely the whole story. Consumers often exhibit individual differences in their preferences. This has lead to preference mapping techniques that rely on data from each individual subject (MacFie & Thomson, 1988).

Yackinous (1999) investigated consumer preference of Ranch dressing on both an individual and average basis. When examined on an aggregate level, consumer ratings barely reached the mid-point on the hedonic scale (i.e., mean liking ratings ranged between “dislike slightly” and “neither like nor dislike”), and developers received only limited insight into how to formulate the product. When examined at the individual level, preferences varied widely across individuals, and sub-populations (or consumer segments) were identified. In effect by averaging liking scores across sub-populations, overall liking of products was suppressed due to this variability. By identifying the consumer segments, the developer was able to determine that additional market share could be obtained by marketing two products instead of one. Further, Tang et al. (2000)
found that when the preference mapping method explains only a small portion of the variance, such as when only about half of the consumers in a study can be fitted, a considerable amount of information is lost in attempting to predict liking and important information may be masked.

Various methods may be employed to segment consumers on the basis of preference. IPM is the method often touted for identifying groups of consumers. This is largely due to the fact that the product space is derived from consumer data alone (Guinard et al., 2001; Meullenet et al., 2007). However, the segments may be difficult to interpret (Arditti, 1997). Cluster analysis is another tool commonly employed to segment consumers based on preferences (Hottenstein et al., 2008). Hierarchical Cluster Analysis (HCA) is the method most commonly employed in the literature to look for population sub-groups (for examples, see Arditti, 1997; Jaeger et al., 2003; Martinez et al., 2002; Pagliarni et al., 1997). Agglomerative HCA begins by assuming each consumer is in a cluster alone. The consumers are systemically clustered into groups where variation within a group is minimized and variation between groups is maximized. A researcher must also take care to avoid overfitting the data by ensuring preference patterns are sufficiently different by graphically comparing means of the identified consumer groups. It is also important that the number of consumers in a cluster still represents a meaningful portion of the population. Some recommend that a cluster contains at least 20% of the total number of consumers tested (Meullenet et al., 2007), while others recommend maintaining sub-groups of at least 50 consumers (MacFie, 2007).

Anecdotally, a common complaint of segmenting consumers by preference is the difficulty in tying the segments to meaningful demographics or other information to aid in marketing the products developed to these unidentified consumer segments. In one
study, a large group of consumers (n=378) across 5 countries evaluated several cod products and answered questions regarding attitudes and fish consumption in addition to typical demographic questions. HCA was applied to obtain 4 clusters of consumers each with differing patterns of liking. While there was no direct link between cluster and country, ties to demographic and attitudinal information was found (Sveinsdottir et al., 2009). While it is not always possible to find useful ties between preference segments and characterizing information, it could be that the correct characterizing information was not collected and researchers should consider expanding questionnaires beyond their current norms.

Hottenstein et al (2008) found evidence that consumers can be assigned to “preference segments” as a result of serving order rather than real differences in liking. Several remedial measures were suggested for adjusting the data to address “serving-order-by-segment” effects. Adjustment factors varied based on treatment of samples and serving order. While the number of preference segments remained the same before and after the adjustment factors, assignment of consumers to preference segments was altered. The number of samples assessed was found to have an impact on the “serving-order-by-segment” effect, particularly in studies containing six or less samples (Hottenstein et al., 2008).

Another possible pre-processing step is to remove consumers who do not appear to discriminate the products in terms of preference. When hedonic ratings are flat, it is difficult to interpret a consumer’s true preference. On one hand, the consumer may not have a preference or is not expressing a preference due to boredom or other factors. On the other hand, the consumer may have a strong but similar preference for all products.
In either case, these consumers will not contribute to understanding which product features drive preference. Thus, such individuals are sometimes removed prior to analysis allowing for the development of a clearer picture (Arditti, 1997).

However, no guidelines exist for determining how to select low variance consumers or the maximum percentage of consumers that can be removed while still retaining sensible results. Arditti (1997) used HCA to identify a group of consumers who showed no difference in preference for the samples. All such consumers were removed and no further analyses were conducted on these consumers. While the reason for removing the consumers was clear, 43% of the consumers were eliminated from the analysis. It is difficult to know if removing this many consumers revealed true drivers of liking or instead distorted the underlying structure of the liking space. The IFPrograms™ software used to conduct LSA has provisions for excluding low variance raters (0-99%) who are said to sit in the middle of the map, possibly obscuring segmentation of discriminating consumers (IFPrograms Manual, 2009; Meullenet et al., 2007). The Institute for Perception suggests it may be normal to exclude up to 20% of the consumers in a study without consequence (D.M. Ennis, Seeing the Market through the Eyes of the Consumer, Institute for Perception Short Course, November 4, 2009). Even if a large number of consumers are excluded from the analysis, if the basis of their removal was flat hedonic responses, then these excluded individuals should equally like the products optimized for the more discriminating consumers.

Considerations for choosing the best model

In preference mapping, as in any modeling where regression techniques are used, it is important to assess the adequacy of the model as measured by the amount and type of error. Even though a reasonable amount of error is expected, by testing the
properties of the residual error, or residuals, of a model, researchers are given clues as to the model’s adequacy. From these tests, it can be determined if the residuals are acceptable or if remedial measures (such as squaring the data) need to be applied to correct inadequacies (Mendenhall & Sincich, 2003). Specific tests, commonly available in statistical analysis packages, may be employed to test if the residuals conform to three major assumptions (detailed below) upon which the regression techniques rely.

The first assumption is that errors are uncorrelated, or random. When residuals are not random, it can be an indicator that an important independent variable (such as a sensory descriptor) has been excluded from the model or is not adequately described by the model (for example, a quadratic function denoted by a linear term, requiring the addition of the appropriate squared term). A Runs Test is appropriate to test for random residuals.

The second assumption is that error variance is constant, or, in other words, residuals are homoscedastic. Heteroscedastic variances need to be investigated in more detail. For example, when a sensory descriptor is at high intensity or low intensity, a boundary effect can occur and the variance can be distorted. Levene’s Test is appropriate when residuals are normally distributed. The Modified Levene’s Test is more robust and appropriate to use when residuals are not normally distributed. If either version of the Levene’s test indicates heteroscedasticity, then data transformations (such as squaring, taking the square root or taking the log of the raw data) or the removal of outliers may be necessary.

The third assumption is that errors are normally distributed with a mean of 0. When normality is not found, there is a tendency for the model to underestimate liking (possibly resulting in a missed opportunity) or to overestimate liking (possibly resulting in an unsuccessful product launch). As in the cases of heteroscedasticity, data
transformations and/or the removal of outliers can help overcome non-normality. To ensure optimum utility of the model, it is important to check each assumption with each model, re-checking as the retained information in the model is altered. A Normality Test is appropriate to determine if residuals are normally distributed (J.C. Parcon, Assessment of Model Adequacy Workshop, Givaudan Flavors Corporation, July 21, 2009).

In selecting the best model, the principle of parsimony is widely accepted. The principle states that simpler models with fewer parameters are preferred because they are easier to interpret. In a study of EPM, various statistics were calculated to help guide the researcher in choosing the most parsimonious model, including the Fisher test of nested models (Schlich, 1995) and typical regression statistics such as $R^2$. First, the researcher must decide if the simpler vector model was sufficient to explain the data instead of the more complex elliptical model. Second, if the vector model was sufficient, the researcher then must decide how many components to retain in the model (Arditti, 1997). Such an approach is ideal and avoids the pitfalls of conducting an analysis without testing assumptions are met.

