Sensory analysis of brandy: the application of rapid profiling methodologies

by

Leanie Louw

Dissertation presented for the degree of

Doctor of Philosophy (Agricultural Sciences)

at

Stellenbosch University

Department of Viticulture and Oenology, Faculty of AgriSciences

Supervisor: Prof Pierre van Rensburg

Co-supervisor: Prof Marius Lambrechts

Co-supervisor: Dr Hélène Nieuwoudt

Co-supervisor: Prof Tormod Næs

March 2014

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated) that reproduction and publication thereof by Stellenbosch University will not infringe and third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 09/12/2013

Summary

Brandy has a high alcohol content of at least 36% alcohol by volume and a complex volatile chemical structure. Due to this, the sensory evaluation of brandy is challenging to execute and must be carefully managed to avoid panellists' sensory and mental fatigue. Although rapid sensory profiling methods, such as projective mapping (PM), are commonly used for fast moving consumer goods, it was unclear whether these methods could accommodate the difficulties associated with brandy evaluation.

The primary objective of this study was to validate PM, also called Napping®, as a reliable tool for brandy evaluation. Two variations of PM were tested: global napping (GN) where products are evaluated based on overall sensory perception, and partial napping (PN) where the product perception is broken down in to separate sensory modalities, in this case appearance, aroma and in-mouth sensations. Several practical aspects of PM were investigated in order to optimise the method for brandy evaluation and to gain understanding into practical methodological aspects that have not been fully understood at the onset of this study.

The results showed that both GN and PN delivered reliable results, but that PN was more reproducible and better suited to larger sample sets (10 products). The concept of in-mouth sensations was found to be ineffective in extracting useful information on mouthfeel differences in the product set and that retronasal flavour should be separated from basic tastes and mouthfeel. A verbal instruction to the panellists was sufficient to obtain reliable information on mouthfeel differences; it was not necessary to use black glasses or nose-clips to eliminate the influence of appearance and flavour perception. In response to the insufficient reproducibility of GN, the sorting method was tested and validated as a suitable screening method that delivered reliable product maps of larger sample sets (10 products) in one replication, compared to GN which has to be replicated several times.

A recent hypothesis states that the elongated horizontal dimension of a rectangular PM tasting sheet could be used to elicit more prominent product differences, while tasting sheets without an elongated axis, such as circles or squares, could be used to bring forth more subtle differences. This hypothesis was challenged by testing rectangular, square and round tasting sheet shapes. Although differences were observed between the shapes in terms of product configuration and panellist performance, the practical value of using different tasting sheet shapes to obtain specific information could not be established.

The impact of high alcohol content and product complexity on panellist performance in PM was tested and it was concluded that product complexity did influence the panellists' performance, particularly for high alcohol products. It was also found that prior knowledge of a products' high alcohol content influenced the panellists' descriptive language to include more alcohol-related terms.

A new panellist performance measure was developed, namely the Relative Performance Index (RPI). This measure can be used to monitor trained panellists' performance in the PM task more effectively to thereby ensure reliable results. The outcomes of this study extended the brandy sensory evaluation toolbox. Practical measures were identified that can be used to overcome the challenges associated with the sensory evaluation of complex high alcohol products.

Opsomming

Met ‘n alkohol inhoud van mintens 36% alkohol per volume en ‘n komplekse vlugtige chemise samestelling is brandewyn ‘n uitdagende produk om sensories te evalueer. Maatreëls moet in plek gestel word om te verhoed dat sensoriese paneellede nie uitgeput raak nie om sodoende betroubare resultate te verseker. Hoewel vinnige sensoriese toetsmetodes, soos projeksiekartering (PM) en sortering, gereeld gebruik word vir die evaluasie van vining bewegende verbruikersgoedere was dit weens die voorafgenoemde redes onduidelik of die metodes geskik sou wees vir brandewynevaluasie.

Die oorhoofse doelwit van hierdie projek was om PM, ook genoem Napping®, te valideer as ‘n geskikte metode vir die sensoriese evaluering van brandewyn. Twee variasies van PM is ondersoek: oorhoofse kartering (GN) waar die produkte op grond van ‘n algehele sensoriese waarneming evalueer word, en gedeeltelike kartering (PN), waar die verskillende sensoriese waarnemingsmodaliteite afsonderlik gemeet word. In hierdie geval is voorkoms, aroma en algehele binne-mondse ervaring gemeet. Verskeie aspekte van GN en PN is ondersoek ten einde die metodes vir brandewynevaluasie te optimiseer asook lig te werp op praktiese metodologiese aspekte waaroor daar nog onsekerheid was by die aanvang van hierdie studie.

Die resultate het getoon dat beide GN en PN betroubare inligting oor brandewyn kan oplewer, maar dat PN meer herhaalbaar en beter geskik is vir groter produkstelle (10 produkte). Dit is egter bevind dat die konsep van *algehele binne-mondse waarneming* nie doeltreffend was om mondgevoelverskille tussen produkte uit te lig nie. Smaak- en mondgevoelwaarnemings moet liefs onderskei word van binne-mondse geurwaarneming. Dis bevind dat dit voldoende is om hierdie onderskeid deur skriftelike instruksies tot die paneel te bewerkstelling en dat die gebruik van swart glase en neusknippies nie nodig was om betekenisvolle inligting oor mondgevoel te bekom nie. Na aanleiding van die swak herhaalbaarheid van GN, is sortering getoets en gevalideer as ‘n geskikte siftingsmetode wat betroubare produkkaarte van groot produkstelle (10 produkte) kan oplewer met slegs een evaluasie, teenoor GN waarvoor meer as een sessie benodig word.

‘n Onlangse hipotese stel dat die langer horisontale dimensie van reghoekige PM proekaarte gebruik kan word om meer prominente verskille tussen produkte uit te wys, terwyl proekaarte sonder verlengde asse, soos sirkels en vierkante, gebruik kan word om meer subtiele verskille uit te wys. Hierdie hipotese is bevraagteken deur reghoekige, vierkantige en ronde PM proekaarte teenoor mekaar te toets. Hoewel die resultate vanaf die verskillende proekaarte effens verskil het in terme van produk konfigurasie en die paneel se taakverrigting, kon die praktiese waarde van die gebruik van produkkaarte met verskillende vorms om spesifieke inligting te bekom nie bevestig word nie.

Die impak van hoë alkohol en produkkompleksiteit op paneellede se PM taakverrigting is ondersoek. Dit is bevind dat produkkompleksiteit wel hul taakuitvoering beinvloed, en dat hierdie invloed groter is vir hoë alkohol produkte. Dit is ook gevind dat paneellede se beskrywende taalgebruik beïnvloed om meer alkoholverwante woorde in te sluit word indien hul kennis dra van ‘n produk se hoë alkohol inhoud.

‘n Nuwe vaardigheidsberaming vir die meet van paneellede se taakverrigting is ontwikkel, naamlik die Relatiewe Taakverrigtingsindeks (RTW). Dit kan gebruik word om paneellede se vaardigheid in PM meer doeltreffend te monitor en sodoende betroubare resultate te verseker. Die uitkomste van die studie het die poel beskikbare metodes vir die sensoriese evaluering van brandewyn suksesvol verbreed en praktiese maatreëls is geidentifiseer om die uitdagings van ‘n komplekse hoë alkohol produkte te oorkom.

Acknowledgements

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

Prof Marius Lambrechts, Distell Ltd and Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, for investing his time and resources in this project, for his continuous support and for sharing his wealth of knowledge regarding brandy production.

Prof Pierre van Rensburg, Distell Ltd and Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, for his invaluable guidance and support and, as always, for coming to the rescue in tricky situations.

Dr Hélène Nieuwoudt, Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, for her critical evaluation of my work from an academic perspective.

Prof Tormod Næs, Nofima, Norway, for sharing his expertise in sensory data analyses, for his endless patience, and for welcoming me into his home.

Dr Sulette Oelofse, Distell Ltd, for being available to soundboard practical aspects around experimental design as well as for her assistance with the panel facilitation and data analyses.

Marelize Scheepers, Distell Ltd, for supporting me in this endeavor.

Dr Martin Kidd, Deparment of Mathematics, Stellenbosch University, for his assistance with some of the data analyses.

Muriel Adams, Distell Ltd, for her assistance with the preparations in the sensory laboratory and for measuring massive amounts of projective mapping coordinates.

The Distell Sensory Panel, for their hard work and perseverance, but also for their friendship and support.

Karin Vergeer, Institute for Winebiotechnology, Department of Viticulture and Oenology, Stellenbosch University, for her assistance with administrative matters and efforts in the preparation of this manuscript.

Distell Ltd. for funding.

My friends and family, for their patience, support and understanding.

The Heavenly Father, for keeping me safe and grounded.

Preface

This dissertation is presented as a compilation of eight chapters. Each chapter is introduced separately and is written according to the style of the journal Food Quality and Preference in which Chapter Three and Chapter 7 were published. Well known Latin abbreviations are not italicized, according to the journal style.

Chapter 1 **General Introduction and project aims**

Chapter 2 **Literature review**
Brandy sensory evaluation and rapid sensory profiling

Chapter 3 **Research results**
Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products

Chapter 4 **Research results**
Optimisation of the partial napping approach for the successful capturing of mouthfeel differentiation between brandy products

Chapter 5 **Research results**
Rapid sensory evaluation of brandy: Can sorting be used as a one session screening tool?

Chapter 6 **Research results**
The effect of tasting sheet shape on product configurations and panellists' performance in sensory projective mapping of brandy products

Chapter 7 **Research results**
Trained sensory panellists' response to product alcohol content in the projective mapping task: observations on alcohol content, product complexity and prior knowledge

Chapter 8 **General discussion and conclusions**

Table of Contents

Chapter 1

Introduction and project aims

Chapter 1: Introduction and Project Aims

1.1 Introduction

Brandy is a distilled spirit made from fermented grapes and is consumed in many countries across the world, including France, Spain, Chile, Portugal and South Africa. The product has a complex volatile composition consisting of compounds from various chemical families: alcohols, carboxylic acids and their esters, aromatic compounds, furan derivatives, ketones and terpenes (Janáčová et al., 2008). A large number of these compounds are present at odour-active concentrations, contributing to a wide range of perceivable aroma nuances as is illustrated on the South African brandy aroma wheel (Jolly and Hattingh, 2001). Its high alcohol content, a minimum of 36% alcohol by volume (ABV), (European Union, 2008) adds further to complexity.

In South Africa, the largest volume share proportion of the spirit market traditionally belonged to the brandy category. However, the position of brandy as the top selling spirit product in South Africa is seriously threatened by the growing popularity of imported whiskies. In order to regain and protect brandy's top position in the South African spirit beverage market, it is important to understand consumer's needs and expectations of spirit products. Part of the process of gaining such understanding is the use of sensory research to ensure that consumers' taste expectancies are met (Piggot, 2012).

The complexity and high alcohol content of brandy make sensory evaluation of the product very challenging. The combination of sensory fatigue induced by high concentrations of ethanol, a chemosensory irritant, and mental fatigue induced by the perceived product complexity, means that sensory evaluation of brandy must be strictly controlled to ensure valid results. For conventional profiling, several practices such as sample dilution, expectoration after evaluation, forcing panellists to rest between samples with controlled time delays and reduction of sample set sizes through incomplete block designs, have been put in place to control the onset of fatigue during sensory evaluation (Louw & Lambrechts, 2012). The latter is especially useful for studies investigating larger sample sets.

An aspect of spirit beverage sensory evaluation that has remained unexplored is the application of rapid sensory profiling methods, such as projective mapping (Risvik et al., 1994), sorting (Lawless et al., 1995) and flash profiling (Dairou & Siefferman, 2002). These methods are increasingly being used to understand food and beverage products' sensory characteristics, as well as consumers' responses thereto. Part of the appeal of rapid sensory profiling methods is their time and cost-effectiveness compared to conventional sensory profiling methods (Valentin et al., 2012). In addition to this, rapid sensory profiling methods have been found useful to explore complex sample attributes that are pertinent in a product, but not easily measurable with conventional descriptive profiling due to lack of consensus on their definitions (Albert et al., 2011). At the onset of this study, it was unclear whether rapid sensory profiling methodologies would be suitable for brandy evaluation, since previous reports have suggested that product complexity may influence panel reliability (Delarue & Siefferman, 2004; Nestrud & Lawless, 2010). Furthermore, rapid sensory profiling methods rely on comparative product presentation where all products are presented to the panellist simultaneously. This limits the measures that can be taken to reduce sensory fatigue especially in larger sample sets. To date, rapid sensory profiling methods have not been validated for spirit beverages. Without rapid sensory profiling methods, the range of tools for the sensory profiling of spirit beverages is very small, limiting the versatility with which sensory analysis can be applied to brandy studies.

Different types of rapid sensory profiling methods have been developed that are suitable for consumer (Dooley et al., 2010), expert (Pagès, 2005) and/or trained panels (Ross et al., 2012). The most common rapid sensory profiling methods adapted for trained panels are projective mapping (Risvik et al., 1994), sorting (Lawless et al., 1995) and flash profiling (Dairou & Sieffermann, 2002). Projective mapping and sorting are holistic methods based on the assessment of the overall sensory similarity between samples as perceived by the panellist. Compared to other rapid profiling methods, these methods are more intuitive and representative of the way in which consumers interact with products. Both projective mapping and sorting have been reported to deliver reliable results when applied to wine and beer (Chollet et al., 2011; Perrin et al., 2008). Projective mapping provides product specific information and information regarding degree of difference between samples, while sorting provides descriptive information on groups of products only, and no indication of the degree of difference between groups (Valentin et al., 2012; Varela & Ares, 2012). Flash profiling is an attribute-based method which has been found to be more discriminative than projective mapping and sorting, with a higher yield in interpretable descriptive information (Albert et al., 2011; Veinand et al., 2011). However, flash profiling requires more sessions and panellists are more susceptible to sensory fatigue due to the need for re-tasting (Veinand et al., 2011), making it potentially less suitable for brandy evaluation.

Projective mapping has been widely accepted as a valid and flexible method. The method involves panellists to evaluate samples and position them on a tasting sheet according to perceived similarity. There are still many practical aspects of the method that may impact on the overall results, as well as on panellists' performance in the task. Of specific interest is the shape of the tasting sheet, which has, to date, not been standardised. Recent investigations showed that panellists are influenced by the tasting sheet geometry, although it is unclear whether this has a substantial effect on the overall results, and if so, what the effect is (Dehlholm, 2013; Hopfer & Heymann, 2013).

Performance differences among individual panellists and poor panellist consistency have been reported for projective mapping (Kennedy, 2010; Risvik et al., 1997). The factors influencing panellists' performance in projective mapping, especially in terms of alcoholic beverages, are not fully understood. It is likely that product complexity and high alcohol content are both factors that impact on panel performance in projective mapping, but the relative importance of these two indicators has not been established. Understanding the factors that impact on panellists' performance allows the analyst to put measures in place to ensure optimal panel performance. This would increase the potential of rapid sensory profiling methodologies to become work-horse methodologies for explorative sensory evaluation of alcoholic beverages.

1.2 Project aims

Based on these considerations, the main aim for this study is to investigate the application of projective mapping to spirit beverages as an effective rapid sensory profiling tool for the brandy industry. The following five research goals have been identified:

1. The validation of projective mapping and variations thereof for brandy evaluation.
2. The establishment of a suitable rapid sensory profiling method for large sample sets of brandy for screening purposes.
3. The investigation of the role of product complexity and product alcohol content on panellist performance in projective mapping.

4. Further investigation of the influence of projective mapping tasting sheet geometry on overall results and panellist performance.
5. The identification of key control measures for monitoring panel performance and ensuring good quality data.

1.3 References

Albert, A., Valera, P., Salvador, A., Hough, G., & Fiszman, S. (2011). Overcoming the issues in the sensory description of hot served food with a complex texture. Application of QDA®, flash profiling and projective mapping using panels with different degrees of training. *Food Quality and Preference, 22,* 463-473.

Chollet, S., Lelièvre, M., Abdi, H., & Valentin, D. (2011). Sort and beer: Everything you wanted to know about the sorting task but did not dare to ask. *Food Quality and Preference*, *22,* 507-520.

Dairou, V., & Sieffermann, J.M. (2002). A comparison of 14 jams characterized by conventional profile and a quick original method, flash profile*. Journal of Food Science*, *67,* 826-834.

Delarue, J., & Siefferman, J.M. (2004). Sensory mapping using Flash Profile. Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. *Food Quality and Preference*, *15,* 383-391.

Dehlholm, C. (2013). *Descriptive sensory evaluations: Comparison and applicability of novel rapid methodologies.* PhD Thesis, ISBN 978-87-7611-592-0.

Dooley, L., Young-seung, L., & Meullenet, J.F. (2010). The application of check-all-that-apply (CATA) consumer profiling to preference mapping of vanilla ice cream and its comparison to classical external preference mapping. *Food Quality and Preference, 21,* 394-401.

European Union (2008). Regulation (EC) No 110/2008 of the European Parliament and of the Council of 15 January 2008. *Official Journal of the European Union*, L 39/30-21.

Hopfer, H., & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. *Food Quality and Preference, 28,* 164-181.

Janáčová, A., Sadecka, J., Kohajdova, Z., & Spanik, I. (2008). The identification of aroma-active compounds in Slovak brandies using GC-sniffing, GC-MS and sensory evaluation. *Chromatographia, 67,* S113-S121.

Jolly, N. P., & Hattingh, S. (2001). A brandy aroma wheel or South African brandy. *South African Journal of Enology and Viticulture, 22,* 16-21.

Kennedy, J. (2010). Evaluation of replicated projective mapping of granola bars. *Journal of Sensory Studies, 25,* 672-684.

Lawless, H.T., Sheng, N., & Knoops, S.S.C.P. (1995). Multidimensional scaling of sorting data applied to cheese perception. *Food Quality and Preference, 6,* 91-98.

Louw, L., & Lambrechts, M.G. (2012). Grape-based brandies: production, sensory properties and sensory evaluation. IN: *Alcoholic beverages: sensory evaluation and consumer research* (Ed. J. Piggot). Cambridge: Woodhead Publishing Limited.

Nestrud, M.A., & Lawless, H.T. (2010). Perceptual mapping of apples and cheeses using projective mapping and sorting. *Journal of Sensory Studies*, *25,* 390-405.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire Valley*.* *Food Quality and Preference, 16,* 642-649.

Perrin, L., Symoneaux, P., Maitre, I., Asselin, C., Jourjon, F., & Pages, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: Case of ten wines from Loire valley. *Food Quality and Preference, 19,* 1-11.

Piggot, J. (2012). *Alcoholic beverages: Sensory evaluation and consumer research*. Cambridge: Woodhead Publishing Limited.

Risvik, E., McEwan, J.A., Colwill, J.S., Rogers, R., & Lyon, D.H. (1994). Projective mapping: a tool for sensory analysis and consumer research. *Food Quality and Preference, 5,* 263-269.

Risvik, E., McEwan, J.A., & Rødbotten, M. (1997). Evaluation of sensory profiling and projective mapping data. *Food Quality and Preference*, *8,* 63-71.

Ross, C.F., Weller, K.M., & Alldredge, J.R. (2012). Impact of serving temperature on sensory properties of red wine as evaluated using projective mapping by a trained panel. *Journal of Sensory Studies*, *27,* 463-470.

Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012). Quick and dirty but still pretty good: a review of new descriptive methods in food science*. International Journal of Food Science and Technology*, *47(8),* 1563-1578.

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International*, *48,* 893-908.

Veinand, B., Godefroy, C., Adam, C., & Delarue, J. (2011). Highlight of important product characteristics for consumers. Comparison of three descriptive methods performed by consumers. *Food Quality and Preference, 22,* 474-485.

Chapter 2

Literature review

Brandy sensory evaluation and rapid sensory profiling

A Section of this manuscript were published as
Grape-based brandies: production, sensory properties and sensory evaluation
In: *Alcoholic beverages: sensory evaluation and consumer research*
Ed: J Piggot.

Chapter 2: Brandy sensory evaluation and rapid sensory profiling

2.1 Chapter overview

This review is a discussion of the sensory profiling methods that are currently being used to evaluate brandy as well as the methods that could potentially be applied but have not been validated yet for high alcohol products. The first section of the review is focused on the current state of brandy sensory evaluation and has been published in *Alcoholic beverages: sensory evaluation and consumer research (2012).* At the end of this section, the potential of rapid sensory profiling (RSP) methods for brandy evaluation is highlighted, setting the scene for the second section on RSP methodologies. This section starts off with an overview of RSP during which three types of RSP methods will be introduced. The remainder of the manuscript will be focused on similarity-based methods, specifically projective mapping (PM) and sorting. It will consist of a synopsis of common data analysis tool for these methods, in-depth discussion of the development of each method, important findings and applications thereof. The section will be concluded with a critical review of the typical features of studies on PM and sorting and the identification of gaps in the literature.

2.2 Defining brandy[1]

The word 'brandy' is derived from the Dutch word 'brandewijn' which can be literally translated as 'burnt wine'. This refers to brandy being a distilled spirit originating from fermented grapes. For a grape-based spirit to be considered a brandy, it must have been aged for at least 6 months in oak casks smaller than 1000L in capacity or 12 months with oak receptacles. However, the term brandy is sometimes incorrectly used to refer to other fruit-based spirits e.g. plum brandy. Another variation on brandy is the so-called pomace brandies that are produced from the grape skins, a by-product of wine production e.g. grappa. (Robinson, 1999). Several European nations produce brandies, including France, Germany, Bulgaria, Italy, Portugal and Cyprus. Brandy is also produced in the New World wine producing countries such as the USA and Australia. Each production area has its own production regulations and some brandies of origin are very strictly controlled (Robinson, 1999). Well-known and protected styles of brandy include the Slovak 'Karpatske brandy special', Portuguese Lourinhã brandy, Spanish Brandy de Jerez, Wachauer Weinbrand and Weinbrand Durnstein from Austria as well as the Deutcher Weinbrand and Pfälzer Weinbrand from Germany and also South African brandies, Cognac, Armagnac and South American Pisco.

2.3 Sensory evaluation of brandy

2.3.1 Practical considerations for dealing with high alcohol beverages

The high alcohol content of brandy increases the risk of sensory fatigue. To reduce this risk, samples are usually diluted with odourless water to an alcohol content of ranging from 20% to 23% ABV prior to sensory analysis (Du Plessis et al., 2002; Peña y Lillo et al., 2005). Brandy samples are ideally served at room temperature. The dilution of an alcoholic liquid with water causes an exothermic reaction and

[1] Section 2.2 and 2.3 have been published in: Louw, L., & Lambrechts, M.G. (2012). Grape-based brandies: production, sensory properties and sensory evaluation. IN: *Alcoholic beverages: sensory evaluation and consumer research* (Ed J Piggot). Cambridge: Woodhead Publishing Limited. p292-294.

therefore a short resting period should be allowed between dilution and serving. It was found that the pungency of whisky samples increase at higher temperatures (Jack, 2003). Although the effect of temperature on brandy aroma has not been tested, it is possible that the volatility of brandy aromas may also be affected by rises in temperatures. Even so, it has been suggested that brandy samples may also be nosed at full strength, as the aldehydes are more prominent at this alcohol strength, while fatty acids and esters are more prominent at approximately 20% ABV (Quady & Guymon, 1973).

Another method of reducing sensory fatigue is to evaluate samples by nosing only. A study conducted on Pisco evaluation (which per definition also qualifies as a brandy) showed that there was a strong correlation between orthonasal and retronasal odour perceptions. The authors suggested that the olfactory analysis of Pisco can be conducted based on orthonasal perceptions only. By excluding the ingestion of the product, sensory fatigue can be limited (Peña y Lillo et al., 2005). However, some gustatory attributes such as mouth coating can differ significantly between brandies and may be interesting to measure. Palate cleansers such as spring water and soda biscuits (Peña y Lillo et al., 2005) or just water (Caldeira et al., 2010) can be used when brandy products are tasted for gustatory attributes.