**The purpose of this research investigation**

Preference mapping studies consume a considerable amount of time and money. For this reason, there is an expectation of success as a result of the outputs. When such investigations are unsuccessful, many unanswered questions arise and it is important to determine the root cause to avoid a reoccurrence of the same issue in future studies. Preference mapping studies are necessarily complex and error can be introduced at each step. These steps include but are not limited to product selection, descriptive analysis, consumer testing, data analysis and interpretation of the outputs.
The purpose of this investigation is to examine these issues and develop best practices for conducting preference mapping studies. A dataset was obtained and analyzed using various preference mapping techniques. Each of the resulting models was evaluated by statistical criteria to determine model adequacy. Model outputs were then compared and patterns were noted. Input data was then reviewed in an attempt to unveil common characteristics of the input data leading to inferior or superior models. The overall outcome is a series of best practices for collecting sensory and consumer data as well as for modeling the information.
Chapter 2: Materials and Methods

Overview

A large group of commercially available Swiss-type cheeses were assessed by descriptive analysis with trained panelists for flavor and feeling factors as well as by consumers for overall liking.

Materials

Member companies of the Swiss Cheese Consortium provided 15 Swiss cheeses for testing. Specifically, ten domestic Swiss cheeses (S06 to S15), 4 domestic Baby Swiss cheeses (B1 to B4) and 1 imported Swiss Emmenthal (E5) were studied. The domestic Swiss cheeses represented typical characteristics of the Midwest and Northeast regions of the United States. All of the domestic Swiss cheeses were produced with pasteurized or thermalized milk while the Swiss Emmenthal was produced with raw milk. Fat content of the cheeses ranged from 29% to 34% and protein content ranged from 24% to 26% (Liggett, Drake, & Delwiche, 2008).

Blocks of cheese ranging from 2kg to 15kg arrived in the manufacturer’s original packaging. Prior to evaluation, cheeses were cut into uniform 2cm cubes and cubes with eyes large enough to distort the cube shape were discarded. For all sensory and consumer assessments 3 cubes were served in a 1oz translucent plastic soufflé cup (Solo Cup Co., Urbana, IL, USA) labeled with a neutral 3-digit number. Serving temperature was approximately 10°C (Liggett et al., 2008).
Descriptive Analysis

Descriptive analysis of Swiss cheese flavor was conducted at North Carolina State University in Spring 2004. All methods and procedures were in compliance with institutional human subjects regulations. Trained descriptive panelists (n=12) evaluated all 15 cheeses using the Spectrum method. Each panelist had at least 100h of experience evaluating cheese flavor. A descriptive language originally developed for Cheddar cheese flavor (Drake, McIngvale, Gerard, Cadwallader, & Civille, 2001) was adapted to Swiss cheese (see Table 2.1). Panelists rated each attribute on a 15pt universal intensity scale. Universal intensity scales decrease the time needed to train and align panelists to a new type of product and ratings upon them are less prone to shifting over time than are relative intensity scales (Drake & Civille, 2003).

Cheeses were evaluated one at a time in triplicate following a balanced complete block design. Panelists evaluated samples isolated in individual booths free from external distractions. Responses were recorded on paper ballots. Panelists were instructed to expectorate all samples and cleanse the palate with ambient temperature spring water between each sample (Liggett et al., 2008).

Consumer Assessment of Liking

Consumer testing was conducted at The Ohio State University in Spring 2004 and all methods and procedures were approved by the Office of Responsible Research Practices. Untrained consumers (n=101; 48 male and 53 female; ages 18 years to 65 years) were recruited on campus based on willingness to assess Swiss cheese.
Table 2.1. Lexicon used to profile Swiss cheeses

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Definition</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooked/milky</td>
<td>Aromatics associated with cooked milk</td>
<td>Skim milk heated to 85°C for 30 min</td>
</tr>
<tr>
<td>Whey</td>
<td>Aromatics associated with Cheddar cheese whey</td>
<td>Fresh Cheddar whey</td>
</tr>
<tr>
<td>Diacetyl</td>
<td>Aromatic associated with diacetyl</td>
<td>Diacetyl</td>
</tr>
<tr>
<td>Milk fat</td>
<td>Aromatics associated with milk fat</td>
<td>Fresh coconut meat, heavy cream, δ-dodecalactone</td>
</tr>
<tr>
<td>Vinegar</td>
<td>Aromatics associated with vinegar</td>
<td>Distilled white vinegar, acetic acid</td>
</tr>
<tr>
<td>Dried fruit</td>
<td>Aromatics associated with dried fruits, specifically peaches and apricots</td>
<td>Dried apricot half</td>
</tr>
<tr>
<td>Fruity</td>
<td>Aromatics associated with different fruits</td>
<td>Fresh pineapple, ethyl hexanoate</td>
</tr>
<tr>
<td>Sulfur/eggy</td>
<td>Aromatics associated with cooked eggs</td>
<td>Hardboiled egg, mashed</td>
</tr>
<tr>
<td>Sulfur/cabbage</td>
<td>Aromatics associated with cooked cabbage</td>
<td>Boiled cabbage, dimethyl trisulfide</td>
</tr>
<tr>
<td>Cheesy/butyric acid</td>
<td>Aromatics associated with butyric acid</td>
<td>Butyric acid</td>
</tr>
<tr>
<td>Brothy</td>
<td>Aromatics associated with boiled meat or vegetable stock</td>
<td>Canned potatoes, Wylers low sodium beef broth cubes, methional</td>
</tr>
<tr>
<td>Nutty</td>
<td>The nut-like aromatic associated with different nuts</td>
<td>Lightly toasted unsalted nuts, unsalted cashew nuts, unsalted wheat thins</td>
</tr>
<tr>
<td>Sweaty</td>
<td>Aromatic associated with human sweat</td>
<td>Hexanoic acid</td>
</tr>
<tr>
<td>Cowy/phenolic</td>
<td>Aromas associated with barns and stock trailers, indicative of animal sweat and waste</td>
<td>Bandaids, p-cresol, phenol</td>
</tr>
<tr>
<td>Sour</td>
<td>Fundamental taste sensation elicited by acids</td>
<td>Citric acid (0.08 % in water)</td>
</tr>
<tr>
<td>Bitter</td>
<td>Fundamental taste sensation elicited by various compounds</td>
<td>Caffeine (0.08% in water)</td>
</tr>
<tr>
<td>Salty</td>
<td>Fundamental taste sensation elicited by salts</td>
<td>Sodium chloride (0.5 % in water)</td>
</tr>
<tr>
<td>Sweet</td>
<td>Fundamental taste sensation elicited by sugars</td>
<td>Sucrose (5% in water)</td>
</tr>
<tr>
<td>Umami</td>
<td>Chemical feeling factor elicited by certain peptides and nucleotides</td>
<td>MSG (1 % in water)</td>
</tr>
<tr>
<td>Prickle</td>
<td>Chemical feeling factor of which the sensation of carbonation on the tongue is typical</td>
<td>Soda water</td>
</tr>
<tr>
<td>Metallic</td>
<td>Chemical feeling factor elicited by metallic objects in the mouth</td>
<td>Aluminum foil</td>
</tr>
</tbody>
</table>

According to self-reported usage, 73% of respondents consumed Swiss cheese at least once per month and 35% of respondents consumed Swiss cheese at least once per week (Liggett et al., 2008).