The question of the number of samples to be served in one sitting is often raised during the sensory evaluation of alcoholic beverages. The risks of sensory adaptation, fatigue and inebriation that are coupled to large sample sets can reduce the quality of the responses obtained. It is difficult to determine a maximum number of samples that can be tasted; the ease with which the panel will be able to evaluate the samples will depend on the extent of the differences between the samples, the complexity of the lexicon as well as the level of training of the panellists. However, reports on the descriptive sensory analysis of brandy vary from four to twelve samples, both full strength and diluted (Caldeira et al., 2002, 2006, 2010; Du Plessis et al., 2002; Janáčová et al., 2008; Van Jaarsveld et al., 2009). Du Plessis et al. (2002) also reported the evaluation of brandy distillate samples at 23% ABV in a series of 12 triangle tests, amounting to 36 distillate samples. The panel were able to detect significant differences in 9 of the 12 triangle tests which suggest that they were not affected by sensory fatigue.

In small sample sets, comparative sample presentation can be considered instead of monadic sequential presentation. Monadic sequential tests are generally preferred because they provide more independent responses than comparative sample presentations, in addition to providing better control over time delays between samples (Stone & Sidel, 2004). With comparative sample presentations, the evaluation process remains essentially the same as with monadic sequential presentation, i.e. the taster evaluates each sample completely before moving on the next sample. However, the taster has the advantage of having the entire sample range in front of him and can thereby orientate himself in the sensory space represented by the sample set. This may cause the taster to feel more in control of the task when presented with the entire sample set, which may provide a subtle psychological advantage in a difficult sample set. It has been observed on several occasions that more reliable data can be obtained from comparative sample presentation than with sequential monadic presentation of brandy, provided that the sample set is not too large, i.e. five or six products at most.

2.3.2 Evaluation of brandy colour

Final brandy samples can differ in colour as brandy colour is rectified with flavourless caramel to obtain the desired hue and intensity. Wood maturation can also affect brandy colour. Typical brandy colour hues include straw-yellow, golden, topaz and greenish (Canas et al., 2009). In some research scenarios it may be beneficial to remove the bias caused by colour perception. Blue glasses can be used to deal with this problem (Du Plessis et al., 2002; Peña y Lilo et al., 2005).

2.3.3 Evaluation of brandy quality

Apart from the usual sensory analyses focussing on single attributes as perceived by one of the five senses, brandy is sometimes also evaluated for organoleptic quality. Reported quality measurements include 20 point scales (0=without quality to 20=maximum quality) and 6-point category scales (0=completely unacceptable to 5=outstanding quality) (Caldeira et al., 2010; Steger & Lambrechts, 2000). Quality scores can be correlated to sensory attributes. For example, significant correlations (r > 0.7) have been found between Lourinhã brandy quality and toasted, woody, vanilla, body, flavour persistence and spicy. Furthermore it was found that flavour complexity has a highly significant impact on brandy quality. A negative correlation between brandy quality and the tails, green and caoutchouc (rubber) characters were reported by the same authors. Bitterness and burning did not appear to have an impact on quality (Caldeira et al., 2006).

2.3.4 Sensory evaluation of brandy over extended periods of time

Brandy can be matured for anything from six months to twenty years, which can mean that brandy evaluation projects can last for several years. One example is the study by Caldeira et al. in 2006 that reported on flavour evolution that takes place in brandy over five years of maturation. Ideally, one would want the trained panel that participates in the evaluation to be constant to ensure consistent results. However, in reality this is not always possible.

It is therefore very important that the sensory language used to describe brandies remains constant within a working environment. The existence of a common flavour language will ease the incorporation of new brandy panellists. A common flavour language will also make it easier to relate results from independent studies within a working environment. The Brandy Aroma Wheel for South African brandies (Jolly & Hattingh, 2001) provides a standardised list of descriptors that can also be used for brandies from other countries (Figure 2.1). The reference standards act as a supplement to the aroma wheel and can be used as a tool to align the word-flavour associations between panellists. Unfortunately, many of the references were commercial flavour keys from suppliers that have merged with other companies since the publication of the article. The most important factor is to implement a standardised descriptor list in a working environment with appropriate reference standards. References standards can be replaced by alternatives by matching the alternative with the control, as long as the reference standard provokes the same word-aroma association.

Panel performance monitoring is vital during long term sensory studies. Panellists’ reliability can be tested by including blind replicates into a training session or by using statistical tools such as Tucker plots or Pearson’s correlations (Caldeira et al., 2006; Tomic et al., 2010). Regardless of which method is used, it is very important to monitor panel performance on a regular basis to ensure that the quality of results remain consistent over time.

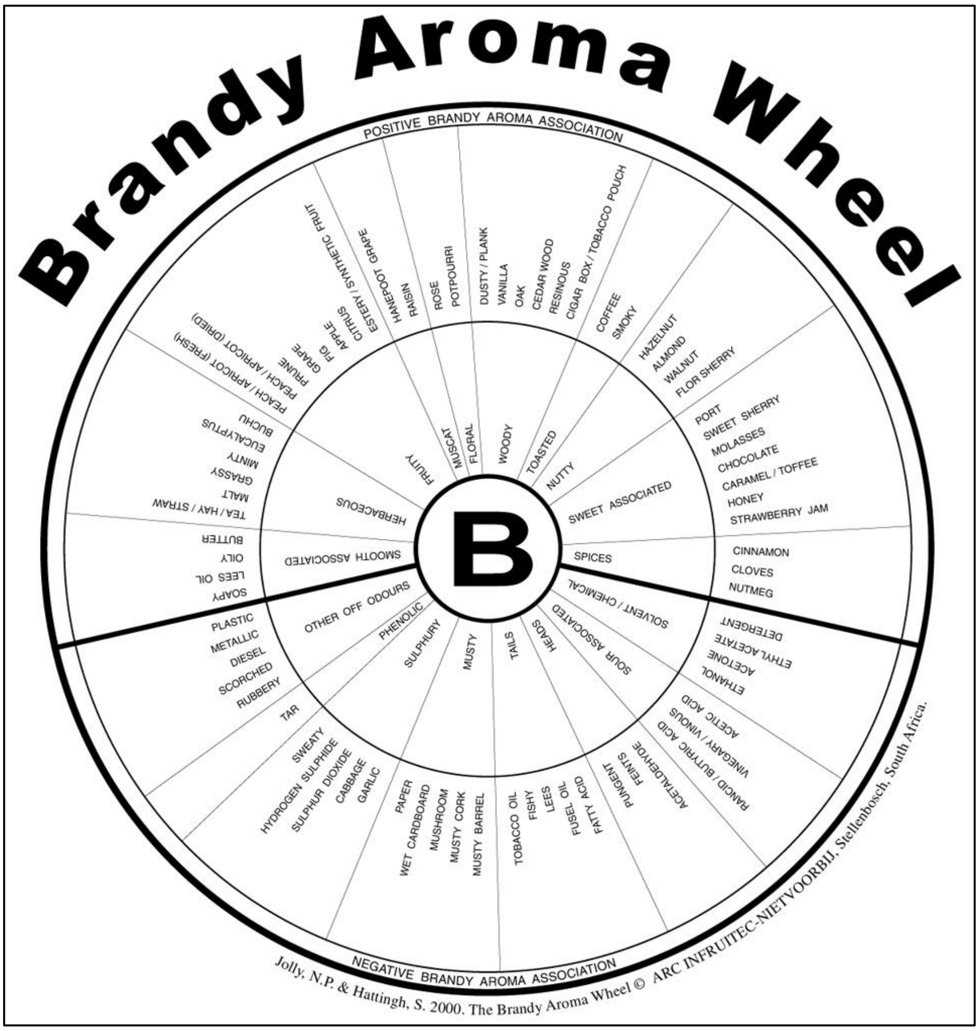

Figure 2.1 Aroma wheel developed for South African brandies (Jolly & Hattingh, 2001).

2.3.5 Brandy and consumer research

Published data on quantitative consumer research regarding brandy intrinsics are not common as it is usually funded by industry. Brandies are sometimes included in conceptual consumer studies testing for drinking habits of alcoholic beverages in general (Scriven et al., 1989). There are several possible reasons why brandies may not commonly be subjected to consumer preference or acceptability testing. The limitations in sample servings due to the high alcohol content may be one reason. Another may be that companies and research institutes are reluctant to invest in consumer research due to the long term implications of product re-formulation.

Nevertheless, should intrinsic consumer research be conducted on brandy, the basic principles of best sensory practices should still be taken into account. For consumer testing, samples should generally be presented in such a way to mimic as closely as possible the way the consumer would normally consume the product (Meilgaard et al., 2007). In the case of blended brandies that are rarely consumed neat, it may be necessary to present the product in a soft drink mixer with ice. The question of carriers can become quite complicated as consumers in the same target market may have different preferences for mixers e.g. cola, ginger ale or even orange juice. Of course, the dilution that occurs with ice may be a further complication. On the one hand, the consumer may not be able to relate to a neat, blended brandy at room temperature, which may mean that acceptability scores may be skewed towards the negative part of the scale. A further consideration is the time of day at which the tasting is conducted.

The tasting of spirit products is generally not recommended for early in the morning (Meilgaard et al., 2007).

2.3.6 Future trends

There is massive scope for sensory research in brandy, judging by the limited number of publications in this area. One area of interest would be the application of rapid profiling techniques such as flash profiling (Dairou & Sieffermann, 2002) and Napping® (Perrin, et al., 2008). These techniques are valuable screening tools with wide-spread applications in industrial environments. However, the effect of the high alcohol content of brandy on the effectiveness and reliability of such techniques must be investigated.

2.4 Rapid sensory profiling methods

2.4.1 An overview of rapid sensory profiling

The aim of sensory profiling is usually to obtain a two-dimensional configuration, or product map, that visualises the perceived differences in a set of samples. For such a map to be useful, it must comply with as many as possible of the following requirements (Lawless et al., 1995):

i) it must have a good fit to the data
ii) blind duplicates should plot close together
iii) the map should communicate the descriptive attributes and consumer preferences associated with the product set
iv) the map should be able to be plotted using only a few dimensions
v) it should be interpretable
vi) it should have a pay-off by either suggesting a new hypothesis or confirming a previous one

Conventional sensory profiling methods, such as QDA™ (Stone et al., 1974), can successfully be used to obtain perceptual maps compliant to these criteria. Such methods involve familiarising and training a panel to consistently rate the intensity of a defined set of attributes perceivable in a product. Due to the often extensive training phase, these methods can be cost- and time intensive.

RSP methods have been developed as fast and cost-effective complementary methods to conventional sensory profiling with the aim of establishing a more efficient link between trained sensory panels and consumer perceptions (Risvik et al., 1994). As it became clearer that RSP methods comply with the above mentioned criteria, it was proposed that RSP methods could also be used as alternatives to conventional profiling methods (Cartier et al., 2006; Dehlholm et al., 2012a).

In a broad sense, RSP methods are based on the principle of generating a multivariate map of the sensory differentiation in a product set, similar to that of conventional profiling, but without, or at least with very little, panel training (Valentin et al., 2012). By eliminating the training phase, the overall time spent on collecting data can be reduced from between 20 and 10 hours (usually spread over several weeks) to less than an hour (Dehlholm et al., 2012a; Johnson et al., 2013). However, the time and cost efficiency of RSP means sacrificing defined and quantified sensory attributes on which the product configuration is based. Some methods, such as flash profiling, are based on single attributes giving an indication of relative intensity, but these are not clearly defined and differ from panellist to panellist (Albert et al., 2011). In similarity-based methods such as PM and free sorting, the sensory attributes explaining the differentiation between samples are based on frequency of citation. In these cases, large amounts of undefined, idiosyncratic sensory descriptors are generated that the analyst need to process in order to interpret the product maps (Dehlholm et al., 2012a).

2.4.2 A summary of three categories of rapid sensory profiling methods

Many different RSP methods exist, each with their specific benefits and limitations. Two excellent review papers are available that provide very comprehensive reports on the theory, implementation, advantages and disadvantages of rapid profiling techniques (Varela & Ares, 2012; Valentin et al., 2012). RSP methods can roughly be categorised into three main categories. The first are attribute-based and include flash profiling (Dairou & Siefferman, 2002) and Check-All-That-Apply (CATA) (Adams et al., 2007). A second category consists of methods based on pre-selected reference samples or anchors: polarised sensory positioning (Teillet et al., 2010), pivot profile© (Thuillier et al., 2007) and recently, also polarised PM (Ares et al., 2013). The third category, which will be the focus of this review, is based on an overall appraisal of sample similarity. These methods include PM (also known as Napping®) (Pagès, 2003 and 2005; Risvik et al., 1994), sorting (Lawless et al., 1995) and variations thereof e.g. partial napping (Pfeiffer & Gilbert, 2008) and multiple free sorting (Dehlholm et al., 2012a).

In brief, PM requires panellists to position a set of samples on a sheet of paper according to perceived similarity, where very similar samples are positioned close together and very different samples are positioned far apart (Figure 2.2). An optional step requires panellists to provide words to describe the differences between the samples. The coordinates of each sample's position are collected, compiled and analysed to generate a consensus sample configuration (Figure 2.3). The descriptors, collected on a citation basis, are used to explain the dimensions of the consensus configuration. PM studies typically involve panels of 8-20 tasters, tasters testing between 8 and 13 products.

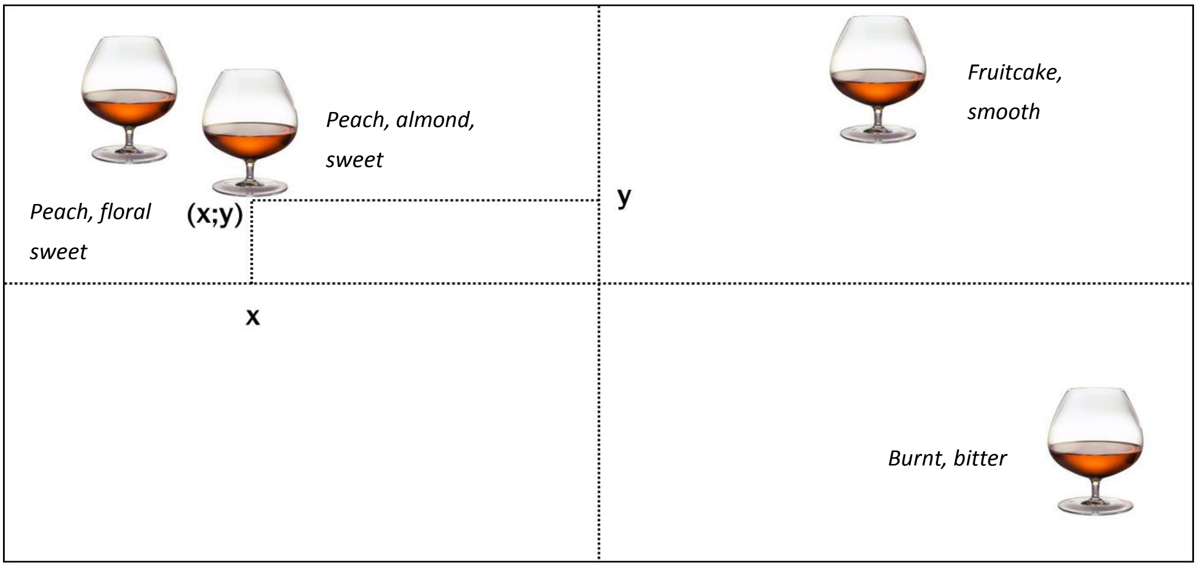

Figure 2.2 PM tasting sheet with coordinates and descriptive terms (own data).

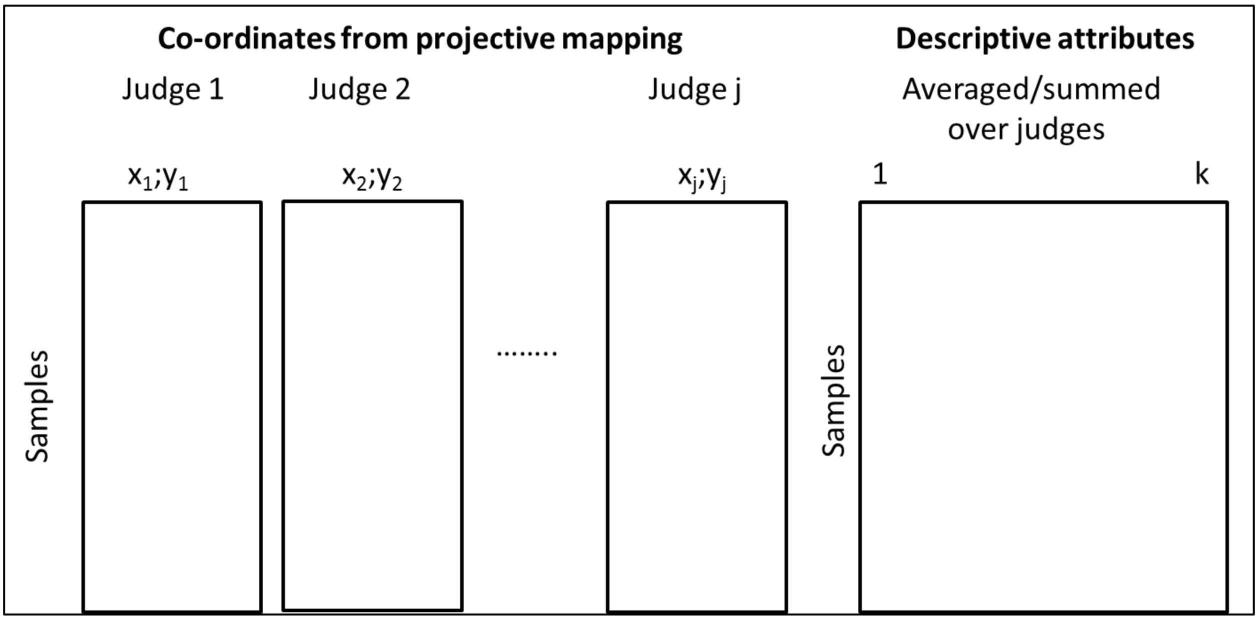

Figure 2.3 Data matrix for PM data. The data consists of J two-dimensional data matrices containing the x-and y-coordinates of J judges. A supplementary data table for K descriptors can be added. The descriptors can consist of intensity ratings averaged over judges or the sum of the descriptive citations from each judge.

Free sorting requires panellists to simply sort samples into groups according to similarity and to provide words to describe each group. In order to generate a sample configuration to visualise the similarities and differences between samples, distance matrices based on co-occurrences in groups are used as a basis for data analysis. Sorting studies typically involve between 10 and 20 products. Sorting panels are usually larger than PM panels with around 20-30 panellists.

2.5 Studies on the development, optimisation and application of similarity-based rapid sensory profiling methodologies

Both PM (Risvik et al., 1994) and sorting (Lawless et al., 1995) originate from the fields of psychology and social sciences (Oppenheim, 1996; Miller et al., 1986). Since PM and sorting have been established in the mid-nineties, many studies have focused on validating these methods for the evaluation of foods and beverages (Albert et al., 2011; Chollet et al., 2011; Dehlholm et al., 2012a; Falahee et al., 1997; Nestrud & Lawless, 2010). Most studies have involved direct method comparisons, including data obtained from conventional profiling. Studies have also focused on determining whether panel training affects results in terms of overall product configurations as well as repeatability, accuracy and reproducibility (Barcenas et al., 2004; Chollet et al., 2011; Dehlholm et al., 2012a; Kennedy, 2010). Since the mid-2000’s there has been an increase in publications using RSP methods to address research questions unrelated to methodologies. Examples include research towards understanding concepts such as minerality and astringency in wine as well as understanding drivers behind consumer acceptability of products (Ballester et al., 2005; Carrilo et al., 2012a; Gawel et al., 2001). However, there are still many questions regarding the factors that affect the effective implementation of RSP methods and many of the recent methodology orientated papers aim at answering these questions (Chollet et al., 2011; Dehlholm, 2013; Hopfer & Heymann, 2013).

This section is focused on the development of PM and sorting. It starts with an overview of the most common data analyses tools used in PM and sorting studies. Following this, the development of these methods will be briefly described. Subsequently, the key findings from studies towards the optimisation of PM will be detailed, after which the same will be done for the discussion of sorting in section 2.5.3.

2.5.1 Common data analyses tools used in projective mapping and sorting studies

2.5.1.1 Generation of consensus product maps

Due to the complex multivariate nature of RSP methods, advanced multivariate data analysis tools are needed to analyse the results in order to produce interpretable product configurations. The descriptive data generated by similarity-based methods can require considerable pre-processing. The analyst need to compile a list of descriptors used, identify and pool words with similar meanings and remove terms that are infrequently used (Moussaoui & Varela, 2010; Cartier et al., 2006).

The basis of data analysis for these methods is to find a consensus configuration based on the preliminary configuration from each panellist. The most common methods involve finding patterns in attribute correlations of several blocks of data. For PM, these include multiple factor analysis (MFA) (Pagès, 2005), the INDSCAL method (Barcenas et al., 2004), general procrustus analysis (GPA) (Risvik et al., 1994) and procrustes multiple factor analysis (PMFA) which incorporates procrustes rotation into the MFA analysis (Kennedy, 2010). Multidimensional scaling (MDS) was investigated as an alternative data analysis protocol to GPA and uncovered more underlying dimensions of differentiation than GPA (King et al., 1998). Despite this, the application of MDS for PM data analyses has not been reported since. Sorting data are most commonly analysed with MDS (Lawless et al., 1995), but have also been subjected to DISTATIS (Abdi et al., 2007) and correspondence analysis (CA) (Bouteille et al., 2013). The FAST method, based on multiple correspondence analysis (MCA) and MFA, was proposed to obtain optimal representation of all the panellists and all the samples in the final consensus configuration (Cadoret et al., 2009).

2.5.1.2 Identification of similar samples

After obtaining a product configuration from PM or sorting, cluster analysis is often applied to identify similar samples (Albert et al., 2011; Cartier et al., 2006; Nestrud & Lawless, 2010). Confidence ellipses, based on various bootstrapping methods, have also been used to visualise product confidence intervals in sensory results (Abdi et al*.,* 2009; Cadoret & Husson, 2013; Dehlholm et al*.,* 2012b; Santosa et al*.,* 2010).

2.5.1.3 Evaluation of reliability and stability of results

In order to ensure valid results, it is necessary to evaluate the reliability and stability of the product configurations. Most commonly, the RV coefficient is applied to evaluate similarity between two configurations, be it between replicated sessions, between methods or between panellists (Abdi, 2007; Robert & Escoufier, 1976). The Pearson's correlation coefficient between the inter-sample distances of two product configurations has also been used as a measure of similarity (Cartier et al., 2006).

Another check-point includes the use of duplicate samples. The Euclidean distance between two duplicate samples, relative to the maximum inter-sample distance in the product configuration, can be used as a tool to determine accuracy. This measure has been suggested as a panel performance monitoring tool for PM and has been called the People Performance Index (PPI) (Bertucciolli, 2011; Hopfer & Heymann, 2013). The same ratio has also been referred to as $D_{r\%}$, which is perhaps more appropriate, since the index is not limited to the accuracy of individuals, but can also be used to evaluate the accuracy of an overall product configuration (Torri et al*.,* 2013).

Blancher et al. (2012) proposed a resampling method to assess the stability of sorting task results. This approach is based on the idea that a method can be considered stable if the product map can be reproduced under the same circumstances. Perhaps this definition of reproducibility is too narrow; a truly reproducible method would also deliver stable results when performed by different panels.