Cheeses were evaluated one at a time following a balanced complete block design. In individual booths free of external distractions, consumers assessed each cheese for overall liking using the 9pt vertical hedonic category scale where 1 = dislike extremely, 2 = dislike very much, 3 = dislike moderately, 4 = dislike slightly, 5 = neither...
like nor dislike, 6 = like slightly, 7 = like moderately, 8 = like very much and 9 = like extremely (Peryam & Pilgrim, 1957). Responses were entered directly into a computer using Compusense five data collection and analysis software (version 4.6, Compusense Inc., Guelph, Ontario, Canada). Consumers were allowed to swallow or expectorate samples and were instructed to cleanse the palate with ambient temperature spring water between each sample (Liggett et al., 2008).
Chapter 3: Data Analysis

Overview

Three main categories of preference mapping techniques were identified for investigation. The three categories were internal preference mapping (product space constructed from consumer liking information simultaneously), external preference mapping (product space constructed from sensory descriptive information) and hybrid preference mapping (product space constructed by integrating consumer and sensory information). One popular version of each technique was chosen to represent each of these categories. Tests of model adequacy were employed to assess stability, model fit ($R^2$), and predictive model quality ($Q^2$).

In order to understand the characteristics of both the descriptive data and the consumer data, each data set was analyzed separately prior to preference mapping. After preference mapping, these characteristics were related to the various models to reveal which characteristics contributed any apparent model breakdowns.

Descriptive Data Analysis

Descriptive datasets were evaluated using standard descriptive statistics (average, range, etc.) and a series of three-way analysis of variance (ANOVA) with product, panelist and replication as main effects. When ANOVA of an attribute indicated a significant difference (product $p<0.05$), Fisher’s LSD was used to determine which products differed significantly on that attribute. Principal components analysis (PCA) was
employed to aid in the visualization of the sensory space covered by the products, the identification of atypical products and outliers, and determination of highly correlated variables and issues with collinearity. Analyses were conducted using XLSTAT™ (Version 2009.4.03, Addinsoft, New York, NY, USA).

**Consumer Data Analysis**

Consumer datasets were evaluated using standard descriptive statistics (average, range, etc.) and a two-way analysis of variance (ANOVA) with product and consumer as main effects. When ANOVA indicated a significant difference between products (product \( p<0.05 \)), Tukey’s HSD was used to determine which products differed significantly in overall liking. Analyses were conducted using XLSTAT™ (Version 2009.4.03, Addinsoft, New York, NY, USA).

**External Preference Mapping: Response Surface Approach**

For external preference mapping (EPM), the response surface approach was employed. Descriptive analysis data were subjected to PCA in order to generate a map of the product space. While most descriptors were correlated with either the first or second principal components, three components were retained because two descriptors (nutty and prickle) were well correlated with the third component and did not load on the earlier components. Overall liking for each consumer was then regressed onto the space resulting in a contour plot of consumers’ appreciation for the products. From this plot, sub-groups possessing similar patterns of appreciation could be visualized. Model adequacy of the EPM was measured by correlating standardized liking scores of each consumer with the predicted liking scores for each individual from the preference map.
(\(R^2\) surface) and the number of consumers that significantly fit the product space. While specific guidelines for minimum criteria were not found in the literature, \(R^2\) values and number of fitted consumers greater than 50% are generally considered acceptable (Heyd & Danzart, 1998; Meullenet et al., 2007). Predictive quality of the models is not a standard output of the software used and was not considered. As is typical, positive and negative drivers were determined by visual inspection of scatter plots of mean overall liking scores versus mean intensity ratings and their correlations. Analyses were conducted using XLSTAT™ (Version 2009.4.03, Addinsoft, New York, NY, USA).

**Internal Preference Mapping: Landscape Segmentation Analysis**

For internal preference mapping, Landscape Segmentation Analysis® (LSA) was employed. Overall liking scores were used to generate a two-dimensional contour map depicting density of consumers’ ideal points and the relative position of the Swiss cheese products within that liking space. This plot also allows for the identification of sub-groups possessing similar patterns of appreciation. Descriptive data were then projected on the map to aid in the interpretation of underlying reasons for product acceptance. For LSA, model adequacy was determined by analyzing residuals from mean overall liking and predicted overall liking from the model as well as model fit for both products (\(R^2_{products}\) ) and consumers (\(R^2_{consumers}\)). In addition, mean liking scores and predicted liking scores from the model were correlated (\(R^2_{surface}\)). Predictive quality of the models (\(Q^2\)) is not a standard output of the software used and was not considered. Positive and negative drivers were determined quantitatively by correlating the descriptor to the product locations on the map. A descriptor is considered a driver when significantly correlated to the product space (correlation coefficient, \(p < 0.05\)). When the correlation is not significant, the descriptor is considered a non-driver. Optimal product
profiles were obtained from the software by strategically selecting points on the map, determining the coordinates, and relating them to product attributes. Analyses were conducted using IFPrograms™ (Version 8.5.1222 [8.7.0305], The Institute for Perception, Richmond, VA, USA).

**Hybrid Preference Mapping: Partial Least Squares Regression**

Partial Least Squares Regression (PLS) was employed to simultaneously relate consumer overall liking data (dependent variable) and descriptive attributes (independent variables) to generate a map to visualize the product-liking space in multiple dimensions. For PLS, model adequacy was determined by analyzing residuals from mean overall liking and predicted overall liking from the model as well as measures of model fit ($R^2$) and predictive quality of the model ($Q^2$). The latter two measures were also used to determine the number of dimensions retained in the model. Positive and negative drivers were determined quantitatively by examining cumulative variable importance on each dimension and standardized regression coefficients. Optimal product profiles were obtained by conducting reverse PLS using product factor scores as independent variables and descriptive attributes as dependent variables. Analyses were conducted using XLSTAT™ (Version 2009.4.03, Addinsoft, New York, NY, USA).

**Testing Model Residuals**

Three measures were employed to assess residuals of preference models. A Runs Test was conducted to determine if a model’s residuals were random. The Runs Test is based on the null hypothesis that data are randomly distributed and a p-value greater than 0.05 is desirable. Levene’s Test was used to determine if a model’s residual variances were homoscedastic. Based on the null hypothesis that variances are
homogeneous, a conservative p-value greater than 0.25 is desirable. However, Levene’s Test is only considered reliable if the residuals are normally distributed. When residuals are not normally distributed, the Modified Levene’s Test should be used. A Normality Test was used to determine if a model’s residuals were normally distributed. Based on the null hypothesis that residuals are normally distributed, a conservative p-value greater than 0.15 is desirable. Analyses were conducted using XLSTAT™ (Version 2009.4.03, Addinsoft, New York, NY, USA). If the residuals from any of the models failed to meet these requirements, then appropriate steps were taken to correct the deviations and the cause and effect of these deviations were discussed.

Segmenting Consumers

Cluster analysis is a statistical technique commonly used to explore and group data into categories based on shared characteristics. Based upon similar patterns of individual liking scores for each of the products, consumers were segmented using agglomerative hierarchical clustering (complete linkage, city-block distance). Analyses were conducted using XLSTAT™ (Version 2009.4.03, Addinsoft, New York, NY, USA).
Chapter 4: Results and Discussion

Descriptive Data Results

The Swiss cheeses (Baby Swiss, Swiss Emmenthal and Swiss) differed significantly (p<0.05) in their flavor profiles (see table 4.1). Overall, mean intensity ratings for the descriptors fell between 0 and 4. While these scores may seem low, they were collected using a 15-pt universal scale (Meilgaard et al., 2007) which is meant to span the perception of all food products for a given descriptor. These intensities also align with published literature values for cheeses (Drake et al., 2001).