2.5.1.4 Evaluation of panellists' projective mapping and sorting strategies

Identification and understanding of clusters of panellists with different sorting strategies can be important. Several possible approaches have been proposed: DISTATIS (Abdi et al., 2007), SORT-CC (Qannari et al*.,* 2010) as well as a technique involving a hierarchical clustering strategy based on the Adjusted Rand Index combined with a partitioning algorithm around the consensus partitions of the various clusters (Courcoux et al*.,* 2013). In projective mapping studies, the X- and Y-coordinates of individual panellists' are compared as well as the physical distribution of samples of each panellist's projective mapping sheet in order to extract pertinent information on trends in the data (Dehlholm, 2013; Hopfer & Heymann, 2013; Pagès, 2005)

2.5.1.5 Handling of missing data

Recently, the regularised iterative MFA was developed as a strategy to handle missing data during MFA (Husson & Josse, 2013). The study showed that interpretable product configurations could be obtained from incomplete PM and sorting datasets, creating the possibility of using incomplete block designs to accommodate larger sample sets. However, the missing values were generated by removing values from a complete set, and not from a dataset with true missing values where all the products were not evaluated by all panelists, as would be the case in an incomplete block design. It is uncertain whether the validity of the method, as reported by the authors, would hold in a real-life scenario, since the data used in the study were based on evaluations where the panellists had the opportunity to see and compare all the products. Although all the data points were not included in the computation, it was based on a holistic evaluation. Since PM and sorting are both based on evaluating a set of samples relative to one another, it is uncertain whether the panellists approach to the samples would change if the sample set itself changed. For example, if one sample that were very differentiating based on a fruity note were removed from a sample set, the panelist may cease to view fruitiness as an important discriminating factor. This may not be a nail in the coffin for the regularized iterative MFA; Kennedy and Heymann's study on PM of chocolates showed that three panels delivered comparable results even though only 9 out of 12 samples presented were the same over all three panels (Kennedy & Heymann, 2009). The performance of regularised iterative MFA is dependent on the strength of the relationships between variables, which in the case of PM would be the panellists' coordinates. In PM, a high degree of individual differences has been observed by several authors (Kennedy & Heymann, 2009; Risvik et al., 1997) and it is unlikely that good results will be obtained with a high number of missing values. Nevertheless, the strategy provides a possibility to overcome the most important shortcoming of both PM and sorting, namely the need to serve all samples at the same time.

2.5.2 Development and optimisation of projective mapping

This section starts with a summary of the development of the basic form of PM to the way it is used at present. Key findings from research identifying practical methodological aspects will be discussed. An overview of variations on the PM method will be provided. Finally, a synopsis of the studies where PM has successfully been applied will be presented.

2.5.2.1 First feasibility assessment of projective mapping for application to food products

The first report of PM entailed the evaluation of five chocolates which, as the authors admitted is a small sample set (Risvik et al., 1994). A trained panel performed three replicates, the data of which were analysed with GPA. The study showed excellent repeatability (RV>0.9) (Risvik et al., 1994). In this specific study, PM appeared to be more consistent than conventional profiling. The authors suggested that the two-dimensional configuration obtained with PM was based on two-dimensional data, while in the case of conventional profiling, the two-dimensional space was approximated from a multidimensional variable space, allowing more room for variation between replicated sessions. Although this study was the first to show the potential of PM as a tool for sensory analysis and consumer research, the combination of a small sample set size and a trained panel, may have given overly optimistic results. In a follow-up study, by the same authors, these shortcomings were addressed by testing the method on a slightly larger (n=7), but more homogenous sample set (Risvik et al., 1997), this time using a small, untrained, consumer panel (n=8). The consumers agreed on the major differences between the samples, although there were individual differences between the consumers' responses related to minor sample differences. Furthermore, a comparison with trained panel conventional profiling data showed that consumers were able to extract the same major differences as the trained panel, but not the smaller differences. The results obtained with a consumer panel suggested that no training is needed to perform PM. The authors made an interesting observation regarding the consumer sample; as the consumer group did not represent a specific target market, their responses may have been more heterogeneous and therefore could have introduced less homogenous results. However, at this point conventional profiling was still needed to provide insight into the attributes that would drive differentiation on a PM product configuration as a descriptive phase has not been incorporated into the PM method yet. PM was therefor still regarded as a complementary tool.

2.5.2.2 Development of Napping®

In 2005, Pagès (2005) reintroduced PM under the name Napping®, in reference the French word "nappe", meaning table cloth. In his study, ten wines were evaluated with PM by a panel of wine professionals using a 60 x 40 cm sheet for the positioning of the samples. In addition to this, wine experts also evaluated the wines for a set of 25 descriptors on a ten point scale. The data were analysed with MFA according to the data matrix shown in Figure 2.3. The co-ordinate data from the PM were not scaled, to retain information on the panellists' use of the horizontal and vertical axes. This use of the rectangular tasting sheet and the analyses of the unscaled data with MFA, are reportedly the distinctive features of the Napping® method as opposed to PM (Dehlholm et al., 2012a). The descriptive data were added to the co-ordinate data matrix as supplementary data i.e. given very low weights so as not to contribute to the final consensus configuration. By adding the descriptive data in this way, the authors were now for the first time able to explain the Napping® configuration by correlating the descriptive data to the MFA factors. Based on the use of unscaled data, Pagès (2005) reported that the panellists tended to use the horizontal axis more prominently than the vertical axis. He furthermore noted that the subjects primarily based their PM on olfaction; one sample was particularly sweeter than the others, but this characteristic was not taken into account when the panellists positioned the wines on the tasting sheet. It was suggested that Napping® should be used as a complimentary tool to conventional descriptive profiling rather than an alternative tool, since descriptive data are necessary to interpret the axes of the napping consensus configuration (Pagès, 2005).

2.5.2.3 Incorporation of descriptive terminology

The problem of obtaining meaningful descriptive data directly from PM was soon solved by Perrin et al. (2008), who introduced the concept of combining ultra-flash profiling with Napping® to enrich the method with descriptive information. In this adaptation, panellists supply descriptors for each sample which are then captured for all panellists on a citation basis. Initially, the descriptors provided by each panellist were kept separate, but later publications reduced the potentially massive number of descriptors by adding terms used by more than one panellist (Dehlholm et al., 2012a; Kennedy, 2010; Moussaoui & Varela, 2010). Although repeatability was not tested, this definitive work was the first indication that Napping® could be used as a stand-alone alternative to conventional profiling in situations where a rough overview of the samples is sufficient. However, some studies have found that descriptors generated with PM, and by extension Napping®, can be more difficult to understand than those generated from other rapid profiling methods (Santos et al., 2013). This observation was made in the view that PM can deliver more holistic and complex descriptors in addition to singular descriptors that are used in conventional descriptive profiling and attribute-based methods (Reinbach et al., 2013).

2.5.2.4 Comparisons of trained, expert and consumer panels

The question of trained panels vs. consumer panels was further explored by Barcenas et al. (2004) in a study on ewe's milk cheeses. PM was performed in triplicate with a trained panel and an untrained consumer panel. Trained panels proved to be more repeatable and their observations were more homogeneous compared to consumers. Again it was found that the trained panellists and consumers agreed on the major sample differences but not the smaller, more subtle differences. As the authors rightly mentioned, the trained panellists were trained to recognise and describe subtle differences that might go by unnoticed for consumers.

Two studies on French wines proposed that PM was well suited to product experts, a third type of panel that have received very little attention in the validation of PM (Pagès, 2005; Perrin et al., 2008). Pagès' (2005) results also showed variability among the respective tasters. Ten to fifteen percent of the tasters struggled with the task, and like Barcenas et al., (2004) the author suggested that Napping® may not be ideally suited to untrained consumers. In another study, which also focussed on the role of expertise on PM results, it was shown that chefs, as product experts, were found to be better at positioning duplicate samples close to one another than inexperienced consumers (Nestrud & Lawless, 2008). Supporting this statement, Moussaoui and Varela (2010) reported that despite duplicate samples being positioned close together on the consensus product configuration, the consumers who participated in their study did not necessarily realise that there were duplicated samples in the product set. Product experts seemed to be more idiosyncratic in their approach compared to the consumers, an observation which the authors attributed to the experts' sensitivity towards more subtle differences between the samples (Nestrud & Lawless, 2008). It appears from these studies that, although PM was initially developed as a tool for consumer studies, it may be more suitable for expert and trained panels. A study on the sensory evaluation of chocolates suggested that training prior to conducting PM may be beneficial (Kennedy & Heymann, 2009). Another study reported a degree of learning by the panellists towards the last of three PM sessions based on a decrease in time required to perform the task and also a slight increase in internal consistency as measured by RV coefficients between successive replications (Kennedy, 2010). This increase in performance also translated to the consensus product configurations. The evidence provided by the author to support these statements was very limited. Nevertheless the concept that training can improve the effectiveness of PM panellists begs further research and the development of more appropriate tools to monitor individual panellist performance during training.

2.4.2.5 Panellists' criteria for differentiation in projective mapping

A very recent study used PM to understand the perception of wine aroma differences by wine experts and naïve consumers (Torri et al., 2013). The authors found that the projective maps from wine experts and consumers differed from results obtained by a trained panel and conventional descriptive profiling.

The reason for the difference was ascribed to the observation that, while the descriptive profiling showed product differentiation based on aromatic properties, the differentiation obtained from the experts and consumers were based on overall quality and liking respectively. For example two wines that smell completely different may be considered equal in quality (or liking) and were therefore positioned close together during PM. The authors commented that using PM as a complementary tool for descriptive analysis will only be valid if the differentiating criteria for the PM is based on the same aspects that were measured in conventional descriptive profiling. Therefore it can be concluded that PM outcomes are bound to differ from those obtained from other methods if the mapping criteria differ (Deegan et al., 2013; Nestrud & Lawless, 2008; Torri et al., 2013). A study that compared the use of various profiling methods for the evaluation of prebiotic yogurts using untrained panels provided further insights on the importance of panellists' criteria for sample differentiation in PM. It was suggested that differences observed between PM outcomes and those from a Check-All-That-Apply (CATA) method may have been due to the fact that the criteria that the participants used to conduct the PM was not considered in the list of descriptors on the CATA list (Cruz et al., 2013).

2.5.2.6 Projective mapping of complex product types

A study on the PM of two sample sets, ten apples and ten cheeses respectively, revealed that product type and complexity may influence panellists' ability to perform the task (Nestrud and Lawless, 2010). Other studies also mentioned concerns that the homogeneity of a sample set will influence the consistency of PM results (Kennedy & Heymann, 2009; Risvik et al., 1994). The first PM report specifically focussing on complex products was given by Albert et al. (2011) in their study on fish nuggets. These types of products are mostly evaluated in terms of kineasthetic attributes, e.g. tenderness, which are generally difficult to quantify. The authors reported that PM elicited attributes referring to appearance and flavour, which were omitted in the conventional profiling task due to lack of agreement between panellists and lack of reliable quantification. From this study it seems that PM can aid in gathering information on attributes that are difficult to measure in complex food products that would be lost during conventional profiling.

2.5.2.7 Repeatability and reproducibility of projective mapping

Kennedy (2010) reported that, over three sessions, repeatable overall consensus configurations could be obtained despite the low repeatability observed for the individual consumers participating in the study. Interestingly, between the overall consensus configurations there were better agreement on the first two factors than on the third factor. Kennedy hypothesised that although the panellists were consistent in their observations of the major differentiating attributes, they may have changed their criteria regarding the smaller sample differences. In contrast, Hopfer and Heymann (2013) reported insufficient repeatability and the authors strongly recommended building replicated sessions into PM studies. The contradicting results of these two studies can partly be explained in terms of the size and homogeneity of the tested sample sets. Kennedy (2010) studied eight fairly differentiating granola bars, while Hopfer and Heymann studied 20 red wines, a much more complex and homogeneous product type. In a study on liver pâtés, Dehlholm et al. (2012a) reported, substantial differences between the PM results from two panels, suggesting the PM results could not be reproduced as effectively as conventional descriptive profiling. On the other hand, Kennedy and Heymann (2009) reported good reproducibility, testing chocolate. As product type appears to play a role in repeatability, and possibly reproducibility, it seems prudent to validate the PM for a specific product type before implementation. Varela and Ares (2012) indicated that attention needs to be given to the evaluation the reproducibility of PM; judging by the contradiction in the very recent literature mentioned above, this is still the case.

2.5.2.8 Investigations of the upper-limit for sample set size for projective mapping

King et al. (1998) stretched the sample set size from the norm of 8-12 products to 18 products, in this case commercial snack bars. The authors reported that a meaningful product configuration could be obtained with a consumer panel (King et al., 1998). However, the consumers were specifically instructed to evaluate the samples based on usage occasion and liking. Only one replicate was captured and therefore, it is unclear whether the results of such a large sample set could be reproduced, especially with an untrained panel (King et al., 1998). Hopfer and Heymann (2013) attempted to define an upper-limit for the number of samples that can be evaluated with PM, and for this purpose they tested 20 wines in terms of overall sensory perception. Based on the high variability found in the judges' ability to position blind duplicate samples close to one another, the authors concluded that the sample set may have been too large. Pagès' (2005) recommendation for a maximum of 12 samples seems to be appropriate for the present.

2.5.2.9 The effect of projective mapping tasting sheet shapes on overall results

Two studies have recently investigated the effect of PM tasting sheet geometry on the consensus product configuration, comparing squares and rectangles (Hofper & Heymann, 2013) and also rectangles with circles (Dehlholm, 2013). Both studies reported that panellists adjust their approach to the PM task according to the tasting sheet shape. The lengthened axis of the rectangular tasting sheet has in both cases been said to provide a dominant direction of differentiation which panellists use to project the main product differences on. This is similar to what Pagès (2005) observed, namely that panellists tended to use the horizontal axis more prominently than the vertical axis. In contrast, a more symmetric shape, with equal perpendicular bisectors through the centre point, has been argued to allow for more subtle differentiating dimensions to be captured. Interestingly, in Hopfer and Heymann's study (2013), the judges seemed to only respond to differences on the horizontal axes and not on the vertical axes. Specifically, the judges tended to treat a rectangle with a long vertical axis like a square, since they did not use the whole length of the vertical axis to convey their perceptions of the products. It was found that the overall consensus configuration obtained respectively from square and rectangular sheets differed significantly (Hopfer & Heymann, 2013). It must be noted that the sample set tested in their study was abnormally large for projective mapping, introducing more room for variance. Dehlholm (2013), on the other hand, found that although the panellists treated the round and rectangular tasting sheets differently, the overall consensus configurations were very similar. The concept of using a specific tasting sheet shape to obtain specific results, is built around the notion that panellists' presumably react differently to different tasting sheet shapes. In Dehlholm's (2013) study, two different panels evaluated the round and rectangular shapes, making a direct comparison from a panellists point of view impossible. Although it seems that panellists are influenced by the shape of the tasting sheet, it is still uncertain whether tasting sheet design has a substantial effect on overall results.

2.5.2.10 Limitations of projective mapping compared to other descriptive methods

Dehlholm et al. (2012a) presented a very comprehensive evaluation of RSP methods, including PM. The study involved the evaluation of nine different liver pate samples by two trained panels. One PM replication was captured for each panel. The results were compared to those from conventional profiling, flash profiling, partial napping and multiple free sorting. It was observed that, out of all the methods tested, PM was the least discriminative. On the other hand, PM was found to be the fastest method among those tested. The authors maintain that differences observed between PM and conventional profiling should be viewed in the light that the two methods are based on different concepts and therefor answers different sensory analytical questions. The descriptive process during PM is focussed entirely on the sample, whereas the focus is split between sample and attribute in the case of conventional profiling and other attribute-based methods. According to the authors this allows for more complex semantics and subsequently richer descriptive information (Dehlholm et al., 2012a). Conversely, Moussaoui and Varela (2010) observed, in a study on hot beverages, that PM was in fact less rich and descriptive than flash profiling, a method that was also tested by Dehlholm et al. (2012a). Nevertheless, the two studies used different types of panels, Moussaoui and Varela (2010) compared the methods using consumer panels, while Dehlholm et al. (2012a) used trained panels. In intuitive descriptive methods such as PM, the richness of the vocabulary is likely to depend on the descriptive ability of the individuals. PM has been critised for being difficult to explain to naïve consumers (Veinand et al., 2011), while other authors commented that assessors experienced it as a game (Pagès, 2005).

2.5.2.11 Variations of projective mapping

Although PM is by definition an intuitive task, the idea of a structured version with labelled axes was also entertained (King et al., 1998). However, structured PM was not as effective in revealing meaningful inter-sample relationships as unstructured PM. The concept was brought up again during a presentation at the Eurosense conference in 2010 (Ferrage, 2010), but has not been explored further.

The idea of a modality-based PM procedure, where panellists are required to perform PM only on one sensory modality (e.g. olfaction) was suggested as a means to bypass any loss of information caused by panellists being forced to transfer multi-dimensional sensory perceptions onto a two-dimensional space (Pagès, 2005). This concept was later presented under the name of partial napping (Pfeiffer & Gilbert, 2008) as a middle ground between the intuitive napping approach and the analytical conventional profiling approach. In partial napping, the sensory perception of a product set may be divided into several sensory modalities. This is illustrated in Figure 2.4, which shows five brandies being evaluated for appearance (b), olfaction (d) and in-mouth sensations (c) as opposed to a holistic PM approach (a). The data from the different modalities are treated with hierarchical multiple factor analysis or HMFA (Le Dien & Pagès, 2003) in order to obtain a consensus configuration incorporating information from all the modalities. HMFA follows the same principle as MFA, a two-step data analyses approach where variables from several data matrices are weighted by the eigenvalues obtained from a preliminary PCA on each data matrix. These weighted variables are then subjected to a second PCA to obtain a consensus configuration based on all the data matrices. In addition to this, HMFA takes the hierarchical structure of the data into account. The rationale behind the method is to scale each level of the hierarchy, starting with the lowest level in the hierarchy i.e. the individual judges for each modality.

To differentiate between partial napping and the original Napping® method, the latter is now often referred to as global napping (Dehlholm et al., 2012a; Pfeiffer & Gilbert, 2008). Comparisons between conventional profiling, partial and global napping showed that, of the two napping configurations, the partial napping product configuration was the most similar to conventional descriptive profiling product configuration (Dehlholm et al., 2012a; Pfeiffer & Gilbert, 2008). Partial napping was also reported to provide more detailed descriptors, making it easier to interpret the results (Pfeiffer & Gilbert, 2008). Partial napping was further shown to be very reproducible and more discriminating than global napping (Dehlholm et al., 2012a). It was advocated as a suitable method to obtain a holistic representation, but with focus directed to modalities of specific importance to the researcher.

Initial studies used partial napping to focus more analytically, at the time of sample evaluation on separate modalities and afterwards, during the data capturing stage, building up a consensus product map based on several modalities (Dehlholm et al., 2012a; Pfeiffer & Gilbert, 2008). However, partial napping can also be applied in such a way as to force panellists to focus on a specific modality, disregarding the other modalities altogether. In an example on beer, data were captured for smell and taste, while the modalities of appearance and mouthfeel were intentionally ignored (Giacalone et al., 2013). Although the results suggested that the influence of appearance could not be completely excluded without the use of dark glasses, the partial napping approach appeared to be effective in avoiding the influence of mouthfeel. A similar restrictive approach was followed for evaluating the texture of soy- and cow's milk gels. The panellists were issued with nose clips to reduce the influence of flavour (Grygorczyk et al., 2013).

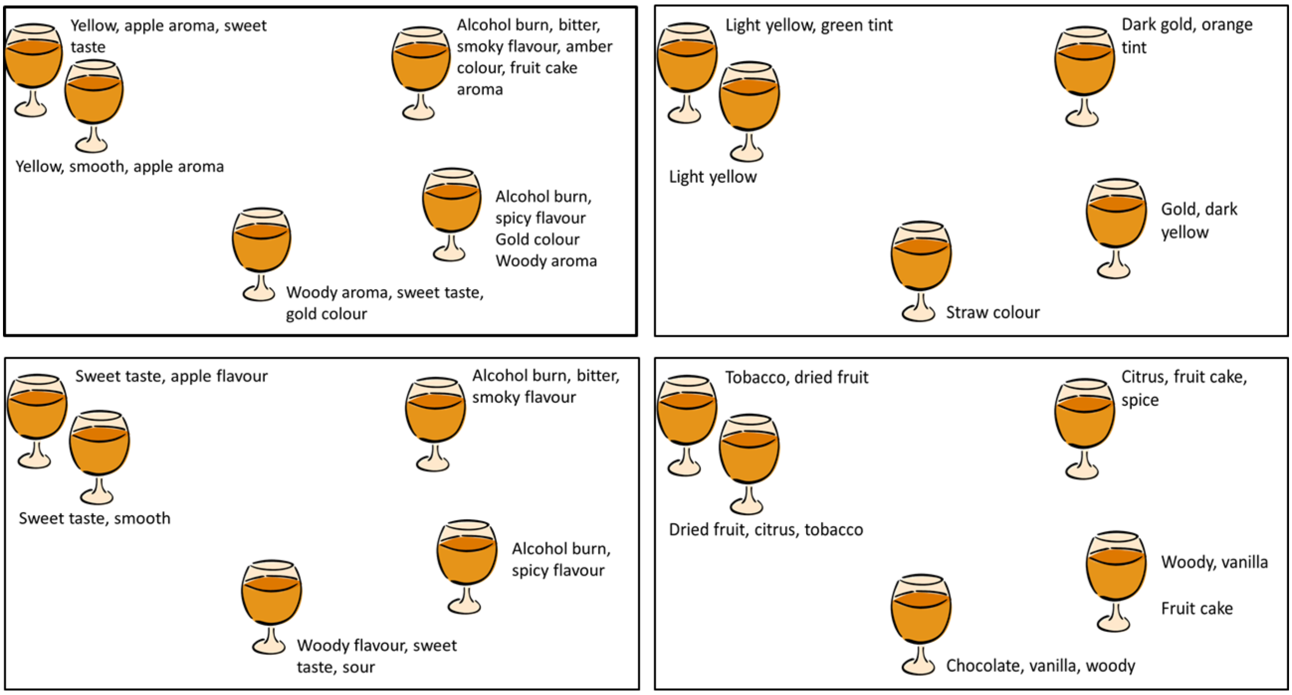

Figure 2.4 Four projective mapping sheets showing the difference between global napping (a) and partial napping based on appearance (b), in-mouth sensations (c) and olfaction (d) (own data).

Sorted napping, a combination of Napping® and sorting, is another variation on the PM method. In this case, panellists are instructed to indicate groups of similar samples after positioning the samples (Pagès, Cadoret & Lê, 2010). The added benefits of combining the two procedures are not well documented.

A very recent variation on PM is polarised projective mapping (Ares et al., 2013). This variation combines PM with a relatively new method called polarised sensory positioning (PSP) in order to address the respective shortcomings of each method, namely that product maps from PM can not be compared over time and that in PSP, samples are evaluated only relative to pre-selected reference samples. PSP involves evaluating samples against a fixed set of reference products or poles (Teillet et al., 2010). For polarised PM, three poles or reference samples need to be selected, based on product knowledge as well as positions on the first two dimensions of a preliminary PM consensus configuration of relevant products. The three poles are then pre-positioned on a PM tasting sheet according to their position on the first PM configuration. The tasters are asked to position their samples relative to the positions of the poles and to describe the differences. The data were analysed with MFA in the same fashion as normal PM. The preliminary report on this method showed promise, especially as a product development tool. A potential advantage over PM is that there is no limitations on sample set size, since the data from many sessions testing different products against the reference products can be potentially be used to build a global map. However, this claim must still be verified. A second advantage is that samples were evaluated at different times can be compared directly. The main disadvantage is that, like PSP, thorough knowledge regarding the product category is needed to identify suitable poles. However, as the authors indicated, rapid sensory profiling methods can be used for this purpose; therefor the required knowledge can be obtained in a short time-span. Further research is needed to fully understand the applications and practical considerations of polarised PM.

2.5.2.12 Applications of projective mapping

Apart from method development and validation studies discussed before, PM has been applied as a tool in typicality judgement of wine where wines were rated by experts according how typical the wines were of their region and also evaluated with PM. It was found that the first two factors of the PM sample configuration corresponded to both appellation and typicality (Perrin & Pagès, 2009). Another wine study used PM to test the effect of serving temperature on the sensory perception of red wines (Ross et al., 2012). In this study it was found that red wines served at 10°C and 16°C were perceived as less flavourful and more sour and astringent than wines served at 22°C. Furthermore, the panel was more repeatable in their assessments when the wines were served at 22°C than when served at cooler temperatures. In addition, PM was made part of a Qualitative Multivariate Analysis technique (QMA) which the authors proposed as a tool to create a holistic understanding of the consumer product experience (Drake et al., 2009). Two studies have used PM to understand the drivers behind consumer acceptability of nutritionally enriched biscuits and green tea respectively (Carrillo et al*.,* 2012a; Kim et al., 2013). The latter study reported a particularly interesting finding on cross-cultural differences in green tea perceptions that may not have been elicited in a traditional preference mappings study. The Korean consumers' internal preference map (based on acceptability) differed more from their projective map (based on overall perception) than what was observed for the French consumers. This highlighted that French consumers positioned the products according to how much they liked the products whereas, Korean consumers focussed more on objective sensory differences. The authors hypothesised that sample familiarity affects consumers' perception of green tea (Kim et al., 2013). Another interesting application study showed PM being used as a screening tool to narrow down 23 biscuit samples based on packaging to 10 samples for intrinsic sensory evaluation. As the authors rightly mentioned, 23 brands would have been overwhelming to evaluate and position in a PM exercise; for this reason the biscuit brands were first evaluated based on packaging after which 10 differentiating samples were chosen for further analysis (Carrillo et al*.,* 2012b).

2.5.3 Development and optimisation of the sorting task

Sorting is a very intuitive method during which panelists are required to sort samples according to overall sensory similarity. Product configurations are obtained by analysing sample co-occurrences in the groups made by the panellists. As with the previous section, this section will start with an overview of the development of the sorting method. Key findings regarding practical considerations for application will be discussed. Variations on the sorting task will be presented after which studies where sorting has been successfully applied will be summarised.

2.5.3.1 First feasibility assessment of the sorting task for application to food products

The feasibility of the sorting task as an explorative tool to study sensory variation of food was confirmed by Lawless et al. in 1995. The study comprised sorting of two cheese sample sets, the first consisting of 16 different types of cheeses, while the second consisted of a more homogeneous sample set of 13 blue cheeses (Lawless et al., 1995). The data were analysed with MDS to generate a sample plot. In addition, several descriptive attributes were rated on line scales, the mean values of which were regressed onto the sample configuration plot. Duplicated pairs of cheeses, and Youdan pairs (same brand with batch variations) positioned close together. This confirmed the reliability of the method, an aspect which has been supported by results from other studies (Falahee et al., 1997).

2.5.3.2 Comparison of trained vs. untrained panellists

Consistent consensus results have been reported for both trained and untrained panels, with good agreement between assessors for the trained panel (Cartier et al., 2006; Chollet et al., 2011). However, it has been suggested that visual heterogeneity in a sample set can facilitate the panel's consistency (Cartier et al., 2006). It has been reported that although similar overall results can be obtained from trained and untrained panellists, less experienced tasters used fewer and simpler attributes during sorting than trained panellists (Lawless et al., 1995). A study on breakfast cereals confirmed that similar product maps can be obtained with trained and untrained panels, but that the meaning of more complex words, such as *crispy*, differed between the two panels. The two panels also made different associations between sweetness and olfactory perceptions (Cartier et al., 2006). A similar observation was made in relation to two expert panels from different cultures; the panels generated very similar product configuration, but complex characters such as *round* or *structured* were interpreted differently (Bécue-Bertaut & Lê, 2011).

2.5.3.3 Panellists' criteria for sorting

During the sorting task, panellists are normally encouraged to sort the samples based on any criteria they want to use. Sorting task studies have shown many different patterns and approaches to sorting criteria. In some studies, overall results suggested that untrained assessors sorted samples according to liking (Falahee et al., 1997), while others reported the consistent use of sensory perceptual differences (Cartier et al., 2006; Chollet & Valentin, 2001). Both trained and untrained assessors have been reported to change their sorting strategy and/or criteria over two sorting task sessions, with a 20 minute delay between sessions (Chollet & Valentin, 2001). Trained assessors completely changed their sorting strategy from a purely sensory approach to a technical approach where six the commercial, unflavoured, beers and six flavoured beers were sorted by type (unflavoured vs. flavoured) rather than by sensory perceptions. The untrained assessors' assessments also differed from one session to the other, but seemed to have remained focus on the sensory perceptual differences rather than categorical differences (flavoured vs. unflavoured) (Chollet & Valentin, 2001). In a study on the effect of information present on packaging material on consumer perception of milk desserts, the consumers' sorting results were to some extent influenced by information provided on the products' packaging information, even though the attributes cited during the sorting task related only to sensory characteristics. The authors suggested that since the sorting task was not guided, the panellists were free to think about sensory and non-sensory aspects, and that the latter may have influenced their sorting decisions, even if they did not mention it in their group descriptions (Vidal et al., 2013). Similarly, wine experts have been observed to use hedonic terminology in sorting tasks, even though differentiation is based on perceived differences and not liking (Bécue-Bertaut & Lê, 2011). Veramendi et al. (2013) observed that although similar configurations could be obtained from trained panellists and untrained consumers, the trained panellists' groups and descriptions did not completely reflect the consumers' perceptions. For this reason it was suggested that collecting consumer perceptions is an important complement to trained panel data as this could provide guidance for marketing and consumer communication.

2.5.3.4 Panellists' consistency in the sorting task

Falahee et al. (1997) found that only 20% of the untrained assessors provided fairly consistent results, but that despite this, a reliable overall configuration, as assessed by the proximity between duplicated products as well as similar types of products, could be obtained. Other studies have also reported poor performance among individuals. In a beer study, duplicated samples were often sorted into different groups (Chollet et al., 2011). The authors attributed this to carry-over effect. Moreover, a study on reconstituted wine also reported that only around 30% of the panellists sorted duplicated samples into the same group (Sáenz-Navajas et al., 2012). In this case it was hypothesised that the difficulty that panellists experienced in sorting duplicate samples into the same group was related to the amount of sensory information that panellists had to deal with in the sorting task.

2.5.3.5 Repeatability of the sorting task

Despite poor internal consistency on the part of individual panellists, repeatable overall consensus configurations have been reported for the sorting method (Cartier et al., 2006; Chollet et al., 2011). Sample set size and homogeneity have been implicated to affect repeatability (Chollet et al., 2011). These factors should be taken into account when using the sorting task to evaluate products.

2.5.3.6 Investigations of the upper limit for sample set size for sorting

Chollet et al. (2011) recommended a sample size of between 9 and 20 beers; the upper limit is set to avoid short-term memory overload and sensory fatigue which impacts negatively on results, while their study showed that evaluating too few samples resulted in a less stable sorting map. Another study suggested that up to 25 samples can be evaluated if the evaluation is based on olfaction only, but that less samples could be accommodated if the samples were to be tasted also (Pombioni et al., 2004).

2.5.3.7 Comparison of sorting with conventional profiling in terms of descriptive output

It has been reported that the unstructured descriptive information obtained with sorting can be less interpretable than the analytical descriptive data obtained with conventional descriptive profiling (Chollet et al., 2011). Providing the panellists with a list of descriptive terms also did not improve the interpretability of the results. The authors speculated that a shorter list of more relevant terms may have been more helpful than the list of 44 descriptors which included technical terms, which was provided in the study. Untrained panellists have been found to rely more on a list of descriptors if provided, but that it may also guide them to use words that are not necessarily relevant, due to lack of understanding of the meaning of the words (Lelièvre et al., 2008). A study on the inter-relationships between olfaction and texture perception in yogurts highlighted that the integrated perceptual approach of sorting does not reveal the same information as conventional profiling where the product is dissected into measurable attributes (Saint-Eve et al., 2004). In a sense, if the whole is not the sum of its parts, looking at the "whole" through sorting and looking at the "parts" through conventional profiling would understandably not deliver the same results.

2.5.3.8 Variations on the sorting task

Lelièvre et al. (2008) added an intensity rating element to the descriptive phase of the sorting task by asking panellists to indicate the intensity of each group descriptor as either "not" (e.g. not sweet), "a little", "medium" and "very". These terms were converted to intensity scores during the data analysis phase, which were further used to identify which descriptors were truly meaningful. It was found that trained and untrained assessors used the intensity indicators differently. This approach is likely to be more useful with panelists that are trained to consider intensity, as it could possibly deliver arbitrary results from untrained panelists.

Most of the studies discussed above involved sorting exercises where panellists were free to sort the samples according to any criteria. It is also possible to instruct panellists to sort samples according to a pre-defined criterion. In a study on snack bars, panellists were asked to sort the samples according to usage occasion. Although the final product configuration successfully depicted differences in usage occasion, no underlying dimensions of differentiation could be identified, making the benefit this approach very limited (King et al., 1998). In a study on cross-modal sensory interactions it was observed that panellists were able to sort yogurts based on texture according to their instructions, without taking other sensory perceptions into account (Saint-Eve et al., 2004). The authors proposed that lower inter-panellist variability would be observed if the sorting task is focused on a single sensory modality, although the statement was not supported with empirical evidence. This form of restrictive sorting was also used by Deegan et al. (2010) in a study on growth conditions and storage times on vegetables. They conducted a separate sorting exercise focusing only on appearance, since this is relevant to the criteria on which consumers base their purchase decisions for vegetables. The appearance sorting was followed up by a sorting based on all other perceptions in the absence of colour to provide further information on the effect of growth conditions and storage time (Deegan et al., 2010). Johnson et al. (2013) reported the first retronasal odour sorting of wines. The authors took an interesting approach similar to partial napping, where the assessors were asked to sort the samples by orthonasal aroma and to do a separate sorting on retronasal odours. By performing an initial INDSCAL test, they were able to determine whether the results from the orthonasal and retronasal sorting exercises were similar enough to be aggregated into a single MDS solution. It is arguable whether it would make sense to aggregate orthonasal and retronasal sorting maps in this sense; if the configurations are very similar it would be sufficient to use only one of these modalities, since orthonasal and retronasal odour perceptions are often highly correlated (Aubry et al., 1999; Peña y Lilo et al., 1995).

Santosa et al. (2010) proposed a two-stage sorting strategy aimed at extracting more differentiating criteria from consumers. With this method, consumers are asked to sort samples in the usual way, and afterwards they are prompted to sort their groups into sub-groups. The outcomes of the specific study showed that finer product discrimination could be obtained using a two-stage approach. However, no information was available on reproducibility, and it is uncertain at this stage, whether adding a second step would consistently yield useful results. The exact opposite of this approach was proposed, under the name of taxonomic free sorting. The method was developed as an attempt to overcome information lost by the underlying assumption that the groups made by panellists are equally different from each other (Courcoux et al., 2012). Taxonomic free sorting requires panellists to complete the sorting task as usual, after which they have to select two groups that are most similar and for a new group consisting of these two (thereby reducing the number of groups by one). This process is reiterated until only two groups remain, providing a hierarchical structure of groups. This pilot study also showed increased discrimination between products compared to the normal sorting task. The authors proposed that this approach where the hierarchical tree is “built up” instead of “broken down”, as in the case of the study by Santosa et al. (2010), is by comparison intuitively easier for the panellist, less fatiguing and less time-consuming. In both cases, the initial sorting step was taken as the control against which to compare the results from the modified task. The robustness of both approaches should be challenged by comparing it against a free sorting task conducted independently. The authors suggested that the data analyses approach to taxonomic free sorting data must be refined, which will further improve the quality of the results (Courcoux et al., 2012).

In order to uncover the sensory dimensions relating to yogurt freshness, Bouteille et al. (2013) asked panelists, after completing a sorting task, to indicate whether the groups they made were perceived as “fresh” or not. Panellists were allowed to indicate that some samples in the same group were considered fresh, while others were not. Such responses were regarded as missing values. The same was done for liking. This approach resulted in an easily interpretable map showing associations between various yogurt descriptors and perceived freshness.

2.5.3.9 Applications of the sorting task

Sorting has been applied as a tool to understand wine experts' construct of wine typicality (Ballester et al., 2005; Ballester et al., 2008). Furthermore it was used to investigate the conceptual changes regarding wine perceptions that occur with the acquisition of wine expertise, providing supporting evidence that experts differentiate wines based on conceptual differences (e.g. grape variety) as opposed to strictly perceptual differences (Ballester et al., 2008; Solomon, 1997). Sorting was also used in a study investigating the merit of taking consumers' knowledge and expertise into account during consumer studies (Faye et al., 2012). The method was also used as a sample screening tool in a study investigating cross-modal sensory interactions of cheese (Niimi, 2013). Sorting has been applied as a tool to understand relationships between terms used to describe the astringency of red wines (Gawel et al., 2001), and similarly to understand the concept of minerality in Chardonnay wines (Ballester et al., 2013). Useful and interpretable results regarding the effect of growth conditions and storage times of cucumbers and tomatoes were gathered using the sorting method (Deegan et al*.,* 2010).

2.5.4 Critical evaluation of the features of rapid sensory profiling studies

Studies on the application of RSP methodologies are usually focused on comparisons of biplots obtained from the method in question to those from conventional methods, or other RSP methods. These comparisons were often based on interpretation from sample knowledge e.g. appellation, ageing, fruit-content (Barcenas et al., 2004; King et al., 1998; Pagès, 2005). An RV coefficient of 0.700 has been cited as a cut-off point for sufficient similarity between biplots (Cartier et al., 2006; Hopfer & Heymann, 2013; Nestrud & Lawless, 2008), although RV coefficients as low as 0.400 has been found acceptable through statistical significance testing (Chollet et al., 2011; Dehlholm, 2013). More often, rough visual assessments of biplots lead to overoptimistic statements regarding comparable results. The question is not whether two biplots are statistically comparable, but whether the same conclusions could be drawn from them.

Another aspect that lacks from publications on rapid profiling methodologies is feedback from the panellists themselves. One exception is a PM study by Pagès in 2005 where qualitative feedback from the participants was captured. Nestrud and Lawless (2008) commented on insights gained from a post-hoc assessment of the panellists' individual projective mappings. A study by King et al. (1998) was the only study that explicitly asked panellists about ease of use. Some panellists may find the intuitive nature of PM more natural than others; knowledge of panellists' comfortableness with the task may be beneficial in obtaining reliable results.

Despite all the reports on variation in panellist criteria in PM, there are seldom any recommendations on how this phenomenon should be managed. Torri et al. (2013) suggested that panellists should share a common background in order to generate more interpretable maps. As mentioned earlier, PM has been reported to be a suitable method for trained panellists, and reports of a learning effect substantiate the idea that panellists can be trained to be more effective in PM. Yet, there is a lack in appropriate performance monitoring tools to be used for PM studies. Some researchers have used the panellists' physical projective maps to determine their task competency, i.e. whether samples were placed in straight lines, or scattered across the entire sheet (Nestrud & Lawless, 2008; Pagès, 2005), but this approach is labour intensive and does not provide a measure with which a panellist's competency can be quantified. It has been suggested that panellists' performance can be evaluated by their ability to position two duplicated samples close to each other on the projective mapping sheet (Bertucciolli, 2011). This is expressed as a ratio of the Euclidean distance between the two duplicate samples and the maximum inter-sample Euclidean distance in the sample set and has been referred to as the People

Performance Index (PPI) (Hopfer & Heymann, 2013) and also as a $D_{r\%}$ ratio (Torri et al., 2013). The drawback of this ratio is that it provides information on the panellists' consistency in positioning only one product i.e. the duplicated sample. Procrustes Analysis of Variance (PANOVA) has been used to determine overall panel consistency by evaluating total consensus variance. Product residuals from PANOVA has been used to determine whether there were any specific products that the panellists disagreed on (Nestrud & Lawless, 2008). Although this approach provides information the panel's consistency for all samples, it does not provide a single interpretable measure.

One of the more confusing aspects of RSP is the distinction between PM and Napping®. PM was first proposed as a suitable tool for the perceptual profiling of food products by Risvik et al. in 1994. Pagès (2003 and 2005) reintroduced PM under the name Napping®. The two terms are often used interchangeably. Dehlholm et al. (2012a) pointed out in their noteworthy comparison of the most prevalent RSP methods that Napping® is in fact a sub-form of PM. Napping® specifically requires a rectangular tasting sheet, or "nappe". Napping® data are specifically analysed with MFA, using unscaled data to preserve idiosyncrasies in the panellists' use of the X- and Y axes. The original report on PM used GPA to analyse the data and other methods have also since been used. The issue of data analysis method as a distinctive attribute of Napping® vs. PM is perhaps not so pertinent as little difference have been observed between PM data analysed with MFA, GPA, or even PMFA, which also includes a procrustes transformation (Kennedy, 2010; Nestrud & Lawless, 2008). Furthermore, scaling has been found to have little effect on the overall consensus configuration (Dehlholm, 2013). The specification in terms of tasting sheet shape may be of more relevance in view of the recent studies suggesting that panellists do in fact treat different sheet shapes differently, although it does not necessarily impact on conclusions drawn from the overall consensus configurations (Dehlholm, 2013; Hopfer & Heymann, 2013). The relevance of the distinction between Napping® and PM is debatable, but arbitrary use of the terms should for the moment be viewed with caution until the actual significance of tasting sheet shape has been resolved.

Another confusing aspect of Napping® is the spelling; across all the papers dealing with Napping®, there seems to be no consensus on whether the term should be capitalised and the use of the registration symbol also seems to be optional. It appears as if there is a trend to omit the registration symbol when the term is used with a qualifier (e.g. partial napping). Perhaps it is not at all surprising that some authors prefer to use the more general term, PM. Sorting is another method with an inconsistent naming convention, as it is also referred to as free sorting (Faye et al., 2012) and categorisation (Lelièvre et al., 2009).

The nomenclature for method and panel performance has also been inconsistent. The terms repeatability, reproducibility, "replicability" and stability are used interchangeably, sometimes in the same study e.g. that of Barcenas et al., (2004). It is necessary to establish a clear definition for each, as they could refer to the positioning of duplicate samples, the ability of a panel or panellists to generate comparable results in different sessions under the same conditions or, the ability to produce similar results with two different panels, perhaps even with a slight change in the sample set. In this study, the term repeatability will refer to the ability to obtain the similar results with the same panel whereas reproducibility will refer to the ability to obtain similar results using two different panels. The positioning of duplicated samples will be referred to as accuracy.

2.6 Concluding remarks

Although the past two decades of research have clarified several questions regarding the validity of PM and sorting as sensory methods, there are still some uncertainties pertaining to the practical application thereof. Based on the contradictory evidence regarding repeatability and reproducibility of PM, more evidence is needed to establish the robustness of this method. The literature shows that individual differences are observed among panellists and inconsistency on panellist level is often reported. Panellist variability is commonly attributed to changes in projective mapping criteria or strategy. There is a lack of information on what experimental factors influence their performance, including the matter of the projective mapping sheet shape. More importantly, the concept of managing panelist performance to obtain reliable and interpretable PM results is still largely unexplored. There is a need for an interpretable measure that can be used as a panellist performance monitoring tool. Although RSP methods could be a potentially powerful addition to the brandy sensory evaluation toolbox, it is unclear whether similarity-based RSP methods are able to accommodate the unique challenges associated with complex high alcohol beverages such as brandy.

2.7 References

Abdi, H. (2007). Rv and congruence coefficients. IN: *Encyclopedia of Measurement & Statistics*. (Ed. N.J. Salkind). Thousand Oaks, USA: SAGE publications Inc., 849-853.

Abdi, H., Valentin, D., Chollet, S., & Chrea, C. (2007). Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. *Food Quality and Preference, 8*, 627-640.

Abdi, H., Dunlop, J.P., & Williams, L.J. (2009). How to compute reliability estimates and display confidence and tolerance intervals for pattern classifiers using the bootstrap and 3-way multidimensional scaling (DISTATIS). *Neuro Image, 45,* 89-95.

Adams, J., Williams, S., Lancaster, B., & Foley, M. (2007). Advantages and uses of check-all-that-apply response compared to traditional scaling of attributes for salty snacks. *7th Pangborn sensory science symposium. Minneapolis, MN., USA*, *2007.*

Albert, A., Valera, P., Salvador, A., Hough, G., & Fiszman, S. (2011). Overcoming the issues in the sensory description of hot served food with a complex texture. Appliation of QDA®, flash profiling and projective mapping using panels with different degrees of training. *Food Quality and Preference, 22,* 463-473.

Ares, G., De Saldamando, L., Vidal, L., Antúnez, L., Giménez, A., & Varela, P. (2013). Polarized projective mapping: Comparison with Polarized Sensory Positioning approaches. *Food Quality and Preference, 28,* 510-518.

Aubry, V., Etievant, P., Sauvageot, F., & Issanchou, S. (1999). Sensory analysis of Burgundy Pinot noir wines: A comparison of orthonasal and retronasal profiling. *Journal of Sensory Studies*, *14,* 97-117.

Ballester, J., Dacremont, C., Le Fur, Y., & Etievant, P. (2005). The role of olfaction in the elaboration and use of the Chardonnay wine concept. *Food Quality and Preference*, *16,* 351-359.

Ballester, J., Patris, B., Symoneaux, R., & Valentin, D. (2008). Conceptual vs. perceptual wine spaces: Does expertise matter? *Food Quality and Preference, 19,* 267-276.

Ballester, J., Mihnea, M., Peyron, D., & Valentin, D. (2013). Exploring minerality of Burgundy Chardonnay wines: a sensory approach with wine experts and trained panellists. *Australian Journal of Grape and Wine Research, 19,* 140-152.

Barcenas, P., Pérez Elortondo, F., Albisu, M. (2004). PM in sensory analysis of ewes milk cheeses: a study on consumer and trained panel performance. *Food Research International 37,* 723-729.

Bécue-Bertaut, M., & Lê, S. (2011). Analysis of multilingual labeled sorting tasks: application to a cross-cultural study in wine industry. *Journal of Sensory Studies, 26*, 299-310.

Bertuccioli, M. (2011). *The role of sensory analysis to identify factors contributing to wine performance.* Graz, Austria: Keynote lecture at the In Vino Analytica Scientia.

Blancher, G., Clavier, B., Egoroff, C., Duineveld, K., & Parcon, J. (2012). A method to investigate the stability of a sorting map. *Food Quality and Preference*, *23,* 36-43.

Bouteille, R., Cordelle, S., Laval, C., Tournier, C., Lecanu, B., This, H., & Schich, P. (2013). Sensory exploration of the freshness sensation in plain yoghurts and yoghurt-like products. *Food Quality and Preference*, *30,* 282-292.

Cadoret, M., & Husson, F. (2013). Construction and evaluation of confidence ellipses applied at sensory data. *Food Quality and Preference*, *28,* 106-115.

Cadoret, M., Lê, A., & Pages, J. (2009). A factorial approach for sorting task data (FAST). *Food Quality and Preference*, *20,* 410-417.

Caldeira, I., Belchior, A.P., Clímaco, M.C., & Bruno de Sousa, R. (2002), Aroma profile of Portuguese brandies aged in chestnut and oak woods. *Analytica Chimica Acta 458,* 55-62.

Caldeira, I., Mateus, A. M., & Belchior, A.P. (2006) Flavour and odour profile modifications during the first five years of Lourinha brandy maturation on different wooden barrels. *Analytica Chimica Acta*, *563,* 264-273.

Caldeira, I., Anjos, O., Portal, V., Belchior, A., & Canas, S. (2010). Sensory and chemical modifications of wine-brandy aged with chestnut and oak wood fragments in comparison to wooden barrels. *Analytica Chimica Acta*, *660,* 43-52.

Canas, S., Caldeira, I., & Belchior, P. (2009). Comparison of alternative systems for the ageing of wine brandy, wood shape and wood botanical species effect. *Ciência E Técnica Vitivinicola, 24 (2),* 91-99.