Looking across all measured descriptors, only six were present in all assessed Swiss cheeses and these six differed significantly (see Table 4.1). Specifically, these descriptors were Cooked/milky, Milk fat, Vinegar, Sulfur/cabbage, Sweet and Umami. An additional six descriptors differed significantly but were not present in all 15 Swiss cheeses: Whey, Dried fruit and Sweaty (each present in 9 cheeses), Diacetyl (present in 5 cheeses), Brothy (present in 3 cheeses), and Nutty (present in 2 cheeses). Finally, three “binary” descriptors were present in only one cheese. These descriptors include Bitter and Metallic (present in only B1), and Prickle (present in only S11). Fruity, Sulfur/eggy, Cheesy/butyric acid, Cowy/phenolic, Sour, and Salty were not detected in any of the Swiss cheeses (data not shown) (Liggett et al., 2008).

Unlike other published reports (e.g., Beuvier et al., 1997; Lawlor, Delahunty, Wilkinson, & Sheehan, 2003), nutty aroma and flavor were not prevalent in the cheeses assessed in this study. One possible explanation is that domestic and international
Table 4.1. Mean descriptive analysis intensity ratings for Swiss cheeses

<table>
<thead>
<tr>
<th>Attribute</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>E5</th>
<th>S6</th>
<th>S7</th>
<th>S8</th>
<th>S9</th>
<th>S10</th>
<th>S11</th>
<th>S12</th>
<th>S13</th>
<th>S14</th>
<th>S15</th>
<th>LSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooked/milky</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.2</td>
<td>1.4</td>
<td>1.3</td>
<td>1.4</td>
<td>1.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Whey</td>
<td>1</td>
<td>--</td>
<td>1</td>
<td>2</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1.2</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>1.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Diacetyl</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>2</td>
<td>--</td>
<td>2</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1.1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Milk fat</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1.3</td>
<td>1.1</td>
<td>1.4</td>
<td>1.3</td>
<td>1.5</td>
<td>1.7</td>
<td>0.3</td>
</tr>
<tr>
<td>Vinegar</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1.2</td>
<td>1.6</td>
<td>1.1</td>
<td>1.7</td>
<td>1.2</td>
<td>1.4</td>
<td>0.3</td>
</tr>
<tr>
<td>Dried fruit</td>
<td>--</td>
<td>2</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>1.1</td>
<td>--</td>
<td>1</td>
<td>1.2</td>
<td>1.3</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>Sulfur/cabbage</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1.9</td>
<td>2.2</td>
<td>1.6</td>
<td>1.9</td>
<td>1.9</td>
<td>0.3</td>
</tr>
<tr>
<td>Brothy</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0.2</td>
</tr>
<tr>
<td>Nutty</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>2</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>0.3</td>
</tr>
<tr>
<td>Sweaty</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>--</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1.5</td>
<td>1.2</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>0.3</td>
</tr>
<tr>
<td>Bitter</td>
<td>2</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Sweet</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2.7</td>
<td>3</td>
<td>2.6</td>
<td>3</td>
<td>3</td>
<td>2.9</td>
<td>0.2</td>
</tr>
<tr>
<td>Umami</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3.4</td>
<td>3.6</td>
<td>3.4</td>
<td>3.4</td>
<td>3.3</td>
<td>3.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Prickle</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>1.2</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Metallic</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)
0 = not present, 15 = very high intensity, - - = the mean panel value < 0.5, --- LSD was not calculated for binary attributes
Means in a row that differ by more than LSD are significantly different (p<0.05)
Swiss cheeses differ in this attribute, however, Nutty was not found in the imported Swiss Emmenthal either. This difference seems to have arisen instead from the way in which nutty flavor was defined and anchored with physical references. Without references to anchor the term nutty, the concept will not be consistent across studies. It appears the nutty concept in other studies was inconsistent with the nutty reference used in this study (Liggett et al., 2008).

Principal components analysis of the assessed Swiss cheeses explains 71.6% of the product variance in the first three dimensions (see Figures 4.1 and 4.2). There is no visible pattern differentiating the Swiss cheese types as all types are interspersed, e.g., B2 (Baby Swiss), S8 (Swiss) and E5 (Swiss Emmenthal) are located next to each other on the map. The Baby Swiss cheeses are spread across 3 directions of the map. That is, B1 is located away from all cheeses in the direction of Bitter and Metallic; B2 is located near cheeses which are high in Nutty and Sweet; and B3 and B4 are located near cheeses which are characterized by Whey and Cooked/milky.

The first dimension (34.5% explained variance) is characterized mainly by Whey and Cooked/milky in one direction as found in B3, B4 and S6. The opposing direction is Sulfur/cabbage, Vinegar, Brothy and Dried fruit. Dominating characteristics of B2 and E5 are Sulfur/cabbage, Brothy and Dried fruit. S8 is high in Sulfur/cabbage and Vinegar; while S13 is high in Vinegar and Dried fruit. However, Sulfur/cabbage and Vinegar were present in all samples. The second dimension (23.9% explained variance) is characterized mainly by Bitter and Metallic in one direction (which was unique to B1), and Milk fat (as in B4 and S6) and Sweet (as in B2, E5 and S6) in the other direction. Like Sulfur/cabbage and Vinegar, Milk fat and Sweet were present in all samples. The third dimension (13.2% explained variance) is characterized mainly by Nutty (only in B2 and S9) in one direction and by Prickle (only in S11) in the other direction.
Consumer Liking Results

Overall liking of the assessed Swiss cheeses (Baby Swiss, Swiss Emmenthal and domestic Swiss) were found to differ significantly (p<0.05) (see table 4.2). Overall, mean liking ratings fell between 4.4 and 6.0. While these scores may seem low, liking ratings are generally dependent upon the product type or category under investigation as well as the culture. For example, in the United States, liking scores for ice cream are
expected to be higher than for chicken nuggets. In a study on Cheddar cheeses, liking scores ranged from 5.7 to 7.2 on a 9 pt scale (N. Young, Drake, Lopetcharat, & McDaniel, 2004); and in a more recent study on Gouda cheeses, mean liking values ranged from 5.5 to 6.5 on a 9 pt scale (Yates & Drake, 2007). Thus, liking scores in this study are in agreement with the reported literature on similar products (Liggett et al., 2008).

Figure 4.2. PCA of descriptive analysis for Swiss cheeses, PC3 vs PC1
Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)
◆ = product; ○ - attribute
Table 4.2. Mean consumer liking ratings for Swiss cheeses

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mean Overall Liking</th>
<th>Significant Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>B3</td>
<td>6.0</td>
<td>A</td>
</tr>
<tr>
<td>S11</td>
<td>6.0</td>
<td>A</td>
</tr>
<tr>
<td>B1</td>
<td>5.9</td>
<td>A</td>
</tr>
<tr>
<td>S06</td>
<td>5.8</td>
<td>AB</td>
</tr>
<tr>
<td>B4</td>
<td>5.7</td>
<td>ABC</td>
</tr>
<tr>
<td>B2</td>
<td>5.5</td>
<td>ABCD</td>
</tr>
<tr>
<td>S07</td>
<td>5.4</td>
<td>ABCD</td>
</tr>
<tr>
<td>S15</td>
<td>5.3</td>
<td>ABCD</td>
</tr>
<tr>
<td>S10</td>
<td>5.2</td>
<td>ABCD</td>
</tr>
<tr>
<td>S14</td>
<td>5.2</td>
<td>ABCD</td>
</tr>
<tr>
<td>S08</td>
<td>5.0</td>
<td>BCDE</td>
</tr>
<tr>
<td>S13</td>
<td>5.0</td>
<td>CDE</td>
</tr>
<tr>
<td>S09</td>
<td>4.8</td>
<td>DE</td>
</tr>
<tr>
<td>E5</td>
<td>4.8</td>
<td>DE</td>
</tr>
<tr>
<td>S12</td>
<td>4.4</td>
<td>E</td>
</tr>
</tbody>
</table>

Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)
1 = dislike extremely, 9 = like extremely
Means with the same letter (A-E) do not differ in overall liking (Tukey HSD, p<0.05)

On average, B3 (Baby Swiss) and S11 (Swiss) were rated the highest in overall liking and were equally appreciated by consumers. Several other Swiss and Baby Swiss cheeses were liked by consumers. In general, the Baby Swiss cheeses in this study tended to score higher than the Swiss cheeses. Most consumers disliked the imported Swiss Emmenthal as well as certain other Swiss cheeses, such as S09 and S12 (Liggett et al., 2008).
Overall, results showed that the vector model was the best model for 71 out of 101 consumers. This is the simplest model available to explain the relationship between a consumer’s liking and the products under investigation, but vector models cannot indicate optimal levels of positive and negative drivers. The remaining 30 consumers were fitted by more complex models (circular, elliptical and quadratic), all of which allow for ready determination of ideal points. All 101 individual models were then superimposed to construct the contours of the map. Across all consumers, $R^2$ values ranged from 96.3% down to 1.7% with only 44 of the 101 consumers significantly fitting the product space. The resulting overall fit of the predicted liking values to actual liking values ($R^2_{\text{surface}}$) was only 42.3% indicating the model was not very reliable (since the $R^2_{\text{surface}} < 50\%$). Residuals are not standard output of EPM analysis making it impossible to conduct Runs, Levene’s and Normality Tests. Regardless, given the small number of consumers that were significantly fitted and the low overall fit, adequacy and any subsequent interpretation of the preference map would be highly questionable.

Visible inspection of the contour plots revealed that consumer appreciation for the samples was only in partial alignment with the overall liking scores (see Figures 4.3 and 4.4). According to the contour plot, 60-70% of consumers appreciated B1, B3, B4 and S06. This was in alignment as these are 4 out of 5 of the top liked cheeses. However, S11, the second most liked cheese, fell in a region where only 30-40% of consumers appreciate it. Conversely, S12, the least liked cheese, fell in a region where 50-60% of consumers appreciated it. The placement of S11 was rationalized by examining its placement on the third dimension. Recall that S11 was the only sample characterized by Prickle and this attribute loaded highly on the third dimension.
S12, however, cannot be rationalized in this way and it was most likely that poor fit by a large number of consumers was responsible for the discrepancy.

Positive and negative drivers for Swiss cheeses given this product set and these consumers were determined by visual inspection of the trace plots (see Figure 4.5). Diacetyl and Whey were both positive drivers. As Diacetyl and Whey increased, so did Overall Liking. In this data set, acceptance of Diacetyl leveled off suggesting that an additional increase in perceived intensity of Diacetyl would not increase liking any
further. For Whey, however, no plateau or point of maximum liking was found. Above the highest intensity of Whey present in the data set, the impact upon liking was unknown. On the other hand, Sulfur/cabbage was a negative driver. As Sulfur/cabbage increased, Overall Liking decreased. Correspondingly, the two least liked cheeses, S12 and E5, had the highest amount of Sulfur/cabbage.

An optimal profile cannot be determined for this group of consumers. Roughly one third of the map was appreciated by 60-70% of the consumers and the region was too broad to allow for a reasonable estimation of the optimal profile. In addition, there
Figure 4.5a. EPM trace plots of each attribute vs. liking for Swiss cheeses

- - - : Actual data  
- - : Linear equation  
- - : Polynomial equation

Cooked/milky
- $y = 0.69x + 4.4$  
  $R^2 = 0.11$

- $y = 1.1x^2 - 2.4x + 6.5$  
  $R^2 = 0.13$

Milk fat
- $y = 0.37x + 4.8$  
  $R^2 = 0.06$

- $y = 1.8x^2 - 5.7x + 9.7$  
  $R^2 = 0.25$

Sulfur/cabbage
- $y = 0.86x + 6.9$  
  $R^2 = 0.31$

- $y = -0.20x^2 - 0.13x + 6.3$  
  $R^2 = 0.31$

Sweaty
- $y = 0.03x + 5.4$  
  $R^2 = 0.001$

- $y = -0.51x^2 + 0.63x + 5.3$  
  $R^2 = 0.03$

Umami
- $y = 0.47x + 6.9$  
  $R^2 = 0.025$

- $y = 9.3x^2 - 62.3x + 109.5$  
  $R^2 = 0.42$
Figure 4.5b. EPM trace plots of each attribute vs. liking for Swiss cheeses
- - - - : Actual data  
: Linear equation  
: Polynomial equation
Figure 4.5c. EPM trace plots of each attribute vs. liking for Swiss cheeses

- - - - : Actual data  - - : Linear equation  - - - : Polynomial equation
was no region on the map where greater than 70% of consumers appreciated the samples, resulting in very little guidance to product developers.

There are several possible explanations for the lack of success of this preference map. First, consumers may not be well fitted because two equally liked products may have different sensory profiles. Looking at the map, B1 and S06 were located opposite each other on the second axis and yet they were equally liked. This also contributes to the ambiguity of determining an optimal profile. Second, the criteria of the consumers for appreciating the samples may have been different from the sensory information provided by the trained panel. Third, there may have been differences in criteria among the consumers in the test. These latter two hypotheses cannot be tested by EPM alone but could potentially be addressed by pre-processing the data with HCA.

**Internal Preference Mapping: Landscape Segmentation Analysis**

Overall, the space constructed by LSA based on consumer liking fits the products well ($R^2_{products}=99.2\%$). Examining the contour map (see Figure 4.6), product placement was somewhat aligned with consumer liking, but since $R^2_{surface}$ was only 50.2\% the correlation between predicted liking and actual liking was moderate at best. B1, S06 and B4 were located together on the map and all three were reasonably liked. Likewise, S08 and E5 were less-liked and were located near each other on the map. However, the two most liked samples, B3 and S11, were located proximate to the less liked S09. While the Runs Test indicated the residuals were random, the Normality Test indicated the residuals were not normally distributed. Examination of the actual and predicted liking scores showed the model tended to overestimate liking. The most discrepant, S09, was
overestimated by 1.3 hedonic points. This suggested the predicted optimal profile(s) also may be overestimated and thus theoretically optimized products would be unlikely to meet performance expectations.

In terms of consumers, only 49.6% were well-fitted to the map. This suggested consumers in the test were not using the same criteria to determine appreciation for the samples. Further, examination of the contours suggested that multiple consumer segments were present and that a single optimized product would not please all consumers. However, the densest regions of the map were void of samples. This potentially indicated an unmet consumer need or gap in the market and that adjusting the flavor perception of existing samples or creating new ones could increase consumer liking and possibly market share. Five attributes were responsible for driving liking or disliking in this product space. Specifically, Milk fat was a positive driver, and Sulfur/cabbage, Vinegar, Sweet and Dried fruit were negative drivers.