Carrillo, E., Varela, P., & Fiszman, S. (2012 a). Effects of food package information and sensory characteristics on the perception of healthiness and the acceptability of enriched biscuits. *Food Research International, 48,* 209-216.

Carrillo, E., Varela, P., & Fiszman, S. (2012 b). Packaging information as a modulator of consumers' perceptions of enriched and reduced-calorie biscuits in tasting and non-tasting tests. *Food Quality and Preference 25,* 105-115.

Cartier, R., Rytz, A., Lecomte, A., Poblete, F., Krystlik, J., Belin, E., & Martin, N. (2006). Sorting procedure as an alternative to quantitative descriptive analysis to obtain a product sensory map. *Food Quality and Preference, 17,* 662-571.

Chollet, S., & Valentin, D. (2001). Impact of training on beer flavour perception and description: are trained and untrained subjects really different? *Journal of Sensory Studies, 16,* 601-618.

Chollet, S., Lelièvre, M., Abdi, H. & Valentin, D. (2011). Sort and beer: Everything you wanted to know about the sorting task but did not dare to ask. *Food Quality and Preference, 22,* 507-520.

Courcoux, P., Qannari, E.M., Taylor, Y., Buck, D., & Greenhoff, K. (2012). Taxonomic free sorting. *Food Quality and Preference, 23,* 30-35.

Courcoux, P., Faye, P., & Qannari, E.M. (2013). Determination of the consensus partition and cluster analysis of subjects in a free sorting task experiment. *Food Quality and Preference*, *32 (Part A),* 107-112.

Cruz, A.G., Cadena, R.S., Castro, W.F., Esmerino, E.A., Rodrigues, J.B., Gaze, L., Faria, J.A.F., Freitas, M.Q., Deliza, R., & Bolini, H.M.A. (2013). Consumer perception of probiotic yogurt: performance of Check all that apply (CATA), PM, sorting and intensity scale. *Food Research International* (2013), doi:10.1016/j.foodres.2013.07.056.

Dairou, V., & Sieffermann, J.M. (2002). A comparison of 14 jams characterized by conventional profile and a quick original method, the flash profile. *Journal of Food Science 67, (2),* 826-834.

Dehlholm, C., Brockhoff, P.B., Meinert, L., Aaslung, M.D., & Bredie, W.L.P. (2012)a. Rapid descriptive sensory methods – comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference 26,* 267-277.

Dehlholm, C., Brockhoff, P.B., & Bredie, W.L.P. (2012)b. Confidence ellipses: A variation based on parametric bootstrapping applicable on Multiple Factor Analysis results for rapid graphical evaluation. *Food Quality and Preference*, *26,* 278-280.

Dehlholm, C. (2013) *Descriptive sensory evaluations: Comparison and applicability of novel rapid methodologies. PhD Thesis,* ISBN 978-87-7611-592-0.

Deegan, K.C., Koivisto, L., Näkkilä, J., Hyvönen, L., & Tuorila, H. (2010). Application of a sorting procedure to greenhouse-grown cucumbers and tomatoes. *LWT Food Science and Technology*, *43,* 393-400.

Deegan, K.C., Holopainen, U., McSweeney, P.L.H., Alatossava,T. & Tuorila, H. (2013). Characterisation of the sensory properties and market positioning of novel reduced-fat cheese, *Innovative Food Science and Emerging Technologies* (2013), doi: 10.1016/j.ifset.2013.10.003.

Drake, S.L., Lopetcharat, K., & Drake, M.A. (2009). Comparison of two methods to explore consumer preferences for cottage cheese. *Journal of Diary Science, 92*, 5883-5897.

Du Plessis, H.W., Steger, C.L.C., Du Toit, M., & Lambrechts, M.G. (2002). The occurrence of malolactic fermentation in brandy base wine and its influence on brandy quality, *Journal of Applied Microbiology, 92,* 1005-1013.

Falahee, M., & MacRae, A.W. (1997). Perceptual variation among drinking waters: The reliability of sorting and ranking data for multidimensional scaling. *Food Quality and Preference*, 8, 389.

Faye, P., Courcoux, P., Giboreau, A., & Qannari, E.M. (2012). Assessing and taking into account the subjects' experience and knowledge in consumer studies. Application to the free sorting of wine glasses. *Food Quality and Preference*, *28,* 317-327.

Ferrage, A., Nicod, L., & Varela, P. (2010). How does assessors' performance influence the outcome of alternative sensory methodologies? *Eurosense 2010 – A sense of quality*. Vitoria-Gasteiz, Spain, 5-8 September, 2010.

Gawel, R., Iland, P.G., & Francis, I.L. (2001). Characterizing the astringency of red wine: a case study. *Food Quality and Preference, 12,* 83-94.

Giacalone, D., Ribeiro, L.M., & Frøst, M.B. (2013). Consumer-based product profiling: Application of Partial Napping® for sensory characterisation of speciality beers by novices and experts. *Journal of Food Products Marketing, 19:3,* 201-218.

Grygorczyk, A., Lesschaeve, I., Corredig, M., & Duizer, L. (2013). Impact of structure modification on texture of a soymilk and cow's milk gel assessed using the napping procedure. *Journal of Texture Studies, 44,* 236-246.

Hopfer, H., & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. *Food Quality and Preference, 28,* 164-181.

Husson, F., & Josse, J. (2013). Handling missing values in multiple factor analysis. *Food Quality and Preference, 30.* 77-85.

Jack, J. (2003). Development of guidelines for the preparation and handling of sensory samples in the Scotch whisky industry. *Journal of the Institute of Brewing, 109,* 114-119.

Janáčová, A., Sadecka, J., Kohajdova, Z., & Spanik, I. (2008). The identification of aroma-active compounds in Slovak brandies using GC-sniffing, GC-MS and sensory evaluation. *Chromatographia, 67,* S113-S121.

Johnson, T.E., Hasted, A., Ristic, R.,& Bastian, S.E.P. (2013). Multidimensional scaling (MDS), cluster and descriptive analyses provide preliminary insights into Australian Shiraz wine regional characteristics. *Food Quality and Preference*, *29,* 174-185.

Jolly, N. P., & Hattingh, S (2001). A brandy aroma wheel or South African brandy. *South African. Journal of. Enology and Viticulture. 22,* 16-21.

Kennedy, J. (2010). Evaluation of replicated projective mapping of granola bars. *Journal of Sensory Studies, 25,* 672-684.

Kennedy, J., & Heymann, H. (2009). Projective mapping and descriptive analysis of milk and dark chocolates. *Journal of Sensory Studies, 24,* 220-233.

Kim, Y., Jombart, L., Valentin, D., & Kim, K. (2013). A cross-cultural study using Napping®: Do Korean and French consumers perceive various green tea products differently? *Food Research International*, *53,* 534-542.

King, M.C.,Cliff, M.A., & Hall, J.W. (1998). Comparision of PM and sorting data collection and multivariate methodologies for identification of similarity-of-use of snack bars. *Journal of Sensory Studies, 13,* 347-358.

Lawless, H.T., Sheng, N., & Knoops, S.S.C.P. (1995). Multidimensional scaling of sorting data applied to cheese perception. *Food quality and preference*, *6,* 91-98.

Le Dien, S., & Pagès, J. (2003). Hierarchical multiple factor analysis: application to the comparison of sensory profiles*.* *Food Quality and Preference, 14,* 397-403.

Lelièvre, M., Chollet, S., Abdi, H., & Valentin, D. (2008). What is the validity of the sorting task for describing beers? A study using trained and untrained assessors. *Food Quality and Preference*, *19,* 697-703.

Lelièvre, M., Chollet, S., Abdi, H., & Valentin, D. (2009). Beer-trained and untrained assessors rely more on vision than on taste when they categorize beers. *Chemosensory Perception*, *2,* 143-153.

Meilgaard, M.C., Civille, G.V., & Carr, B.T. (2007). Controls for test room, product and panel IN : *Sensory evaluation techniques.* London, CRC press.

Miller, D.M., Wiley, D.E., & Wolfe, R.G. (1986). Categorization methodology: An approach to the collection and analyses of certain classes of qualitative information. *Multivariate Behavioural Research, 21,* 135-167.

Moussaoui, K.A., & Varela, P. (2010). Exploring consumer product profiling techniques and their linkage to a quantitative descriptive analysis. *Food Quality and Preference, 21,* 1088-1099.

Nestrud, M.A., & Lawless, H.T. (2010). Perceptual mapping of apples and cheeses using projective mapping and sorting. *Journal of Sensory Studies, 25,* 390-405.

Nestrud, M.A., & Lawless, H.T. (2008). Perceptual mapping of citrus juices using projective mapping and profiling data from culinary professionals and consumers. *Food Quality and Preference, 19,* 431-438.

Niimi, J. (2013). *Cross-modal sensory interactions of taste and cheese aroma.* (Thesis, Doctor of Philosophy). University of Otago. Retrieved from http://hdl.handle.net/10523/4183.

Oppenheim, A.M. (1966). *Questionnaire design and measurement.* London. Heinemann.

Peña Y Lilo M., Latrille E., Casaubon, G., Agosin, E., Bordeu E., & Martin, N. (2005). Comparison between odour and aroma profiles of Chilean Pisco spirit. *Food Quality and Preference, 16,* 59-70.

Pagès, J. (2003). Direct collection of sensory distances: application to the evaluation of ten white wines of the Loire Valley. *Sciences des Aliments, 23*, 679-688.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire Valley. *Food Quality and Preference, 16,* 642-649.

Pagès, J., Cadoret, M. & Lê, S. (2010). The sorted napping: A new holistic approach in sensory evaluation. *Journal of Sensory Studies*, *25,* 637-358.

Perrin, L., Symoneaux, R., Maitre, I., Asselin, C., Jourjon, F., & Pages, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: case of wines from Loire Valley. *Food Quality and Preference, 19,* 1-11.

Perrin, L., & Pages, J. (2009). Construction of a product space from the Ultra-Flash profiling method: application to ten red wines from the Loire Valley. *Journal of Sensory Studies, 24,* 375-395.

Piombino, P., Nicklaus, S., Le Fur, Y., Moio, L., & Le Quere, J. (2004). Selection of products presenting given flavor characteristics: an application to wine. *American Journal of Enology and Viticulture, 55:1,* 27-34.

Pfeiffer, J.C. & Gilbert, C.C. (2008). Napping by Modality: a happy medium between analytical and holistic approaches. *Proceedings of Sensometrics Conference,* 2008, Canada.

Qannari, E.M., Cariou, V., Teillet, E., & Schich, P. (2010). SORT-CC: A procedure for the statistical treatment of free sorting data. *Food Quality and Preference, 21,* 302-308.

Quady, A.L., & Guymon, J.F. (1973). Relation of maturity, acidity and growing region of Thompson Seedless and French Colombard grapes to wine aroma and quality of brandy distillate. *American Journal of Enology and Viticulture, 24:4,* 166-175.

Reinbach, H.C., Giacalone, D., Ribeiro, L.M., Bredie, W.L.P., & Frøst, M.B. (2013). Comparison of three sensory profiling methods based on consumer perception: CATA, CATA with intensity and Napping®. *Food Quality and Preference* (2013), http://dx.doi.org/10.1016/j.foodqual.2013.02.004.

Risvik, E., McEwan, J.A., Colwill, J.S., Rogers, R., & Lyon, D.H (1994). PM: a tool for sensory analysis and consumer research. *Food Quality and Preference, 5,* 263-269.

Risvik, E., McEwan, J.A., & Rødbotten, M. (1997). Evaluation of sensory profiling and PM data. *Food Quality and Preference, 8,* 63-71.

Robert, P. & Escoufier, Y.(1976). A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Applied Statistics, 25,* 257-265.

Robinson, J. (1999). *Oxford companion to wine (2nd Edition),* Oxford, Oxford University Press.

Ross, C.F., Weller, K.M., & Alldredge, J.R. (2012). Impact of serving temperature on sensory properties of red wine as evaluated using PM by a trained panel. *Journal of Sensory Studies, 27,* 463-470.

Sáenz-Navajas, M., Campo, E., Avizcuri, J.M., Valentin, D., Fernández-Zurbano, P., & Ferreira, V. (2012). Contribution of non-volatile and aroma fractions to in-mouth sensory properties of red wines: Wine reconstitution strategies and sensory sorting task. *Analytica Chimica Acta*, *732,* 64-72.

Saint-Eve, A., Kora, E.P., & Martin, N. (2004). Impact of olfactory quality and chemical complexity of the flavouring agent on the texture of low fat stirred yogurts assessed by three different sensory methodologies. *Food Quality and Preference, 15,* 655-668.

Santos, B.A., Pollonio, M.A.R., Cruz, A.G., Messias, V.C., Monteiro, R.A., Oliveira, T.L.C., Faria, J.A.F., Freitas, M.Q., & Bolini, H.M.A. (2013). Ultra-flash profile and PM for describing sensory attributes of prebiotic mortadellas. *Food Research International, 54,* 1705-1711.

Santosa, M., Abdi, H., & Guinard, J. (2010). A modified sorting task to investigate consumer perceptions of extra virgin olive oil. *Food Quality and Preference, 21,* 881-892.

Scriven, F. M., Gains, N., Green, S.R., & Thomson, D.M.H. (1989). A contextual evaluation of alcoholic beverages using the repertory grid method. *International Journal of Food Science and Technology 24:2,* 173-192.

Solomon, G.E.A. (1997). Conceptual change and wine expertise. *The Journal of the Learning Sciences, 6(1),* 41-60.

Steger, C.L.C., & Lambrechts, M.G. (2000). The selection of yeast strains for the production of premium quality South African brandy base products. *Journal of Industrial Microbiology and Biotechnology, 24,* 431-440

Stone, H., Sidel, J., Oliver, S., Woolsey, A., & Singleton, R.C. (1974). Sensory evaluation by quantitative descriptive analysis. *Food Technology, 28:11,* 24-34.

Stone, H., & Sidel, J.L. (2004). *Sensory Evaluation Practices,* San Diego, Elsevier Academic Press.

Teillet, E., Schlich, P., Urbano, C., Cordelle, S., & Guichard, E. (2010). Sensory methodologies and the taste of water. *Food Quality and Preference, 21,* 967-976.

Thuiller, B. (2007). *Rôle du CO2 dans l'Appréciation Organoleptique de Champagnes – Expérimentation et Apports Methodologiques.* Reims, France: These de l'URCA.

Tomic, O., Luciano, G., Nilsen, A., Hyldig, G., Lorensen, K., & Næs, T. (2010). Analysing sensory panel performance in a proficiency test using the PanelCHECK software. *European Food Research and Technology*, *230,* 497-511.

Torri, L., Dinella, C., Recchia, A., Næs, T., Tuorila, H., & Monteleone, E. (2013). Projective mapping for interpreting wine aroma differences as perceived by naïve and experienced assessors. *Food Quality and Preference, 29,* 6-15.

Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012). Quick and dirty but still pretty good: a review of new descriptive methods in food science. *International Journal of Food Science and Technology, 47(8),* 1563-1578.

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International, 48,* 893-908.

Van Jaarsveld, F. P., Hattingh, S., Minnaar, P., Blom, M. (2009). Rapid induction of ageing character in brandy products – Part 1. Effects of extraction media and preparation conditions. *South African Journal of Viticulture and Enology. 30,* 1-15.

Veinand, B., Godefroy, C., Adam, C., & Delarue, J. (2011). Highlight of important products characteristics for consumers. Comparison of three sensory descriptive methods performed by consumers. *Food Quality and Preference, 22*, 474-485.

Veramendi, M., Herencia, P., Ares, G. (2013). Perfume odor categorization: to what extent trained assessors and consumers agree? *Journal of Sensory Studies, 28,* 76-89.

Vidal, L., Barreiro, C., Gomez, B., Ares, G., & Gimenez, A. (2013). Influence of information on consumers' evaluations using Check-All-That-Apply questions and sorting: A case study with milk desserts. *Journal of Sensory Studies, 28 (2),* 125-137.

Chapter 3

Research results

Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products

This manuscript was published in
Food Quality and Preference

Chapter 3: Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products

Abstract

Rapid descriptive sensory profiling methods are under active exploration in the field of sensory science. Methods such as projective mapping, sorting and Napping® are considered time and cost-effective alternatives to conventional descriptive profiling for generating sensory product maps. In this study, the feasibility of applying rapid sensory profiling methods to the sensory evaluation of high alcohol products was challenged based on the considerable sensory fatigue expected due to high alcohol content (38-43% ABV) and perceived sensory complexity. Napping® and partial napping was compared to conventional profiling on a small sample set of 6 brandies to test the basic validity of these methods. The effect of an increase in sample size on reliability, repeatability and reproducibility was also tested. The results showed that Napping® and partial napping are equally reliable for the evaluation of small sample sets (n=6) of brandy. However, partial napping was shown to be more stable with an increase in sample set size (n=10). The effect of replication and training on reliable Napping® results is also discussed.

3.1 Introduction

The topic of rapid sensory profiling methods is being actively explored in the sensory science community. One of the major motivations behind the development of methods such as projective mapping and sorting, is to provide researchers with scientifically valid tools for sample screening and understanding consumer preferences, without the extensive time and cost impact of conventional sensory profiling methods such as Quantitative Descriptive Analysis™ (Stone et al., 1974).

Projective mapping, was first introduced by Risvik et al. in 1994 as a tool for linking sensory analysis to consumer research. The method requires assessors to position samples on a sheet of paper according to sensory similarity; very similar samples are positioned close together and very different samples are positioned far apart. Several aspects of the method have since been investigated. The results from trained panels have been compared to those from consumers (Barcenas et al., 2004; Risvik et al., 1997), as well as product experts (Nestrud & Lawless, 2008). The method's repeatability (Risvik et al., 1994) and reproducibility (Kennedy & Heymann, 2009) were also evaluated. Pagès (2003) introduced Napping® as a modified version of projective mapping which is restricted to a rectangular framework, or tasting sheet, for data collection, while both square and rectangular shapes have been used for projective mapping. Napping® data have to be analysed with multiple factor analysis (MFA) without scaling the variables, thereby retaining information on how the panellists used the axes of the tasting sheet (Dehlholm et al., 2012; Pagès, 2003). In contrast, projective mapping was originally analysed by GPA and PCA (Risvik et al. 1994) and various other methods have been used by other authors (Hopfer & Heymann, 2013). Perrin et al. (2008) have introduced the concept of combining ultra flash profiling with Napping® to enrich the method with descriptive information. In this adaptation, panellists supply descriptors for each sample which are then captured on a citation basis. Another version, sorted napping, combines Napping® and a sorting procedure where panellists are instructed to indicate groups of similar samples after positioning the samples (Pagès et al., 2010). The added benefits of combining the two procedures are not well documented. Another modification, partial napping (Pfeiffer & Gilbert, 2008) requires panellists to focus on each sensory modality separately, thereby providing a middle ground between the completely holistic napping approach versus the analytical approach of conventional profiling. Partial napping was validated by Dehlholm et al. (2012). In Dehlholm's work, both Napping® and partial napping are presented as valuable methods for generating rapid descriptive results with an overall time save of approximately seven hours and eight hours respectively over conventional profiling. Considering that that the time spent on conventional profiling is spread over several days and that napping results can be generated in one day, using such methods can result in a considerable cost saving in addition to speeding up project throughput (Dehlholm et al*.,* 2012)

Projective mapping methods have been tested on a variety of product types, including: chocolates (Kennedy & Heymann, 2009; Risvik et al., 1994), dried soups (Risvik et al., 1997), ewe's milk cheeses (Barcenas et al., 2004), citrus juices (Nestrud & Lawless, 2008), apples and cheeses (Nestrud & Lawless, 2010), wine (Perrin et al., 2008) and liver pâté (Dehlholm et al., 2012). The application of projective mapping to hot served food with complex textures was also challenged and validated (Albert et al., 2011).

Brandy is distilled from fermented grapes to an alcohol content of between 38% and 43% per volume (ABV) and holds a considerable market share proportion in the South African spirit beverage category. The sensory evaluation of brandy is challenging due to the sensory fatigue caused by high alcohol content and complexity of the volatile fraction (Peña y Lillo et al*.,* 2005). Over 200 volatile compounds have been identified in brandy (Janáčová et al., 2008). Ethanol is regarded as a chemosensory irritant which causes a burning sensation in the mouth at the concentration present in brandy. These perceived burning sensations may have a strong carry-over effect that can be fatiguing (Green, 1988; Prescott & Swain-Campbell, 2000). The mild intoxicating effect of alcohol should also be taken into account. In practice, alcohol intake is often reduced by diluting the samples to an alcohol content of approximately 20% (ABV) where important volatile compounds are prominent and the irritant effect of ethanol is more tolerable, although still more intense than in wine or beer (Jack, 2003; Quady & Guymon, 1973) and expectorating the sample after evaluation. Another approach is to limit the evaluation to nosing only (Peña y Lillo et al*.,* 2005), but this can result in a loss of information on mouthfeel attributes, which may be important from consumer acceptance perspectives. Sensory fatigue limits the number of samples that a taster can evaluate, before losing the sensitivity and concentration needed to generate accurate results.

During conventional descriptive profiling methods, samples are typically evaluated in a monadic sequential manner, i.e. one sample at a time. In the case of spirit products, this allows the sensory analyst to control the breaks between samples, thereby allowing re-sensitising of the panellists' palates. Monadic sequential presentation also provides the possibility of presenting samples in an incomplete block design, thereby reducing the number of samples presented to each taster, or alternatively, in partial presentations, where all the samples are not evaluated in the same session However, projective mapping is a holistic approach and relies on comparative presentation. Thus, controlled breaks, incomplete block designs or partial presentations can not be used to accommodate more samples without affecting the tasters' performance. This introduces a problem, if projective mapping is to be used as a rapid screening tool to narrow down larger sample sets to a few samples that can be evaluated with conventional methods or in consumer testing. Furthermore, it has been proposed that the reliability with which panellists can perform rapid sensory profiling may depend on product type and complexity. Nestrud and Lawless (2010) reported that panellists provided more stable clustering configurations for apples, than for cheese, in addition to generating more attributes for apples. Furthermore, Delarue and Siefferman (2004) found that, in the case of flash profiling of strawberry yoghurts and apricot cheeses, panellists provided results more similar to conventional sensory profiling, for the former product.

Thus, this study was conducted with three goals in mind. The primary goal was to validate Napping® (referred to as global napping for the purpose of this study) as a suitable method for the sensory evaluation of brandy and other spirit beverages containing 38% ABV or more, by comparing the method with conventional descriptive profiling. As part of this goal, three objectives were identified and investigated: accuracy by comparing the positioning of duplicate samples; the panel's ability to replicate results over three sessions, i.e. repeatability; reproducibility or stability of results using different panels. The second goal was to evaluate partial napping as an alternative approach for difficult product types. The final objective was to determine whether global and partial napping would be suitable, i.e. sufficiently reliable, repeatable and reproducible, for screening large brandy sample sets.

3.2 Materials and Methods

3.2.1 Products

Since the effect of the number of brandy samples on the accuracy of the napping method was of interest, two sets of commercial South African brandies were presented in this experiment, the first consisting of six samples and the second of ten samples (Table 3.1). In Sample Set 1, Brandy 1 was presented twice as a blind replicate. Sample Set 2 consisted of the same five samples presented in the first set, along with four new samples while Brandy 6 was presented twice as a blind replicate. The sample numbers in the first sample set correspond to those in the second sample set. The decision to change the blind replicate sample from Brandy 1 to Brandy 6 was due to logistic constraints. The samples represented two of the three South African brandy styles, namely blended brandies and potstill brandies. These differ in terms of the distillation process used, alcohol content and price point, with the latter generally being positioned as more expensive offerings. The samples were stored at room temperature in their original packaging. One hour prior to tasting, the brandies were diluted to 20% ABV with odourless distilled water. The brandies were served at room temperature in standard 250 ml clear tulip-shaped tasting glasses in 30 ml units. The panellists received all the brandies simultaneously, although the brandies were arranged in different, randomised orders on the sample trays, to reduce serving order effects. The samples were coded with 3-digit codes to eliminate bias. Distilled water and cream crackers were available to the panellists for the purpose of palate cleansing.