To investigate further, S09 was removed from the data and reanalyzed (data not shown). While $R^2_{\text{products}}$ (99.2%) and $R^2_{\text{consumers}}$ (51.7%) were fairly consistent, $R^2_{\text{surface}}$ jumped from 50.2% to a more acceptable 72.0%. However, residuals still were not normally distributed. Going one step further, S07 was removed from the data and reanalyzed since its predicted liking too was considerably overestimated (0.8 hedonic points). As before, this analysis resulted in fairly consistent values for $R^2_{\text{products}}$ (99.4%) and $R^2_{\text{consumers}}$ (54.0%), and $R^2_{\text{surface}}$ jumped again from 72.0% to 90.1%. With this model, residuals met the necessary criteria as indicated by the Runs, Levene’s and Normality Tests and optimal profiles thus could be predicted with more confidence. The location of products in the space yielded a nearly identical configuration; however, many more drivers were correlated to the space (see Figure 4.7). As in the original analysis with all samples, Milk fat was a positive driver and Sweet, Vinegar, Dried fruit and
Figure 4.6. LSA contour map of Swiss cheeses
Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)

The lighter the blue, the higher the density of consumer ideal points

- = product
○ = individual ideal point
→ = preference driver
Sulfur/cabbage were negative drivers. When S09 and S07 were removed from the analysis, Nutty and Whey were also identified as positive drivers and Brothy was also identified as a negative driver.

Caution must be taken, however, in interpreting Nutty as a positive driver as it was somewhat problematic. Recall that S09 (removed from this latter analysis), contained the highest intensity of Nutty in the whole product set. After removing S09, Nutty became a binary descriptor present in only B2. In recommending an optimal profile appreciated by consumers, optimal intensity of Nutty was difficult to predict since only a single intensity was considered in the model. Consumers’ reaction to intensities of Nutty between 0 and 1 was unknown, so any suggested optimal intensity would not be well supported. The same was true of Bitter and Metallic, neither of which were predictive of liking in this model.

Since two potential segments of consumers exist in the population, two optimal profiles were taken from the model (see Figure 4.8). Visual inspection of the graph showed the two profiles were similar as were the predicted liking scores. The profiles were so similar it is unlikely consumers or even a trained panel would be able to distinguish the differences not to mention it would be nearly impossible to manipulate natural products to match these profiles. These profiles, then, suggested multiple consumer segments do not exist. A more likely explanation is that a group of consumers was contributing noise to the analysis. Such noise could arise from consumers who give all products a similar score either because they like or dislike all the samples equally, cannot tell differences between the samples, or are simply unmotivated to take the task seriously. These consumers are of little use to product developers because they do not
Figure 4.7. LSA contour map of Swiss cheeses, excluding S09 & S07 Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)

The lighter the blue, the higher the density of consumer ideal points

● = product
○ = individual ideal point
→ = preference driver
show varying appreciation for the products (Meullenet et al., 2007). Further, one would predict that these panelists would accept all such products equally, even one optimized for other consumers.

Another indicator that this noise may have been distorting the map was that the densest region of the LSA map was void of samples. Two methods were employed to explore this possibility. One method, available in IFPrograms™, excludes “low variance raters.” The specific treatment of the data to determine which consumers were low variance raters is not known as it is an undisclosed proprietary feature embedded in the software. Using this feature, the resulting contour plot changed very little and the drivers remained the same (data not shown). A second method, cluster analysis of the consumer liking scores, revealed 2 segments of consumers (see Figure 4.9). One segment of consumers (n=61) gave varying hedonic ratings to the samples which...
spanned a more representative portion of the scale. The smaller segment (n=40) provided hedonic ratings on only a smaller portion of the scale, showing little differentiation of samples (see Table 4.3) (Liggett et al., 2008).

The contours of the LSA map for the discriminating consumers had some notable differences (see Figure 4.10). One of the segments from the previous analysis seemed to disappear and consumers’ appreciation for some products was diminished, specifically, B4, S10 and E5. Although several drivers remained the same (Milk fat, Sweet, Dried fruit and Sulfur/cabbage), some drivers shifted. Additional positive drivers, Diacetyl and Whey, were added. As for negative drivers, Vinegar was taken away and Brothy was added. These additional drivers provided additional insight to product developers as to how manipulations of flavor would impact consumer preference.

Further, an optimal profile for this group of consumers was estimated based on the model. Comparison of the most and least liked cheeses indicated what flavor directions increased liking. In most cases, the intensity of the descriptors needed to be moderated. That is, moderately decreasing Sulfur/Cabbage and Dried fruit from the levels in the least liked sample (S12) would increase liking, as would moderately increasing Diacetyl from the level in the least liked sample. Highest liking would likely be achieved by an absence of Brothy and, for Whey, Milk fat and Sweet, intensities consistent with the most liked sample (B3) should increase liking.

**Hybrid Preference Mapping: Partial Least Squares Regression**

Using PLS, attribute means from descriptive analysis and mean consumer liking scores were analyzed in a single step to determine specifically which attributes were driving liking for the Swiss cheeses under investigation. Based upon model fit ($R^2=93\%$)
Figure 4.9. HCA of consumers clustered by individual liking of Swiss cheeses
The left cluster of 40 consumers rated equal preference for all cheeses and the right cluster of 61 consumers shows varying preference for the cheeses
and predictive model quality ($Q^2=24\%$), 3 dimensions were retained in the model. Runs, Levene’s and Normality Tests indicated this model was adequate for interpretation.

Placement of products on the map aligned along the vector of overall liking as anticipated from the mean overall liking scores (see Figure 4.11). Since the analysis was conducted using mean scores, densities of consumer appreciation for the products cannot be assessed and reliance on a vector model does not provide ideal points, only directional information. Regardless, location of the product descriptors provided insight as to the product characteristics much like a PCA.

Several positive and negative drivers were identified using PLS. Positive drivers were Diacetyl, Whey and surprisingly Bitter (discussed below). Diacetyl, the largest positive driver, characterized B4 and S06, which were 2 of the most liked cheeses. B3 and S11 were characterized by higher levels of Whey. B1, another well-liked cheese, was the only cheese perceived as Bitter in the descriptive analysis and it was the response to this cheese alone that causes Bitter to be considered a positive driver. Negative drivers were Sulfur/cabbage, Brothy, Vinegar and Nutty. In the less liked samples, Sulfur/cabbage and Vinegar were perceived higher in E5 and S08, and Nutty was present in S09. The highest level of Brothy was found in E5.

Using this information, an optimal profile for this group of consumers was estimated based on the model (see Figure 4.12). Comparison of the most and least liked cheeses in the study indicated what flavor directions increased liking. Two descriptors stood out most notably for flavor adjustment. Increased overall liking should be achieved by decreasing Sulfur/cabbage in the least liked sample (S12) to a moderate intensity similar to the most liked sample (B3). In addition, increasing Diacetyl, which was absent from the least liked sample, to a level similar to the most liked sample should increase overall liking. A low level of Bitter should also increase liking; however, care should be
Table 4.3. Mean consumer liking ratings for two segments of consumers

<table>
<thead>
<tr>
<th>Sample</th>
<th>Differentiating Consumers</th>
<th>Non-differentiating Consumers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Overall Liking</td>
<td>Significant Differences</td>
</tr>
<tr>
<td>B1</td>
<td>6.3</td>
<td>A</td>
</tr>
<tr>
<td>B2</td>
<td>5.1</td>
<td>BCD</td>
</tr>
<tr>
<td>B3</td>
<td>6.0</td>
<td>AB</td>
</tr>
<tr>
<td>B4</td>
<td>6.2</td>
<td>A</td>
</tr>
<tr>
<td>E5</td>
<td>4.0</td>
<td>DE</td>
</tr>
<tr>
<td>S06</td>
<td>6.2</td>
<td>A</td>
</tr>
<tr>
<td>S07</td>
<td>5.2</td>
<td>ABC</td>
</tr>
<tr>
<td>S08</td>
<td>4.3</td>
<td>CDE</td>
</tr>
<tr>
<td>S09</td>
<td>4.6</td>
<td>CD</td>
</tr>
<tr>
<td>S10</td>
<td>5.2</td>
<td>ABC</td>
</tr>
<tr>
<td>S11</td>
<td>5.8</td>
<td>AB</td>
</tr>
<tr>
<td>S12</td>
<td>3.6</td>
<td>E</td>
</tr>
<tr>
<td>S13</td>
<td>4.2</td>
<td>CDE</td>
</tr>
<tr>
<td>S14</td>
<td>4.5</td>
<td>CDE</td>
</tr>
<tr>
<td>S15</td>
<td>4.7</td>
<td>CD</td>
</tr>
</tbody>
</table>

Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)
1 = dislike extremely, 9 = like extremely
Means with the same letter (A-E) do not differ in overall liking (Tukey HSD, p<0.05)

... taken in interpreting this finding. Bitter was a binary descriptor and its interpretation as a positive driver is somewhat problematic. Since Bitter was only present in sample B1, consumers’ reaction to intensities of Bitter between 0 and 1.5 is unknown. Suggesting an intermediate amount of Bitter injects a certain amount of risk. The same is true for Metallic and Prickle, which were not drivers in this model.
Figure 4.10. LSA contour map of Swiss cheeses (excluding S09 & S07) for differentiating consumers
Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)

The lighter the blue, the higher the density of consumer ideal points.

● = product
○ = individual ideal point
⇒ = preference driver
Figure 4.11. PLS map of Swiss cheese for PLS2 vs. PLS1
Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)

■ = product
● = attribute
Since PLS uses mean liking scores to build the model, it is possible that individual differences in preference are concealed and it is important to understand the impact this may have on the resulting model. From the previous cluster analysis evidence of two consumer segments existed: one group who differentiated the samples and another who did not. To further investigate, a model was built for the 61 consumers who gave varying hedonic responses. The resulting map was nearly identical (not shown) but there was improvement in the overall quality of the model. While the model was able to explain more variability in fewer components (80.5% to 85.4% with two components) the improvement in the predictive ability of the model was much increased on even the first component, from 7.4% to 22.5%. However, at three components the

Figure 4.12. Optimal profile of Swiss cheeses identified by PLS model derived with all products

Predicted liking of the optimal profile is 6.5.
B3 = most liked sample (mean liking = 6.0)
S12 = least liked sample (mean liking = 4.4)
predictive ability of the model decreased from 39.2% in the original model to 29.4%. Further inspection revealed sample B1 was an outlier. According to the descriptive analysis, B1 was characterized by two unique attributes, Bitter and Metallic. Since the goal of such research is to predict the best possible optimal profile, B1 and its uniquely associated descriptors (Bitter and Metallic) were removed from the analysis and a final model was developed (Liggett et al., 2008).

In this final model, 65% of the descriptive information was used to explain 93% of liking with 49% predictive ability (Liggett et al., 2008). Again, appearance of the map (see Figure 4.13) and the drivers remained fairly consistent; however, the resulting model behind it was more stable for predicting how a new or optimized product would be liked by consumers. Positive drivers of Diacetyl, Whey, Milk fat were identified. Negative drivers included Sulfur/cabbage, Brothy, Vinegar and Dried fruit. An optimal profile was also estimated. In this profile, the most notable suggested changes for the least liked product (S12) included increasing Diacetyl and Milk fat and decreasing Sulfur/cabbage. While Nutty and Sweaty were not considered drivers, presence of these characteristics at low levels should be acceptable to consumers (see Figure 4.14).

**Comparison of Techniques: EPM, LSA, PLS**

Based on overall model fit ($R^2$), LSA and PLS better explained the data than did EPM. This was largely due to the fact that the product space was based at least in part on consumer liking information in both LSA and PLS. Specifically, LSA constructs the product space solely on consumer information and PLS constructs the product space using a combination of consumer and sensory information. Thus, it should not be surprising that LSA and PLS perform better. To a fault, EPM bases the product space on objective information. In the case of sensory descriptive data, a trained panel may use a
Figure 4.13. PLS of Swiss cheeses (excluding B1) for differentiating consumers, PLS2 vs. PLS1
Baby Swiss cheeses (B1-B4), Swiss Emmenthal (E5), Swiss cheeses (S06-S15)

■ = product
● = attribute
mindset very different from a consumer (Cordonnier, 2007; van Kleef et al., 2006) and, as well, may be more sensitive to small differences between the products that a consumer may not discriminate. A similar issue would arise if the objective information was instrumental data or likewise, i.e., instruments may discriminate differences between products that a human cannot. That is not to say the objective product space would never align with the subjective perception of consumers but it is always a potential source of disconnect.

On the basis of individual consumer fit, LSA and EPM were generally in agreement. In each case, $R^2$ was around 50%. Various factors could have contributed to the low fit. As previously mentioned, low fit in EPM could mean that consumers use a

Figure 4.14. Optimal profile of Swiss cheeses for differentiating consumers identified by PLS model, excluding cheese B1 and attributes Bitter and Metallic

Predicted liking of optimal profile is 6.3.
B3 = most liked sample (mean liking = 6.3)
S12 = least liked sample (mean liking = 3.6)
different set of criteria in making hedonic judgments as opposed to the objective perceptions reported in the descriptive analysis. It is also true that products with very different profiles may be equally liked. Since the products are very different, they will be separated on the PCA map making it difficult to fit and explain the patterns of liking. However, these explanations would not hold true for LSA (where the space is set by the consumers) if the same problems were encountered. Low fit could also be attributed to multiple consumer segments present in the population tested. Since PLS uses mean liking scores, any individual differences in preference would be concealed prior to the analysis and would be difficult to uncover. Conversely, since LSA and EPM fit individual consumers, individual differences in product preference should be readily visualized by examining the contour maps. LSA provided some evidence for the existence of two segments and the products tested sit around the periphery of these dense areas. The contour map from EPM of these Swiss cheeses showed little evidence of segmentation instead showing relatively low appreciation for any of the samples tested. It could have been that there were two segments with opposite preferences resulting in a flat surface or that there was a group of consumers contributing noise to the analysis as previously discussed.

Overall, EPM found the fewest drivers but the process of driver selection was more subjective than with LSA and PLS (which found more drivers - see Table 4.4). Problematically, the number of drivers derived from EPM can vary depending upon the strategy employed to select the drivers. However, there were only slight differences in the drivers determined by each preference mapping technique. Diacetyl and Whey were found to be positive drivers in all three techniques. In addition, Milk fat was a driver in
### Table 4.4. Comparison of drivers determined by EPM, LSA and PLS

<table>
<thead>
<tr>
<th></th>
<th>EPM</th>
<th>LSA</th>
<th>PLS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diacetyl</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Whey</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Milk fat</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfur/cabbage</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Dried fruit</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Brothy</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Sweet</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Vinegar</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

* = attribute identified as a driver

LSA and PLS. Again, all three techniques were in agreement with Sulfur/cabbage and Dried fruit as negative drivers. Brothy was a negative driver in both LSA and PLS while Sweet was identified solely by LSA and Vinegar was identified solely by PLS.

Optimal profiles were determined only for LSA and PLS since the optimal region for EPM was too ambiguous to allow for such predictions. Overall, the two optimal profiles followed a similar pattern (see Figure 4.15). For positive drivers Whey and Milk fat, differences between the recommended intensities were likely not significant. For negative driver Brothy, both techniques recommended the optimal intensity to be 0. For positive driver Diacetyl, PLS recommended a higher intensity similar to the best liked B3. For negative driver Sulfur/cabbage, PLS recommended a lower intensity also similar to the best liked B3. For negative driver Dried fruit, both optimal intensities fell between the
least and most liked samples although PLS recommended a lower intensity than LSA. Sweaty was not identified as a driver in either technique but both recommended a low level in the optimal profile.