Table 3.1 List of brandies evaluated

Presented in Sample set 1 6 brandies	Presented in Sample set 2 10 brandies	South African brandy style	Alcohol content at full strength
B 1R		Blended brandy	43%
B 1	B 1	Blended brandy	43%
B 2	B 2	Blended brandy	43%
B 3	B 3	Potstill brandy	38%
B 4	B 4	Blended brandy	43%
B 5	B 5	Blended brandy	43%
	B 6	Blended brandy	43%
	B 6R	Blended brandy	43%
	B 7	Potstill brandy	38%
	B 8	Potstill brandy	38%
	B 9	Blended brandy	43%

3.2.2 Panellists

Two panels were recruited for this study from a pool of trained panellists employed at Distell Ltd. The panellists were screened for sensory acuity according to guidelines of Stone and Sidel (1992). The screening test involved basic taste thresholds, intensity ranking, aroma recall, short term memory for aroma recall, discrimination tests and interaction in a mock panel session. The demographics and tasks of each panel are shown in Table 3.2. Panel 1 performed three sessions of global napping (GN) and partial napping (PN) on each of the sample sets indicated in Table 3.1. Panel 2 evaluated one session of GN and PN on each of the sample sets; their data were used to determine the reproducibility of the GN and PN methods. Neither of the two panels had prior experience in the napping methods and very limited experience in the sensory evaluation of spirit products, although they had experience in the evaluation of other alcoholic beverages. A three month break occurred between the napping and conventional descriptive profiling (CP) evaluations. This break ensured that the panellists were not over-familiarised with the samples, and therefore biased, after working extensively with the same sample set during the napping experiments. The CP was also performed by Panel 1. Two additional panellists, also from the pool of trained panellists at Distell Ltd., participated in the CP procedure, thereby increasing the CP panel size to twelve tasters (Table 3.2).

Table 3.2 Experiment design showing the panel demographics and order in which the brandy sample sets were evaluated

Sample set	Panel Demographic	Session 1	Session 2	Session 3 [a]	Session 4	Session 5	Session 6 [a]	Three months later
6 products	Panel 1 2 men, 8 women aged 28-60	GN [b] Rep 1	GN Rep 2	GN Rep 3	PN [c] Rep 1	PN Rep 2	PN Rep 3	
10 products	Panel 1 2 men, 8 women aged 28-60	GN Rep 1	GN Rep 2	GN Rep 3	PN Rep 1	PN Rep 2	PN Rep 3	
All brandies	Panel 1 2 men, 10 women aged 28-60							CP [d]
6 products	Panel 2 12 women aged 35-60	GN Rep R	PN Rep R					
10 products	Panel 2 12 women aged 35-60	GN Rep R	PN Rep R					

[a] Two weeks after previous session, [b] Global Napping, [c] Partial Napping, [d] Conventional descriptive profiling

3.2.3 Tasting Methodology

Three sensory methods, CP, GN and PN were used to evaluate the six brandies. The CP result was considered the control or benchmark in this study. All of the evaluations were conducted in white tasting booths with controlled air conditioning and lighting to ensure unbiased responses. The data for the CP were captured using Compusense Five Release 5.2 (www.compusense.com) while data for both GN and PN were captured on paper ballots.

3.2.3.1 Conventional descriptive profiling

The nine brandies evaluated in this experiment, plus the blind duplicates of Brandy 1 and Brandy 6, were evaluated according to the methodology described in Lawless and Heymann (1998). A sensory lexicon of defined attributes, which was used to describe the samples, was compiled during three sessions (Table 3.3). Appearance was considered to be representative of the consumer product experience and was evaluated despite the possibility of biasing the aroma and palate perceptions. Reference standards were developed to represent each of the attributes. Colour charts were not used as the different hues were quite obvious to the panel and could be measured analytically. The aroma references were presented to the panel at the start of each training session. The panel was trained further to recognise and rate the intensity of attributes in the samples over a period of eight sessions. The attributes were measured on 100 mm unstructured line scales that were anchored with the terms “none” on the left end and “intense” on the right end of the scale. Panel performance was monitored using PanelCHECK 1.40 (www.panelcheck.com). The final data were captured over three replicate sessions using a Williams Latin Square design (as generated in Compusense Five version 5.20) where all the panellists evaluated all the samples while compensating for first order carry-over effects. Each replicate was captured over two days in a partial presentation. On the first day, the panellists evaluated the first six samples in their specific order of presentation according to the Williams design and on the second day they continued with the next five samples. Thus the data were captured over a total of six sessions. The reason for the partial presentation was to reduce the number of samples tasted per session, thereby reducing sensory fatigue, while retaining a balanced complete block design. The samples were evaluated in a monadic sequential order. In total the CP took 17 sessions (eight weeks) to complete.

Table 3.3 Sensory lexicon developed to evaluate the sensory profiles of South African brandies

Attribute	Description
Amber	n/a
Yellow/gold	n/a
Green	n/a
Dried peach/apricot	Aroma and flavour typical of sun dried fruit with a slightly sour note
Citrus	Volatile, peely, oily aroma and flavour associated with orange peel or naartjies
Apple/pear	Clean, sweet, fresh aroma and flavour associated with green apples or cloudy apple juice
Raisin/Prune	Aroma and flavour of dried fruit skins, stalky, dusty, dry and sour, reminiscent of fruit cake
Honey	Natural, waxy, stuffy aroma and flavour associated with plain honey. Slightly sour
Sherry	Dry, sweet, oxidative character associated with sweet wines
Tobacco/tea/ straw	Dried grassy character with tea-like, straw, dusty, cigar box notes
Oaky	Sharp, dry resinous aroma and flavour associated with oak barrels and old books
Nutty	Aroma and flavour associated with roasted hazelnuts

Caramel	Sweet associated aroma and flavour with a caramelized, sugary note. Sticky aroma and flavour reminiscent of toffees
Vanilla	Light, sweet, fragrant aroma and flavour associated with cake batter and custard. Clean note
Chocolate	Dark chocolate aroma and flavour with a distinct cocoa note
Molasses	Dark savoury note with smoky, syrupy, Marmite-like tones
Sweet spice	General spicy aroma and flavour associated with bun spices with a cinnamon top note
Soapy	Aroma and flavour associated with unscented soap. Sharp, chemical, bitter-associated, industrial
Sweet	A taste and aftertaste on tongue stimulated by sugars
Sour	Basic taste on the tongue associated with acids
Bitter	A sharp taste and aftertaste experienced at the back of the throat, e.g. caffeine, aloe and tonic water
Alcohol burn	The burning/warming sensation caused by high levels of alcohol that lingers in the entire mouth cavity and lips
Drying	The sensation of de-lubrication in the oral cavity
Viscosity/ density	The sensation of the "thickness"/viscosity or weight of the product in the oral cavity
Smoothness	The sensation that the brandy passes easily through the mouth cavity

3.2.3.2 Global Napping

In each GN session performed in this study, the brandies were evaluated according to the Napping® procedure in combination with ultra flash profiling as described by Perrin et al. (2008). The panellists were instructed to arrange the samples on a blank A3 sheet of paper, according to similarity based on their overall sensory perception. After arranging the samples, the panellists had to write down descriptors for each brandy to explain the differences between the samples. The samples were evaluated in three sessions. The first two sessions (Rep 1 and Rep 2) were evaluated two days apart to test for immediate repeatability. A third session, Rep 3, was presented two weeks later to test for intermediate repeatability. The samples were also presented once to a second, independent, panel to test for reproducibility (Rep R).

3.2.3.3 Partial Napping

For PN, the panel repeated the above mentioned Napping® procedure three times per session, each time focussing on a different sensory modality, as described by Dehlholm et al. (2012). Each modality was evaluated on a separate A3 tasting sheet. The samples were first evaluated according to appearance, then according to aroma and then according to all in-mouth sensations: flavour, mouthfeel and basic taste.

As with the GN data, the PN data were captured over three replicate sessions (Rep 1-3) and also with a second independent panel (Rep R) as described in section 2.3.2. The PN data were collected independently from the GN data.

3.2.4 Statistical analyses

Principal component analysis (PCA) was performed on the CP data, averaged over assessors. The averages over assessors were used for the method comparison in a multiple factor analysis (MFA). For the GN and PN data, the X and Y co-ordinates on the paper ballots were measured for each product relative to the centre of the sheet. The descriptors were captured on a citation basis. A value of 1 was allocated to a sample if the descriptor was used for that sample and a 0 if the descriptor was not used. If a descriptor was used by more than one panelist, the citations were summed. The data were analysed with MFA. The co-ordinate data for each panellist were regarded as a separate data table, i.e. if n=10 then 10 tables were used as active variables in the MFA calculations. The descriptor citations were added as a single table of supplementary variables i.e. given very low weights in the MFA computation. RV coefficients were used to evaluate the consensus among panellists. RV coefficients between panellists and the MFA consensus plot >0.5 were regarded as sufficient consensus. In the event where a panellist showed insufficient consensus with the rest of the panel, the data were recalculated without that panellist's data. If the resulting product configuration differed considerably from the original product configuration, the panellist's data were excluded from the final data analysis.

Hierarchical multiple factor analysis or HMFA (Le Dien & Pagès, 2003) was used to combine data from PN modalities, as illustrated by Pfeiffer and Gilbert (2008). HMFA follows the same principle as MFA while taking the hierarchical structure of the data into account. The rationale behind the method is to scale each level of the hierarchy, starting with the lowest level in the hierarchy. In the case of PN, the hierarchical structure progresses from panellist level to modality level to session level to method level. All data analyses were performed in XLStat 2010 (Addinsoft, www.xlstat.com)

3.3 Results and discussion

3.3.1 A visual guide to interpreting RV coefficients

RV coefficients are used as a measure to determine the similarity between two product configurations. To put RV coefficients into perspective visually, Fig. 3.3.1 shows two product configurations chosen because, of all the data collected, their RV coefficient of 0.807 was the closest to 0.800. RV coefficients are used as a measure to determine the similarity between two product configurations and results showed that d an acceptable degree of similarity was obtained between GN and PN. The configuration in Fig. 3.1b is rotated slightly counter-clockwise relative to Fig. 3.1a.The overall clustering of samples are similar. Some differences can be observed; samples B4, B5, B6 and B6R are grouped in a tighter cluster in Fig. 3.1b than in Fig. 3.1a. B1 is close to B2 in Fig. 3.1a, but closer to B9 and B3 in Fig. 3.1b. When interpreting the results in the rest of the paper, one can consider that two configurations with RV coefficients larger than 0.807 will be more similar to each other than the two configurations showed in Fig. 3.1. Likewise, configurations with RV coefficients lower than 0.807 will be less similar.

3.3.2 Validation of two napping methods in the six product sample set

3.3.2.1 Conventional descriptive profiling results

Fig. 3.2 shows the PCA bi-plot obtained from the CP data. The two blind replicate samples, B1 and B1R, are situated close to each other. They were associated with vanilla aroma, citrus flavour, bitterness, sour taste and alcohol burn. B3 differed from these due to its more intense apple/pear and tobacco/straw/tea notes and honey aroma. Summarized, the products on the left of the biplot were more fruity and soapy, greener in colour and more drying.

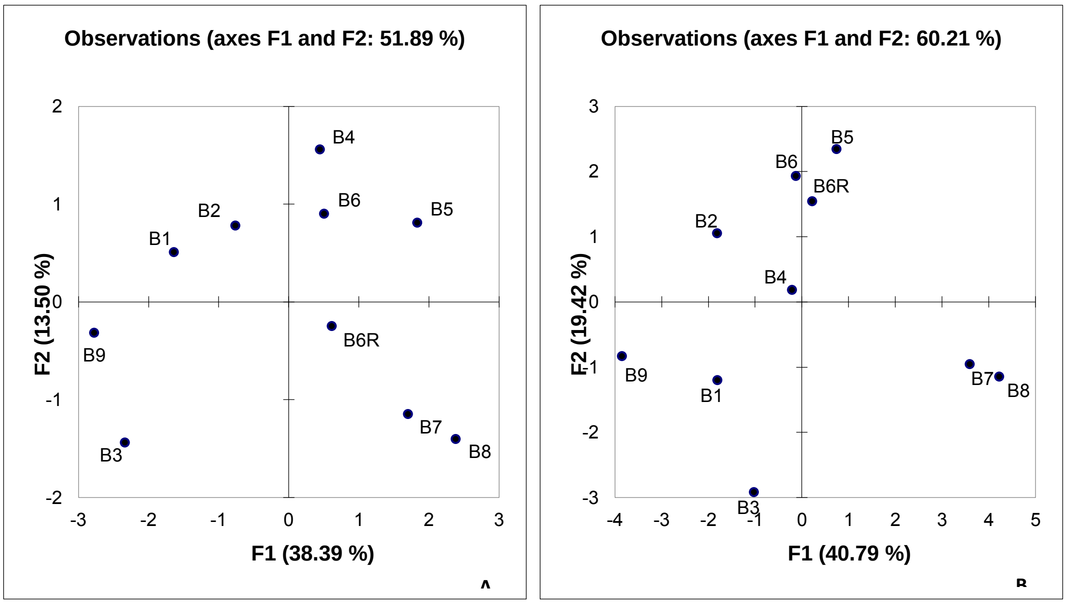

Figure 3.1 Two random product configurations chosen for their RV coefficient of 0.807, the closest to 0.800 in all the data collected. The configuration in graph A is rotated slightly counter-clockwise relative to graph B. The overall clustering of samples is similar, but with minor differences in sample positioning. B1 to B9 denote the brandy samples used. Configurations with RV coefficients exceeding 0.807 are more similar to each other than these, while configurations with lower RV coefficients are more different.

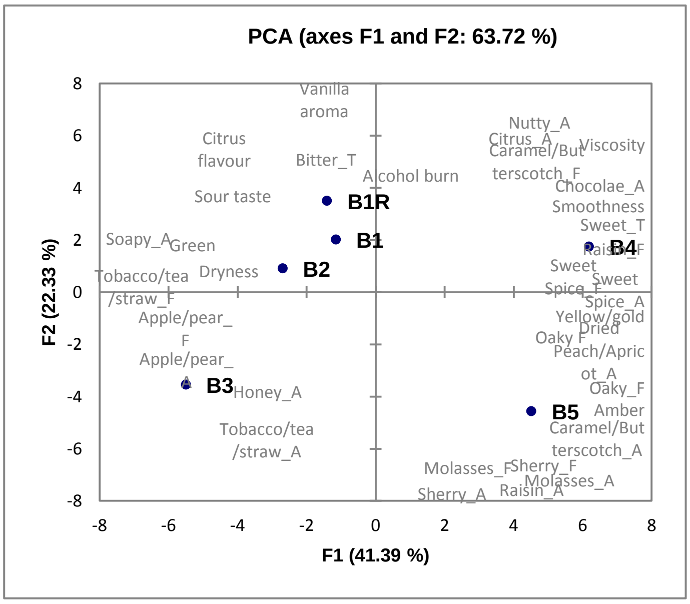

Figure 3.2 PCA biplot showing the differentiation among six brandies obtained with conventional descriptive profiling (CP)

Towards the right of the plot, the samples were associated with dried fruit and oaky aromas, more intense golden colour and sweet spice notes. B4 was associated with smoothness, viscosity and sweet taste. Relative to the other samples it had more intense chocolate and caramel/butterscotch notes. B5 had a prominent amber colour and was associated with sherry, raisins and molasses aromas and flavours, but was less smooth and viscous than B4.

3.3.2.2 Global napping results

Fig. 3.3 shows the product map obtained from the average of three GN evaluations by Panel 1. Samples B4 and B5 were associated with darker colours, described as dark gold and copper-tinged. They were also associated with rich aromas and flavours such as chocolate, coffee and apricots, as well as an oily mouthfeel. The samples on the positive end of F1, namely B1, B2 and B3, were associated with lighter, yellow colours and citrus and fresh apple flavours. The samples were also perceived as less intense. Sample B3 differentiated from the rest towards the negative end of F2 due to cigar box, earthy and herbaceous notes and less of the citrus and fresh fruit notes observed in B1. Although the attributes used to describe the products were not identical, the general sensory characters from the CP were observed in GN.

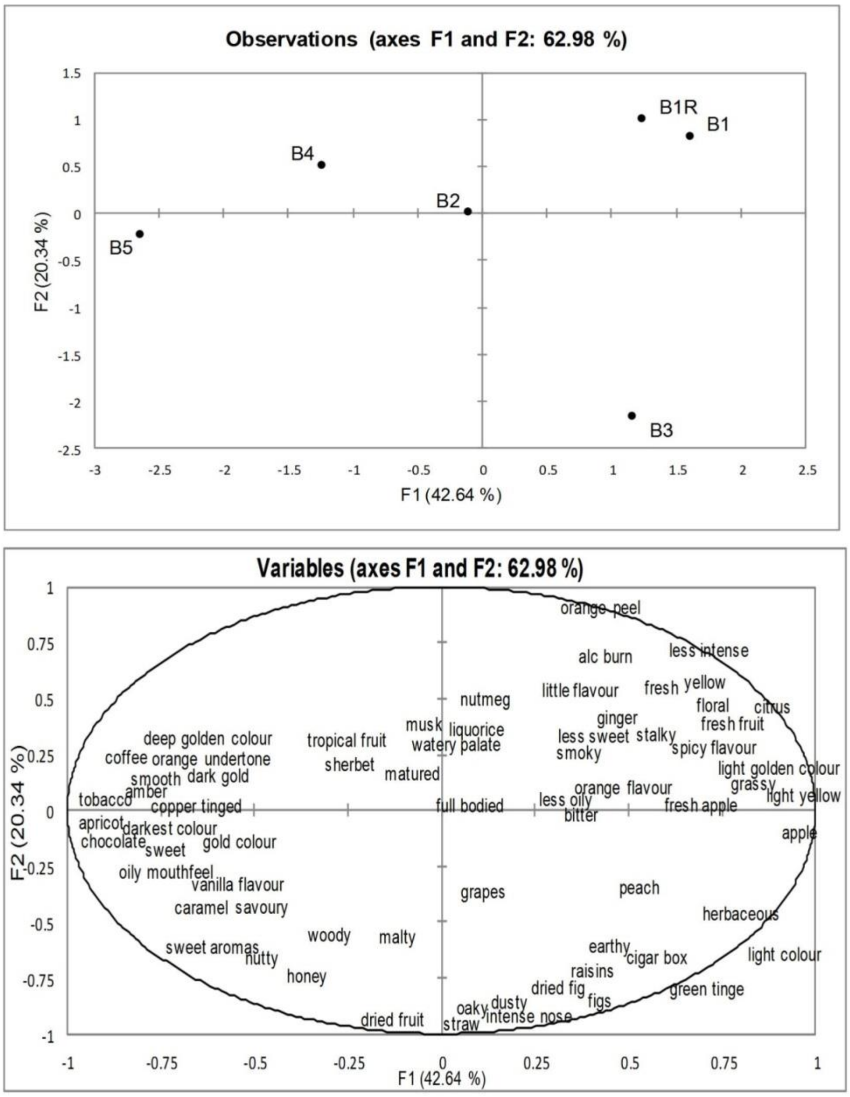

Figure 3.3 Observations and variable plots obtained from Global Napping (GP) of six brandies, B1-B5.

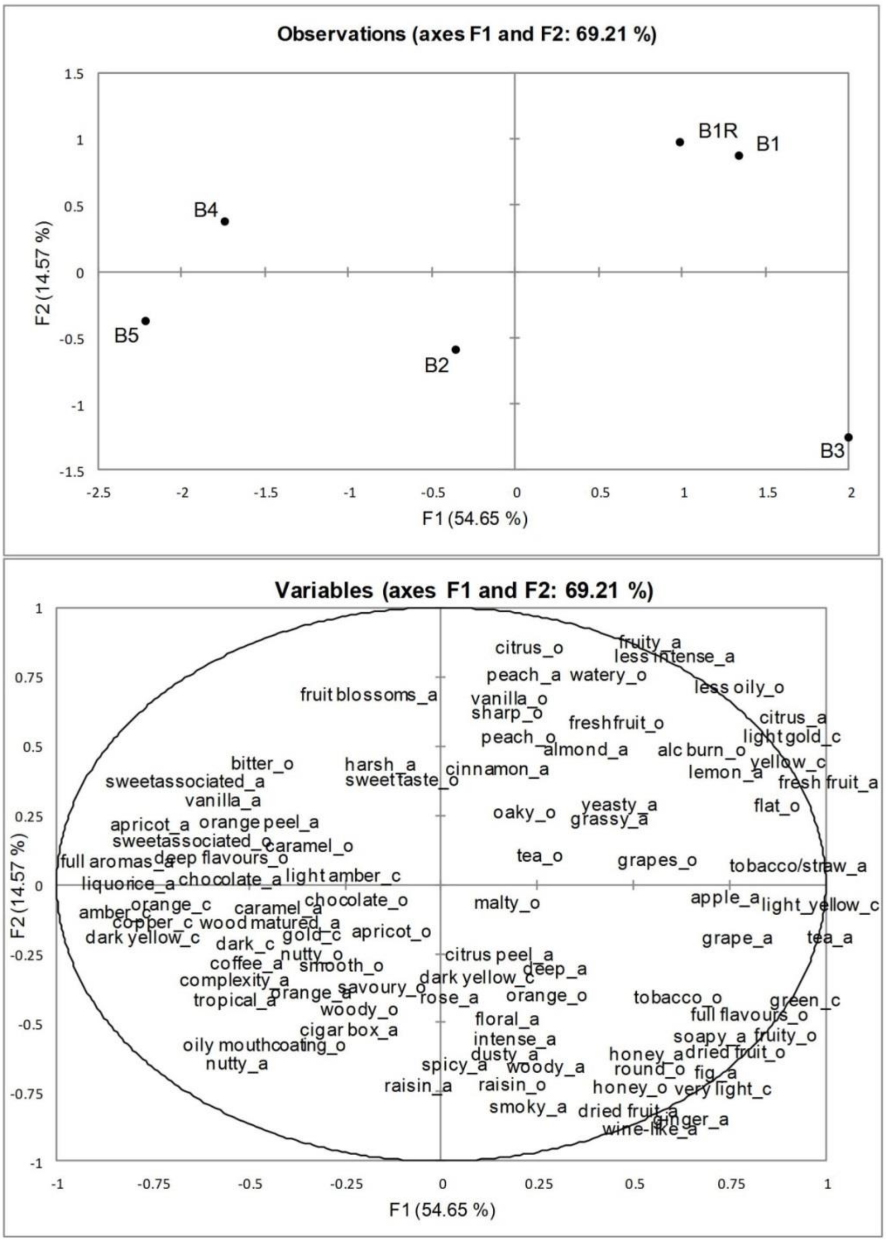

Figure 3.4 Observations and variable plots obtained from partial napping (PN) of six brandies, B1-B5.

3.3.2.3 Partial napping results

Fig. 3.4 shows PN product map. Samples B4 and B5, were again associated with dark yellow and orange colours. Aroma descriptors for these samples included chocolate, wood-matured, coffee, full aromas, nutty, liquorice and apricot. On the positive end of F1, the samples (B1, B1R and B3) were described by a lighter appearance, fresh fruit and tobacco aromas. Sample B3 on the negative end of F2 was associated with sweetness and more intense aromas, especially raisin, smoky and spicy notes. Samples at the positive end of F2, B1 and B1R, were associated with less intense characters, perceivable alcohol burn and a more watery mouthfeel. Associated aroma attributes included citrus notes, almonds, grassy and peach notes. These descriptions are similar to what was observed with CP data.

3.3.2.4 Comparison of the configurations obtained with CP, GN and PN

The RV coefficients for the three methods are shown in Table 4. The RV coefficients between the napping methods and CP were high (RV > 0.881), implying that both methods are reliable methods for the sensory evaluation of a small number of brandies. In this case, the GN configuration had a slightly higher RV coefficient relative to the CP configuration compared to the RV coefficient of PN relative to CP. This contradicts the results communicated by other authors (Dehlholm et al., 2012; Pfeiffer & Gilbert, 2008), which showed that, of the two napping methods, PN were more comparable to CP. The difference between our results and the previous publications may possibly be explained by sample set size. Pfeiffer and Gilbert evaluated eight strawberry yoghurts in their study while Dehlholm et al. evaluated nine liver pâté's, while in this part of the study only six brandies were evaluated. It may be that the sample set size was so small that the difference between GN and PN in terms of similarity relative to CP was negligible. Additionally, in Fig. 3.5a it appears as if the PN observations were mostly positioned closer to the CP observations than the GN observations. It is concluded that PN and GN are equally reliable.

Table 3.4 RV-Coefficients showing the correlation between the overall product configurations obtained with conventional descriptive profiling (CP), Partial Napping (PN) and Global Napping (GN) of 6 brandies over three sessions with two independent panels.

Panel 1				**Panel 2**			
	CP	PN	GN		CP	PN	GN
CP	1.000	0.881	0.916	CP	1.000	0.882	0.828
PN	0.881	1.000	0.939	PN	0.882	1.000	0.856
GN	0.916	0.939	1.000	GN	0.828	0.856	1.000

3.3.2.5 Positioning of the blind replicates as a measurement of accuracy

In Fig. 3.2-4, samples B1 and B1R represent replicates of the same brandy and were situated close to each other in all three plots. Thus a good level of accuracy can be obtained with GN and PN in a small sample set of brandies.

3.3.2.6 The repeatability of GN and PN as measured by Panel 1

RV coefficients were calculated between each of the three GN sessions performed by Panel 1. These coefficients ranged from 0.831 to 0.869 indicating good repeatability. The attributes used to describe the differences between the products were also very similar over the three sessions. The same approach was followed to evaluate the repeatability of the PN method. The RV coefficients between the three individual PN sessions performed by Panel 1 ranged from 0.901 to 0.923, indicating near perfect repeatability for this method. The attributes used to describe the products were also repeatable.

3.3.2.7 Evaluating the reproducibility of GN and PN by comparing results from two, independent, panels

Compared to the results from Panel 1, the results from Panel 2 showed a slightly larger difference between CP and the napping methods (Table 3.4).PN appeared to be the more reliable method. For Panel 2, the RV coefficients of GN and PN relative to CP were 0.828 and 0.882 respectively, indicating fairly good similarity to the reference method (Fig. 3.5b).

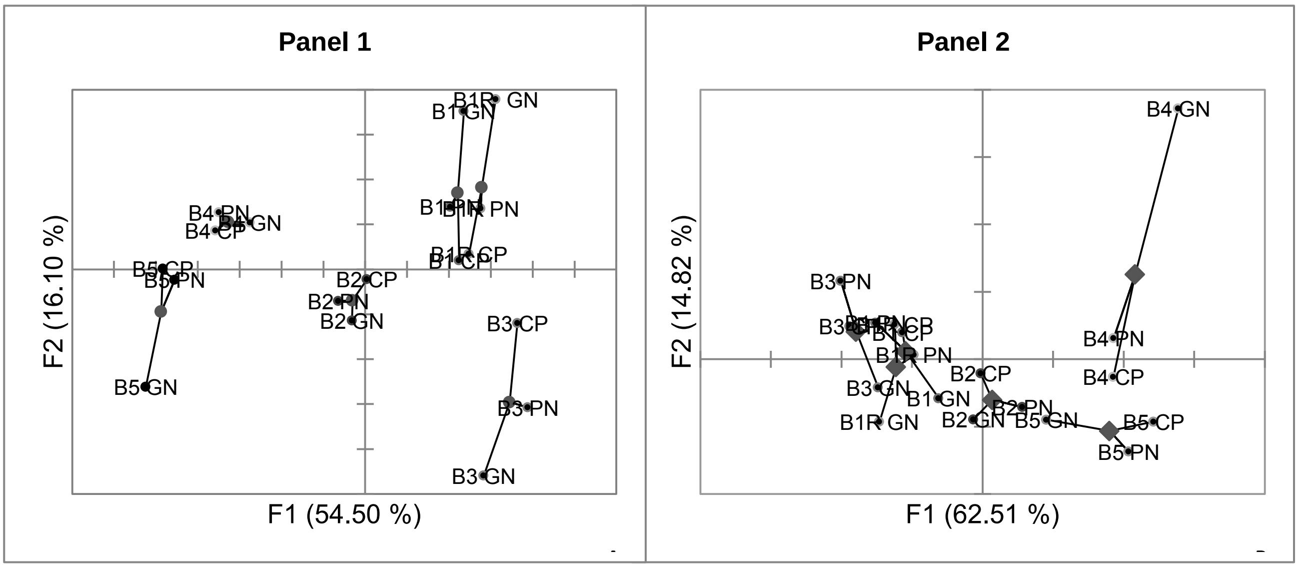

Figure 3.5 Visual comparison of two, independent, panels' evaluations of six brandies with Global and Partial Napping (GN and PN) relative to conventional descriptive profiling (CP). Both panels provided PN and GN results very similar to CP.

However, it must be considered that the CP was done by the first panel and that the two panels likely approached the sample sets differently. Therefore, the GN results of the first panel and the GN results of the second panel was compared and likewise for the PN method. In the case of GN, the Panel 2's evaluation was very similar to Rep 1 (RV=0.875) and to Rep 3 (RV=0.824) from Panel 1, but differed somewhat more from Rep 2 (RV=0.735). This indicates that GN is a reproducible method for sensory evaluation of brandies. Panel 2 tended to use more visual cues when describing the samples, while focussing less on aroma descriptors (not shown). Although their aroma descriptions were not as rich as that of Panel 1, the same general profiles for each of the six samples were obtained.

In the case of PN, the RV coefficients between all the panels' evaluations were higher than 0.909 indicating excellent reproducibility. Although the descriptive words used by Panel 2 differed from those used by Panel 1, Panel 2 expressed a similar sensory perception, i.e. that samples B1, B1R and B3 were lighter and fruitier than B4 and B5, which were darker in appearance and more sweet-associated. When required to focus on aroma as a separate modality, Panel 2 used more aroma terms to differentiate between the samples, than when they performed GN. Since the tasters on Panel 2 were less experienced than those on Panel 1, they may have relied more on the easier, visual cues to differentiate between the samples, as opposed to more complex differences in aroma and mouthfeel. Previous research showed that panellists rely heavily on visual cues in sorting tasks (Lelièvre et al., 2009) as well as descriptive tasks (Parr et al., 2003). It is likely that the same phenomenon may be observed in a Napping task.

3.3.3 The effect of sample size on the validity of GN and PN: increasing the number of samples from six to ten products

3.3.3.1 Comparing the sample configurations

When the sample size increased from six products to ten products, the difference between the napping methods and CP became more apparent. This is highlighted in Fig. 3.6, where the position of the CP observation of each sample is consistently separated from the positions of the napping observations. However, with RV coefficients >0.795, the product configurations of the three methods are still similar (Table 3.5). In this case, the PN method was more similar to CP (RV = 0.861) than the GN method

(RV=0.795) was to CP. This is in agreement with previous publications on nine liver pâtés and eight yoghurts (Dehlholm et al., 2012; Pfeiffer & Gilbert, 2008). However, as shown in section 3.2.4, the opposite was observed in the smaller sample set. It may be argued that separating the different sensory modalities had a significant effect on panel performance with the larger sample set, where it was more difficult to describe the sensory differentiation between the samples, but that the effect was negligible in the smaller and easier sample set.

3.3.3.2 Comparison of the attributes used to differentiate between samples

The CP results showed that sample B4 was perceived as more sweet, smooth and viscous mouthfeel, than the other brandies. This was not pointed out in either the GN or PN product space. In the case of GN, the panel used predominantly colour and aroma descriptors to describe B4; only 13% of the descriptors cited for B4 were based on taste and mouthfeel sensations. The panellists may have been inclined to focus on the easier modalities in a large sample set of highly fatiguing products. In the case of PN, the panel mostly described B4 using retronasal flavour descriptors in two of the three sessions, while its mouthfeel characteristics were only pointed out in one session. Thus the effect of B4's mouthfeel was weakened when the data from the three sessions were combined. By forcing the panel to discriminate between the products based on oral attributes with the PN method, the risk of losing valuable information on this modality is reduced. Nevertheless, conventional profiling methods are known to show out smaller differences between samples more accurately than rapid profiling methods, likely as a result of the intensive training involved in CP methods (Valentin et al., 2012; Varela & Ares, 2012).

Table 3.5 RV-Coefficients showing the correlation between the product configurations obtained with conventional descriptive profiling (CP), partial napping (PN) and global napping (GN) of ten brandies over three sessions.

Panel 1				**Panel 2**			
	PN	GN	CP		PN	GN	CP
PN	1.000	0.950	0.861	PN	1.000	0.746	0.752
GN	0.950	1.000	0.795	GN	0.746	1.000	0.735
CP	0.861	0.795	1.000	CP	0.752	0.735	1.000

3.3.3.3: Positioning of the blind replicates as a measure of accuracy

The relative positioning of the two blind replicates (B6 and B6R) was similar for all three methods, (Fig. 3.6). This is interesting since the panellists did not actually evaluate ten products in a single sitting during the CP. One would expect that, with less samples served in a single sitting, CP would be more accurate than GN and PN, since less sensory fatigue is introduced. One explanation might be that CP is mentally more tiring than Napping® since panellists are required to focus on quantifying subtle differences between samples rather than taking a holistic approach.

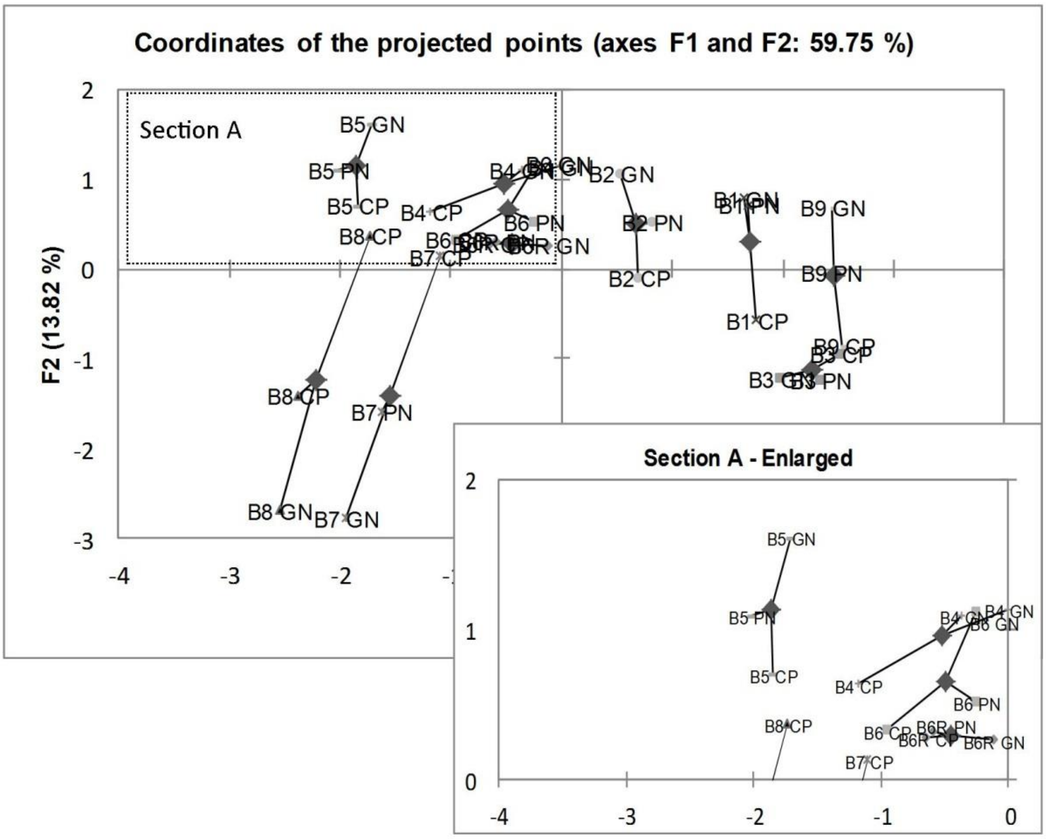

Figure 3.6 Product map obtained with MFA showing the relative sample positioning obtained with conventional descriptive profiling (CP), Global Napping (GN) and Partial Napping (PN) for a sample set of ten brandies together with an enlarged view of section A as indicated.

3.3.3.4 Evaluating the repeatability of GN and PN

The first two GN replications by Panel 1, performed two days apart, provided similar results (RV coefficient = 0.776). However, Rep 3, which was performed a month later, differed more from Rep 1 and Rep2 with RV coefficients of 0.662 and 0.618, respectively. Conversely, high RV coefficients were obtained between Rep 1, Rep 2 and Rep 3 for PN, with RV coefficients ranging from 0.814 to 0.870. This suggests that PN is more suitable for the evaluation of a large number of brandies, than GN.

3.3.3.5 Evaluating the reproducibility of GN and PN by comparing results from two independent, panels

For the larger sample set, the RV coefficients of GN and PN relative to the reference method, as evaluated by Panel 2, decreased from 0.828 and 0.882 (Table 3.4) to 0.735 and 0.752 respectively (Table 3.5). The decrease was to be expected and was also observed with Panel 1. Panel 2 failed to produce GN and PN results sufficiently comparable with the reference method.

A maximum RV coefficient of 0.725 was observed between the individual GN sessions Panel 1 and Panel 2. It seems that GN is not sufficiently repeatable and reproducible in evaluations of larger sample sets of high alcohol beverages.

In the case of PN, the evaluation by Panel 2 was more similar to the evaluations of Panel 1 ($0.786 \leq RV \leq 0.893$), than was the case with GN ($0.701 \leq RV \leq 0.725$). However, despite the higher RV coefficients obtained with PN data, a visual inspection of the data obtained by Panel 2 showed that it was not sufficiently similar to those of Panel 1. This may have been an effect of a lack of experience of the brandy samples of Panel2. By the time Panel 1 started evaluating the larger sample set with PN, they have already practiced the method several times on a smaller sample set and have already been introduced to evaluating larger sample sets with GN. Prior training of panellists on smaller sample sets may possibly increase their ability to accurately evaluate larger sample sets of brandies.

3.3.4. Managing rapid profiling of large sets of spirit products by means of replication

The conclusions reached thus far were based on the combined results of triplicate measurements. It was further asked whether performing only one replication of GN and PN with Panel 1, which also performed the CP evaluation, would provide results similar to CP. RV coefficients were calculated comparing each Panel 1's individual Napping evaluations with their CP results.

3.3.4.1: The case of Global Napping

In Sample Set 1 (6 brandies), each of the three GN replications had an RV coefficient >0.819 relative to CP, showing that one GN evaluation would provide results similar to CP. As discussed in section 3.2.6, the descriptors used in the individual GN sessions were similar to those in the CP lexicon. In Sample Set 2 (ten brandies), the RV coefficients between each of the three GN sessions and CP ranged between 0.658 and 0.765. Here it can be seen that performing replicate measurements increased the reliability of the results in a large sample set of spirit products (RV=0.889).

3.3.4.2: The case of Partial Napping

The same approach was followed for the PN. In Sample Set 1 (six brandies) the individual PN sessions were similar to the CP results (0.880≤RV≤0.904). The descriptors used in the individual PN sessions were similar to those in the CP lexicon as discussed in section 3.2.6. The similarity of the averaged PN sessions to CP (RV = 0.881), was comparable to what was observed for the three individual PN sessions. When the sample size was increased to ten products, the results from the individual sessions were still quite similar to that obtained with CP (0.834≤RV≤0.898). Considering the good repeatability of PN in such a difficult sample set, it is not surprising that the individual sessions provided results comparable to the averaged results.

3.5 Conclusions

In this study, the method of Napping® as both a holistic and modular approach has been revisited in the context of brandy products, which can be very fatiguing to evaluate. GN and PN were reliable, repeatable and reproducible in a small sample set (six brandies), although better results were generally obtained with PN. The differences between the two methods were more apparent when large sample sizes were tested. Since one of the purposes for which rapid sensory profiling methods is advocated, is the screening of large numbers of samples, it is important to determine how effective these methods can be applied to larger sample sizes, and whether they could confidently be used as screening tools for spirit products. Our results showed that PN was more suitable for the evaluation of large brandy sample sets than GN. One of the reasons for this may be that PN is a more focussed task than GN. Pfeiffer & Gilbert (2008) also reported that panellists found it easier to place and describe samples with PN than with GN. Furthermore, the results indicated that training may improve sensory panellists' ability to evaluate difficult sample sets with Napping. More detailed descriptive information may also be obtained with PN, especially with a less experienced panel.

Some suggestions can be made to simplify the management of rapid profiling methods for spirit products. Firstly, in scenarios where small sample sizes are to be evaluated, both GN and PN can be used with equal confidence. However, in larger sample sets, PN is recommended over GN. Hopfer & Heymann's (2013) recommendation to use repeated measurements is supported in this study, especially for GN. Factors to consider when deciding on a Napping approach include the expected degree of differentiation in a sample set, the familiarity of the panel with the method and products used as well as the sample set size. Although GN and PN provide less detailed results than CP, it must be considered whether quantifying subtle differences merit the considerable amount of time spent to train panellists in CP. If a project only requires an overview of the products, it is not necessary to spend weeks of training for CP to generate results that could be obtained in one or two sessions using napping.

Acknowledgements

The authors would like to express their sincere thanks to Distell Ltd for funding the research and to Muriel Adams from the Distell Sensory Research department for her role in sample preparation and data processing.

3.6 References

Albert, A., Valera, P., Salvador, A., Hough, G., & Fiszman, S. (2011). Overcoming the issues in the sensory description of hot served food with a complex texture. Appliation of QDA®, flash profiling and projective mapping using panels with different degrees of training. *Food Quality and Preference 22,* 463-473.

Barcenas, P., Pérez- Elortondo F.J., & Albisu, M. (2004). Projective mapping in sensory analysis of ewes milk cheeses: a study on consumer and trained panel performance. *Food Research International, 37,* 723-729.

Dehlholm, C., Brockhoff, P.B., Meinert, L., Aaslung, M.D., & Bredie, W.L.P. (2012). Rapid descriptive sensory methods – comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference, 26,* 267-277.

Delarue, J., & Siefferman, J.M. (2004). Sensory mapping using Flash Profile. Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. *Food Quality and Preference, 15,* 383-391.

Green, B.G. (1988). Spatial and temporal factors in the perception of ethanol irritation on the tongue. *Perception and Psychophysics, 44,* 108-116.

Hopfer, H., & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. *Food Quality and Preference, 28,* 164-181

Jack, F. (2003). Development of guidelines for the preparation and handling of sensory samples in the Scotch whisky industry. *Journal of the Institute of Brewing, 109,* 114-119.

Janáčová, A., Sadecka, J., Kohajdova, Z., & Spanik, I. (2008). The identification of aroma-active compounds in Slovak brandies using GC-sniffing, GC-MS and sensory evaluation. *Chromatographia, 67,* S113-S121.

Kennedy, J., & Heymann, H. (2009). Projective mapping and descriptive analysis of milk and dark chocolates. *Journal of Sensory Studies, 24,* 220-233.

Lawless, H.T., & Heymann, H. (1998). *Sensory evaluation of food: Principles and practices.* (1st ed). New York, NY: Chapman & Hall.

Le Dien, S., & Pagès, J. (2003). Hierarchical multiple factor analysis: application to the comparison of sensory profiles. *Food Quality and Preference, 14,* 397-403.

Lelièvre, M., Chollet, S., Abdi, H., & Valentin, D. (2009). Beer-trained and untrained assessors rely more on vision than on taste when they categorize beers. *Chemosensory Perception, 2,* 143-153.

Nestrud, M.A., & Lawless, H.T. (2008). Perceptual mapping of citrus juices using projective mapping and profiling data from culinary professionals and consumers. *Food Quality and Preference, 19,* 431-438.

Nestrud, M.A., & Lawless, H.T. (2010). Perceptual mapping of apples and cheeses using projective mapping and sorting. *Journal of Sensory Studies, 25,* 390-405.

Pagès, J. (2003). Direct collection of sensory distances: application to the evaluation of ten white wines of the Loire Valley. *Sciences des Aliments, 23,* 679-688.

Pagès, J., Cadoret, M. & Lê, S. (2010). The sorted napping: A new holistic approach in sensory evaluation. *Journal of Sensory Studies, 25,* 637-358.

Parr, W., White, K.G., & Heatherbell, D.A. (2003). The nose knows: Influence of colour on perception of wine aroma. *Journal of Wine Research, 14,* 79-101.

Peña Y Lilo, M., Latrille, E., Casaubon, G., Agosin, E., & Bordeu E. (2005). Comparison between odour and aroma profiles of Chilean Pisco spirit. *Food Quality and Preference, 16,* 59-70.

Perrin, L., Symoneaux, R., Maitre, I., Asselin, C., Jourjon, F., & Pages, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: case of wines from Loire Valley. *Food Quality and Preference, 19,* 1-11.

Pfeiffer, J.C., & Gilbert, C.C. (2008). Napping by Modality: a happy medium between analytical and holistic approaches. *Proceedings of Sensometrics Conference, Canada, 2008.*

Prescott, J., & Swain-Campbell, N. (2000). Responses to repeated oral irritation by capsaicin, cinnamaldehyde and ethanol in PROP tasters and non-tasters*. Chemical Senses, 25,* 239-246.

Quady, A.L., & Guymon, J.F. (1973). Relation of maturity, acidity and growing region of Thompson Seedless and French Colombar(d) grapes to wine aroma and quality of brandy distillate. *American Journal of Enology and Viticulture, 24 (4),* 166-175.

Risvik, E., McEwan, J.A., Colwill, J.S., Rogers, R., & Lyon, D.H. (1994). Projective mapping: a tool for sensory analysis and consumer research. *Food Quality and Preference, 5,* 263-269.

Risvik, E., McEwan, J.A., & Rødbotten, M. (1997). Evaluation of sensory profiling and projective mapping data. *Food Quality and Preference, 8,* 63-71.

Stone, H., Sidel, J., Oliver, S., Woolsey, A., & Singleton, R.C. (1974). Sensory evaluation by quantitative descriptive analysis. *Food Technology, 28:11,* 24-34.

Stone, H., & Sidel, J.L. (1992). *Sensory Evaluation Practices.* (2nd ed). Orlando, FL: Academic Press.

Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012). Quick and dirty but still pretty good: a review of new descriptive methods in food science. *International Journal of Food Science and Technology, 47(8),* 1563-1578.

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International, 48,* 893-908.

Chapter 4

Research results

Optimisation of the partial napping approach for the successful capturing of mouthfeel differentiation between brandy products

Chapter 4: Optimisation of the partial napping approach for the successful capturing of mouthfeel differentiation between brandy products

Abstract

Partial napping has been validated as a suitable sensory profiling method for brandy evaluation. However, it was found that, compared to conventional profiling, very little useful information could be extracted on brandy mouthfeel when it was evaluated as part of overall in-mouth perceptions. This study aimed to optimise the partial napping method to improve information output on the mouthfeel of brandies. Panellists' proficiency in visual, aroma and in-mouth evaluation of brandies were scrutinised after which three partial napping protocols were tested to identify the most effective solution for the successful capturing of mouthfeel differences between brandies. The results showed that panellists were equally efficient in aroma and in-mouth evaluations, but that in-mouth perception (defined as retronasal flavour, basic taste and mouthfeel) was not a useful construct as it did not contribute to the product configuration that could be obtained with visual and colour assessments alone. Instructing panellists to ignore retronasal flavour delivered more useful results. Using dark glasses and nose-clips to eliminate visual, aroma and retronasal flavour perceptions were not necessary to obtain a reliable and interpretable representation of the mouthfeel differences between brandies. Clear glasses and written instructions were sufficient to generate useful mouthfeel information under conditions more representative of the consumer product experience.