Predicted liking of optimal profile from PLS is 6.3.
Predicted liking of optimal profile from LSA is 6.7.
B3 = most liked sample (mean liking = 6.3)
S12 = least liked sample (mean liking = 4.2)
Chapter 5: Conclusions and Recommendations

Similar types of information were obtained from the three preference mapping techniques under investigation. Each technique provided criteria to assess the model’s adequacy, visual representation of the products and consumers in one space, product characteristics that drive liking, and optimal profiles. However, the processes by which the techniques arrived at this information were quite different. EPM constructed the product space using descriptive sensory information alone and fit consumer liking into the space subsequently. LSA constructed the product space using consumer liking alone and fit the descriptive sensory information into the space subsequently. PLS used descriptive sensory information and consumer liking simultaneously to construct the product space. Each method has its own advantages and disadvantages, each of which should be considered when designing and interpreting preference mapping studies. In this investigation, several challenges arose and as much as possible these challenges were related back to characteristics of the input data or analysis techniques in order to develop a set of recommendations that should lead to successful use of the methods.

One challenge was poor fit of consumers into the space. For EPM, one explanation was that consumers were using different criteria to determine appreciation for the products than those used to set the product space (descriptive sensory information). This was quite likely the case because the descriptive analysis was of flavor alone while consumers likely considered texture as well. Another case in which EPM would provide a poor fit of consumers would be when two products have different
profiles but are equally liked. Alternatively, it could be that consumer experience with the product category is broader than that represented in the test products. In the latter instance, the recommendation is to ensure the products chosen for the test cover the market space and previous consumer experience. Yet another contributor to poor fit of consumers is that a segment of consumers may be contributing noise to the model. A likely cause of noise is consumers who assign roughly equal hedonic ratings to all the samples. There are specific recommendations to overcome this challenge. First, clearly define the target population and screen for these consumers prior to testing. Second, consider LSA or PLS since consumer information is considered in setting the product space. Third, always look for segments of consumers (including low variance raters) at an early stage in the data analysis.

Another challenge was the poor fit of products into the space. The most common reason for poor fit in the space was the presence of binary descriptors, which are descriptors present in only one sample in the product set. A binary descriptor is difficult to interpret since there are only two known data points: 0 and the liking of the intensity that occurs in the one sample. If, for example, liking increases between the point of no intensity and the point of perceivable intensity it is tempting to conclude the descriptor is a positive driver. In fact, consumer preference for an intermediate intensity was not tested and cannot be predicted reliably. If the relationship between intensity and liking is truly linear, then the predicted intensity may be accurate; however, if the relationship between intensity and liking is non-linear, this would not be revealed. The specific recommendation for binary descriptors is to avoid them. No existing analyses, including the three techniques investigated, are robust enough to handle binary descriptors. In fact, it is difficult to imagine how any model could make a meaningful prediction with such limited information.
Another challenge is so-called outlier samples. As was just discussed, binary descriptors are commonly the reason products do not fit the space; however, consumer segmentation can also play a role. If one segment of consumers strongly prefers a product and another segment of consumers strongly dislikes a product, it will be difficult to fit the product into the space. EPM and LSA are particularly susceptible to this issue since both attempt to fit individual consumers. In the worst case, the resulting space will appear as if there is low appreciation for any of the products. Thus, it is also recommended that one always look for consumer segments at an early stage of the data analysis.

Another challenge was determining positive and negative drivers. The most definitive analysis was provided by PLS, which combined weighted variable importance with standardized regression coefficients in order to determine the drivers. While LSA used correlation to the product space to determine drivers, some interpretation was still needed to determine if drivers were positive or negative. EPM was the least definitive since drivers were determined by visual inspection and not by a more quantitative method. In fact, with EPM, researchers working independently on the same data could arrive at varying numbers of positive and/or negative drivers. Thus, PLS and LSA are preferred over EPM.

Theoretically, response surface analysis EPM should make determining optimal profiles easier. However, for this particular dataset, an optimum profile could not be determined because the resulting map was too ambiguous, highlighting an inherent weakness of all forms of EPM. Specifically, two samples can be equally liked and yet have descriptive sensory profiles located opposite each other along some sensory
dimension. In such instances, EPM cannot adequately model the relationship between liking and the descriptive sensory space. This is another reason PLS and LSA are preferred over EPM.

When comparing the difficulties that arose with all three models, it can be seen that many were specific to EPM alone. These difficulties included the poor fit of consumers to the model, difficulty in obtaining preference drivers, and difficulty in obtaining optimal profiles. These issues are almost always related to the fact that EPM sets the product space independent of the consumer. This is a fatal flaw which should not be ignored when the objective of the investigation is to understand the consumer. To meet this objective, PLS and LSA are much more appropriate.

In terms of their outputs, PLS and LSA gave similar results. Six of the seven identified drivers were the same in both models and optimal profiles differed on only a couple of attributes. Even though the optimal profiles were similar, drivers and their relationship to the most and least liked products was easier to interpret in the PLS profile. Further, PLS is conceptually preferable to LSA since consumer and descriptive sensory information are integrated simultaneously, rather than forcing descriptive sensory information into a pre-set liking space.

Although PLS is preferable, it has one major drawback. In PLS models liking data is averaged across consumers, potentially concealing consumer segmentation. There are several techniques which can be used to overcome this disadvantage. First, cluster analysis can be used to reveal segments of consumers that illustrate similar patterns of liking as well as identify consumers who give similar hedonic ratings (indicating no difference in preference) to all products. Subsequently, each of these groups of consumers can be modeled separately with PLS. Alternatively, a technique like LSA can be used to visualize density of consumers on a map. If multiple segments are identified,
consumers can be grouped, averages calculated for each group, and subsequently a
model for each group can be developed with PLS. A more statistically sophisticated (and
complicated) approach would be to use each individual consumer as an independent
variable in the regression and then pool the results. However, with this approach, a high
degree of expertise in statistical programming is needed, making it a less desirable
approach for most sensory and food science analysts.

In the end, the methods must ultimately be matched to the goals of the test.
Understanding the strengths and weaknesses of each technique and following these
recommendations (see Table 5.1) should lead to successful understanding of
consumers and their appreciation for the products under investigation.

Table 5.1. Recommendations toward successful preference mapping studies

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Reasoning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ensure products cover the market space and previous consumer experiences</td>
<td>When consumers use criteria from outside the established descriptive sensory space, drivers will be obscured.</td>
</tr>
<tr>
<td>Clearly define the target population</td>
<td>Undifferentiated consumer groups will introduce noise into the model and drivers will be obscured.</td>
</tr>
<tr>
<td>Look for consumer segments before constructing a preference model (especially for PLS)</td>
<td>Undifferentiated consumer groups will introduce noise into the model and drivers will be obscured, especially with models that average across consumers.</td>
</tr>
<tr>
<td>Avoid products with binary descriptors</td>
<td>Interpretation of drivers and optimal intensities are unreliable when only one data point is used to construct the model.</td>
</tr>
<tr>
<td>Avoid the use of EPM</td>
<td>There is no quantitative method for selecting drivers. EPM cannot model similarly appreciated but perceptually distinct products.</td>
</tr>
<tr>
<td>Use PLS over LSA</td>
<td>Interpretation of drivers and their relationship to products is easier.</td>
</tr>
</tbody>
</table>
References


Cordonnier, S. M. (2007). *Assessments of liking and factors that influence them*. The Ohio State University, Columbus, Ohio, USA.


77