4.1 Introduction

Brandy is a distilled spirit made from fermented grapes; in fact, the word 'brandy' is derived from the Dutch word 'brandewijn' meaning 'burnt wine' (Toerien, 2008). Brandy flavour is influenced during the entire production process. Some flavour compounds present in brandy originate from the inherent aromatic chemical composition of the grapes while others are formed during alcoholic fermentation and it is these compounds that are concentrated during distillation. The extraction of flavour molecules from wood during maturation and finally the evolution of all the above mentioned compounds during ageing further contribute to the final flavour of brandy (Louw & Lambrechts, 2012).

The origin of aroma compounds in brandy can be divided into two categories: those that are formed during base wine production and concentrated during distillation, and those that are extracted or formed during wood maturation. The former, largely consistent of esters, higher alcohols and volatile fatty acids, impacts on the fruity and floral odours. The wood derived aroma compounds, mostly oak lactones, phenolic aldehydes and furanic aldehydes, are generally associated with spicy, woody, sweet-associated and nutty sensory notes (Louw & Lambrechts, 2012). Brandy mouthfeel develops during wood maturation as low molecular weight and hydrolysable tannins are extracted from the oak, and these compounds impact on smoothness, burning, astringency, bitterness and body (Caldeira et al., 2006; Caldeira et al., 2010; Canas et al., 2009). As brandy matures, its body and flavour complexity increases while astringency and alcohol burn decrease. Wood maturation can also affect brandy colour. Typical brandy colour hues include straw-yellow, golden, topaz and greenish (Canas et al., 2009).

Partial napping is a structured version of the projective mapping procedure where panellists are required to perform the projective mapping task based on only one sensory modality (Pfeiffer & Gilbert, 2008). The idea was suggested as a means to bypass any loss of information caused by panellists being forced to transfer multi-dimensional sensory perceptions onto a two-dimensional sheet of paper (Pagès, 2005). This concept was later presented under the name of partial napping (Pfeiffer & Gilbert, 2008), as a middle ground between the intuitive napping approach and the analytical conventional profiling approach. It has initially been applied in step-wise manner by instructing panellists to focus on several separate modalities after which the data are used to construct a consensus map based on all the modalities (Dehlholm et al., 2012; Louw et al., 2013; Pfeiffer & Gilbert, 2008). However, it can also be applied as a restrictive method where panellists are forced to focus on a specific modality and to disregard other modalities completely (Giacalone et al., 2013; Grygorczyk et al., 2013). In such a case no data would be generated on the excluded modalities.

Partial napping has been validated as a reliable and robust screening tool for large sample sets (Louw et al., 2013). The limitations of partial napping for application to brandy were that 1) it generates a lot of data that has to be processed by analyst and 2) differences in mouthfeel were not extracted as successfully as with conventional profiling. It was shown that colour, aroma and retronasal flavour were the dominant differentiating factor between brandies in the mind of the panellists (Louw et al., 2013). Other researchers have suggested that nosing is sufficient to capture the most important sensory characteristics of brandy and whisky due to strong correlations observed between orthonasal and retronasal odour perception (Jack, 2003; Pena y Lilo et al., 2005). However, mouthfeel potentially plays a role in consumer acceptability of spirit products. In fact, references to mouthfeel characteristics is often used to describe products in brand communication; terms such as full-bodied, mellow, soft, round, silky and smooth across the spirit category. Depending on the research objective, it can be critical to obtain information on the mouthfeel perception of brandy.

The principal aim of this study is to optimise the partial napping procedure for brandy evaluation to facilitate more effective capturing of differentiation on mouthfeel. To obtain this goal, the partial napping data from our previous work (Louw et al., 2013) will be scrutinised to determine the panellists' proficiency in the three sensory evaluation modes: visual, orthonasal aroma and in-mouth sensations (consisting of basic taste, mouthfeel and retronasal flavour perception) from which areas for optimisation will be identified. Based on the information gathered, three partial napping protocols will be tested to identify a more effective solution for capturing mouthfeel differences between brandy samples.

4.2 Materials and methods

This study was conducted in two phases. The first phase, referring to all work done on Sample Set 1 (Table 1), is an extension on the work published in Louw et al. (2013) and the methodology for this phase is thoroughly described in said publication. This phase involved additional, explorative, data analyses on the data from Louw et al. (2013) to gain an understanding of the sensory panellists proficiency in evaluating brandy based on three sensory modalities, using the Napping® method. The second phase, referring to all work on Sample Set 2 (Table 4.1), was executed two years after the first phase in response to the outcomes from the first phase. This second phase tested three different partial napping protocols where the perception of colour and odour were excluded in a step-wise manner by using first clear glasses (all sensory modalities perceivable as in normal method), then dark glasses (visual perception excluded) and dark glasses with nose clips (visual and odour perception excluded) in order to identify the best approach for successful elicitation of brandy mouthfeel during partial napping.

4.2.1 Products

Three sets of South African brandy samples were presented in this experiment (Table 4.1). Sample set 1 was also evaluated in Louw et al. (2013). Sample Set 2 consisted of ten brandies unrelated to those in Sample Set 1. In order to assess the accuracy of the results from each sensory modality, each set contained two duplicate samples. The duplicated samples were specifically chosen to present an average profile for the category, based on a bench top evaluation. In Sample Set 1, B6 was presented twice. In Sample Set 2, C9 was presented twice. The samples represented two of the three South African brandy styles, namely blended brandies and potstill brandies. These styles differ in terms of alcohol content; blended brandies contain 43% ABV (alcohol by volume) and potstill brandies 38% ABV. Furthermore, potstill brandies must be 100% pot distilled. Blended brandies are required to contain at least 30% pot distilled distillate, while the remainder may be column distilled, resulting in a less flavourful profile (South African Department of Agriculture, 1989). Blended brandies are sold at a lower price point than potstill brandies. The samples were stored at room temperature in their original packaging. The sample preparation and serving practices were done according to standard sensory practices as described in our previous work (Louw et al., 2013).

Table 4.1 List of brandies evaluated.

Set 1	South African brandy style	Set 2	South African brandy style
B 1	Blended	C1	Potstill
B 2	Blended	C2	Blended
B 3	Potstill	C3	Potstill
B 4	Blended	C4	Potstill
B 5	Blended	C5	Blended
B 6[a]	Blended	C6	Potstill
B 6R[a]	Blended	C7	Potstill
B 7	Potstill	C8	Blended
B 8	Potstill	C9[b]	Blended
B 9	Blended	C9R[b]	Blended

[a]B6 and B6R are duplicates of the same brand; [b]C9 and C9R are duplicates of the same brand.

4.2.2 Panellists and tasting sessions

The panellists for this study were recruited from a pool of trained panellists employed at Distell Ltd. They were screened for sensory acuity according to the guidelines of Stone and Sidel (1992). The screening test involved basic taste thresholds, intensity ranking, aroma recall, short term memory for aroma recall, discrimination tests and interaction in a mock panel session. Ten panellists participated in the partial napping (PN) of Set 1. Two additional panellists participated in the conventional profiling (CP) of these samples, which commenced three months after the PN sessions were completed, in order to reduce bias caused by over-familiarisation with the samples tested. The PN data for Set 2 were captured in a separate experiment which took place two years after the data for Set 1. The panel consisted of twelve women, four of whom also participated in PN of Set 1. The remaining eight panellists were recruited and screened as described above.

4.2.3. Tasting methodology

Two sensory methods, CP, PN were used to evaluate the brandies. All of the evaluations were conducted in white tasting booths with controlled air conditioning and lighting to ensure unbiased responses. CP data were captured using Compusense Five Release 5.2 (www.compusense.com), while PN data were captured on white, unmarked A3-sized paper ballots.

4.2.3.1 Conventional descriptive profiling

Sample Set 1, (samples B1- B9 plus two replicates, Table 1) was evaluated according to standard descriptive profiling procedures (Lawless & Heymann, 1998); the methodology is described in our previous work (Louw et al., 2013). The lexicon included visual, aroma, basic taste, mouthfeel and retronasal flavour descriptors. Three replicate sessions were captured using a Williams Latin Square design (generated in Compusense Five version 5.20). Each replicate was captured over two days in a partial presentation, i.e. on the first day, the panellists evaluated the first six samples in their specific monadic sequential order of presentation according to the Williams design, and on the second day they continued with the next five samples. The reason for the partial presentation was to reduce sensory fatigue, while retaining a balanced complete block design.

4.2.3.2. Partial napping

For PN the brandies from Sample Set 1 were evaluated three times using the Napping® procedure in combination with ultra flash profiling where panellists were instructed to provide descriptive terms after positioning their samples on the tasting sheet (Perrin et al. 2008). In the first evaluation the panellists were instructed to evaluate the samples based on appearance only. In the second and third evaluations, they were instructed to focus on aroma (orthonasal) and in-mouth sensations (retronasal flavour, mouthfeel and basic taste) respectively. Thus, three modalities, "visual", "aroma" and "in-mouth", were captured on separate tasting sheets in the same sitting. This procedure was repeated three times. The first two sessions (Rep 1 and Rep 2) were evaluated two days apart and a third session, Rep 3, was captured two weeks later.

For Sample Set 2, two modalities were captured namely, "aroma" and "taste/mouthfeel", where "taste/mouthfeel" was limited to basic tastes (i.e. sweet, sour, salty, bitter, umami) and mouthfeel attributes. The data were captured over six sessions. In the first two sessions, the panellists received the samples in clear glasses and were asked to first evaluate the samples based on aroma and then on basic taste and mouthfeel. They were instructed to ignore retronasal flavours. In the next two sessions, the samples were served in black glasses and the panellists were again asked to first complete the aroma evaluation and then the basic taste and mouthfeel evaluation (again ignoring flavour). In the final two sessions, the panellists received a set of black glasses with nose clips and they were asked to evaluate the samples based on taste and mouthfeel. After this, they received a second set of black glasses of the sample products but with different blinding codes and were asked to evaluate the samples on aroma. Using these approaches, three sets of data, with two replicates each, were captured; clear glasses (CG), black glasses (DG) and nose clips (NC).

4.2.4. Statistical analyses

A detailed description of the PN data processing and multivariate statistics used for this study is provided in our previous work (Louw et al*.,* 2013) Data analyses were performed in XLStat 2012 (Addinsoft, www.xlstat.com). Principal component analysis (PCA) was performed on the CP data, averaged over assessors. The averages over assessors were used for the method comparison in a multiple factor analysis (MFA). For the PN data, the X and Y co-ordinates on the paper ballots were measured for each product relative to the centre of the sheet. The descriptors were captured on a citation basis. The data were analysed with MFA. The descriptor citations were added as passive or supplementary variables. RV coefficients were used to evaluate the consensus among panellists (Abdi, 2007; Robert & Escoufier, 1976).

Hierarchical multiple factor analysis or HMFA (Le Dien and Pagès, 2003) was used to combine data from the PN modalities, a strategy illustrated in an earlier study by Pfeiffer and Gilbert (2008). RV coefficients were used as a measure to determine how similar two product configurations were relative to each other, using the guideline that an RV coefficient above 0.700 indicates good similarity (Cartier, 2006).
The distance (i.e. perceived difference) between the two duplicate samples relative to the maximum inter-sample distance, expressed as a distance ratio, D_r, was taken as a measure of accuracy (Torri et al., 2013). For the overall consensus configurations, this measure was based on the Euclidean distances between the samples on the MFA consensus configuration. For the panellists, the ratio was based on their raw co-ordinate data. Analysis of Variance (ANOVA) was performed to determine significant differences in panellist accuracy.

Panellist performance was measured using the relative performance indicator (RPI) (Louw et al., submitted for publication). This indicator is based on explained variance after procrustes data transformation and provides a comparison of the product maps generated from each panellist's data. The RPI is calculated using Equation 1

EQ 1: $$Relative\ performance\ index = \frac{\left(sum\ of\ variances - \frac{SSQ}{n_{samples} \times n_{configurations}}\right)}{sum\ of\ variances}$$

Where SSQ = Residual Sum of Squares from procrustes ANOVA after compensating for rotation, translation and scaling during GPA and n = number of samples or configurations, as annotated.

4.3 Results and discussion

4.3.1 Comparison of the visual, aroma and in-mouth partial napping of ten brandies

The panel's performance in three sensory evaluation modes: visual, orthonasal aroma and in-mouth sensations (consisting of basic taste, mouthfeel and retronasal flavour perceptions) were evaluated to identify areas in which the partial napping method for brandy can be improved to illustrate mouthfeel differences more effectively. Results in small and large sample sets, representing an easy and a difficult task respectively, were evaluated in terms of how it compares to conventional profiling, accuracy, repeatability and performance of individual panellists.

4.3.1.1 Comparison of partial napping modalities to conventional profiling

RV coefficients were computed to determine the similarity between product configurations obtained by partial napping and conventional profiling of Sample Set 1, based on visual, aroma and in-mouth perceptions. The panel's visual PN assessments were the most similar to CP results (RV=0.894) based on the same modality, followed by their aroma assessments (RV=0.719). Their in-mouth PN assessments were the least similar to the CP result based on in-mouth attributes (RV=0.657).

Figure 4.1 compares CP with PN in terms of in-mouth perceptions for Sample Set 1. There is no apparent relationship between the CP and PN configurations. For CP, the first dimension was described by a fruity/soapy/tobacco flavour and sour taste direction to the negative end of F1 and a smooth/viscous and sweet associated dimension to the positive. On F2, there was an alcohol burn/drying/bitter/dried peach/apricot direction towards the positive end and a honey/molasses flavour towards the negative end. The first two factors of the PN configuration are difficult to interpret; there were very few descriptors that were strongly correlated (correlation coefficient >0.5) with either F1 or F2. For F1, these were raisin, spicy, cloves, almond and sweet associated flavour on the positive end and none on the negative end. For F2, these were alcohol burn, smoothness and peach flavour on the positive end and vanilla flavour on the negative end. The apparent relationship between alcohol burn and smoothness on F2 seem surprising, but these attributes separate from each other on F3 (not shown).

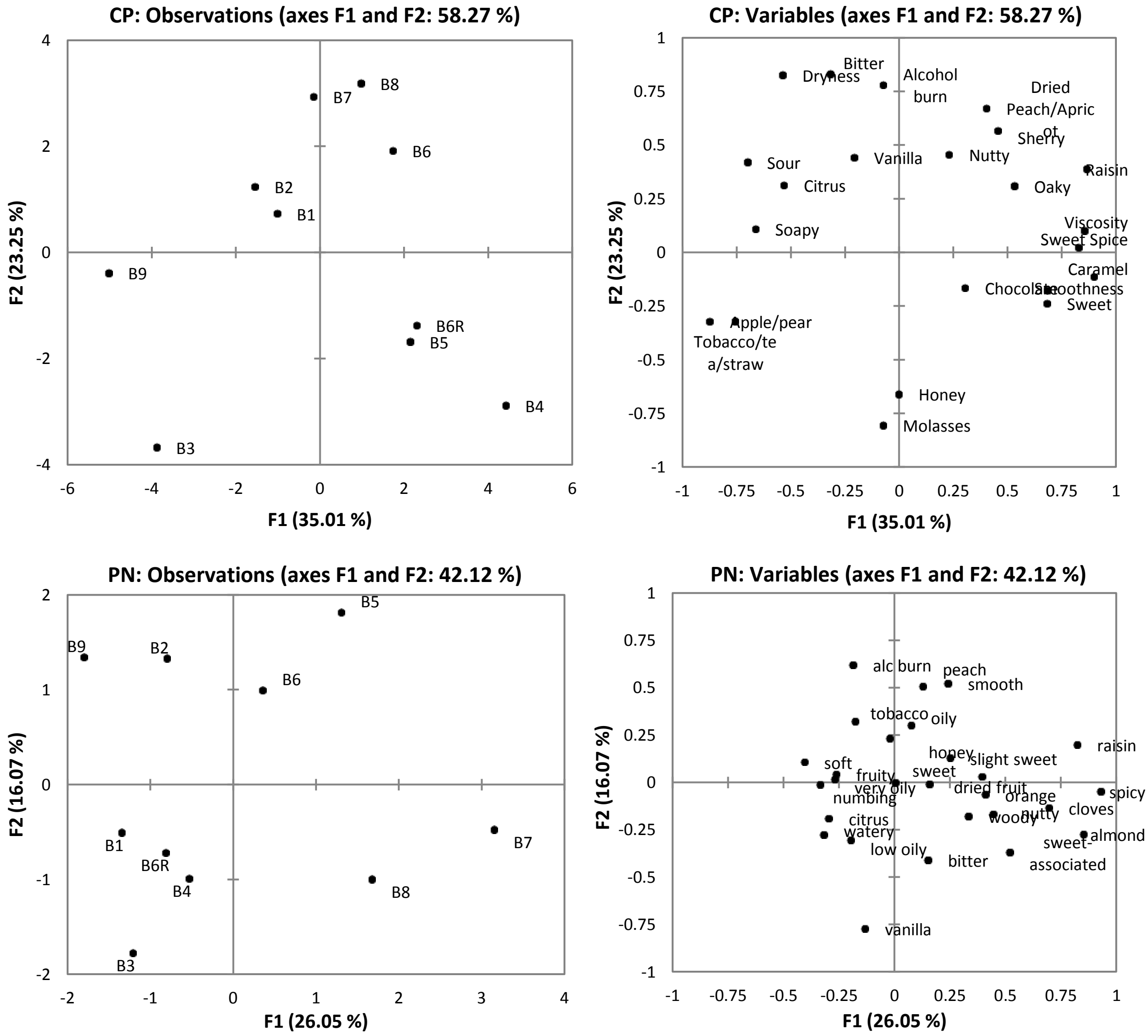

Figure 4.1 Observation and variable plots showing product configurations obtained with conventional profiling (CP) and partial napping (PN) of ten brandies based on in-mouth sensations (i.e. retronasal flavour, basic tastes and mouthfeel). B6 and B6R are duplicates of the same brandy.

The lack of structure in the first two dimensions of the PN configuration could possibly be explained by visual bias. A previous study on beer showed that panellists are heavily influenced by visual cues during sorting exercises (Lelièvre et al., 2009). In fact, when the colour attributes generated during the visual PN exercise were superimposed on the PN in-mouth sample plot (Figure 4.2), it appears as if visual perception influenced the sample differentiation on F1: the negative end of F1 was strongly associated with light colours, while the positive end was associated with dark colours. That being said, the CP results relating to in-mouth descriptors were also influenced by colour; but to a lesser extent (not shown). Several studies have highlighted that people, often subconsciously, seek congruency between what they see and what they smell (Morrot et al., 2001; Lelièvre et al., 2009). With the sense of sight carrying the most weight cognitively, odour associations are often made between the colour of the product and objects of a similar colour (e.g. a red wine smelling like red fruits) (Morrot et al., 2001; Parr et al., 2003). In the case of wine, processes such as oxidation may affect both colour and aroma, therefore, to a certain extent; the influence of colour on odour descriptors may be valid. In contrast, brandy colour may be rectified before bottling with additions of flavourless caramel to obtain the desired colour hue (European Union, 2008; Louw & Lambrechts, 2012; South African Department of Agriculture, 1989); the product colour is therefore, not necessarily related to the volatile composition. However, in view of the odour-colour association concept presented by Morrot et al. (2001) and also Parr et al. (2003), it is possible that lighter yellow brandy (B1 and B9 in Figure 4.2) may be associated with fresh fruits flavours like citrus (Figure 4.1, bottom), while a dark amber brandy (B7 in Figure 4.2) may be associated with raisin flavour (Figure 4.1, bottom), simply due to the colour associations with the respective fruits.

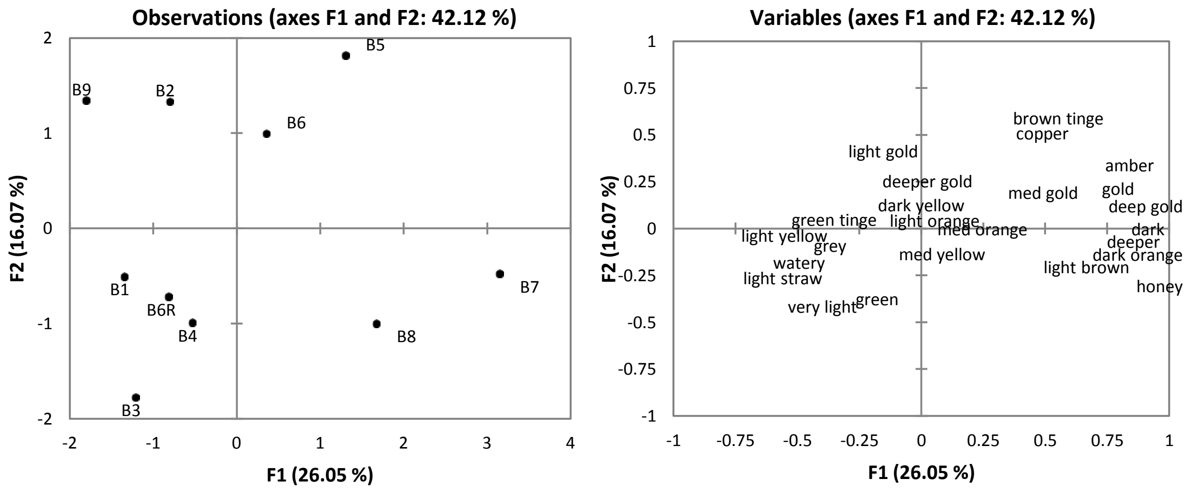

Figure 4.2 Descriptors (right) from visual assessment of ten brandies superimposed onto product configuration (left) from partial napping of the same brandies based on in-mouth sensations. Dimension F1 is strongly associated with colour intensity.

4.3.1.2 Accuracy: Distance between duplicate samples

Accuracy was measured as a ratio (D_r) between the Euclidean distances between the two duplicate samples and the maximum inter-sample distance in each product set. Values approaching 0 indicates good accuracy. The panel's accuracy in visual (D_r=0.28) and aroma assessments (D_r=0.29) were comparable. Their evaluation of in-mouth perceptions were less accurate (D_r=0.39) of the three modalities.

4.3.1.3 Repeatability: RV coefficients and visual evaluation of consensus configurations from three repeated sessions

To determine whether the panel was more repeatable in one of the modalities compared to the others, the RV coefficients between sessions for each of the modalities were calculated. For the visual assessments, the RV coefficients between repeated evaluations of Sample Set 1 exceeded 0.900, indicating that the panel was extremely repeatable in their visual evaluations. Taking 0.700 as a cut-off for good repeatability, the panel was not very repeatable in their aroma assessments (RV: 0.510-0.700). A similar lack of repeatability was observed for the in-mouth sensations (RV: 0.488-0.669). Yet, the PN consensus configurations taking both aroma and in-mouth sensations into account was repeatable (RV: 0.725-0.815); the same conclusions can be drawn from Rep 1, Rep 2 and Rep 3. The evaluation of a single sensory modality in a restricted PN exercise can be compared to GN in terms of repeatability. This is interesting; it appears as if the poor repeatability of GN in large sample sets (Louw et al., 2013) may not be related to the fact the amount of sensory information that the panellists need to take into account (all sensory perceptions in GN vs. just one sensory modality in restricted PN). Not much is known about the reasons for poor repeatability in projective mapping; these results suggest that a holistic task structure does not necessarily affect repeatability.

4.3.1.4 Panellist performance

Individual differences in panel performance were evident from the results (Figures 4.3). The panellists were very consistent in their visual evaluations. Poor panellist consistency was observed for aroma and in-mouth assessments. Panellists 2 and 6 were more consistent in their aroma assessments than in their in-mouth assessments, while panellists 3, 5 and 7 were more consistent in their in-mouth assessments than in their aroma assessments.

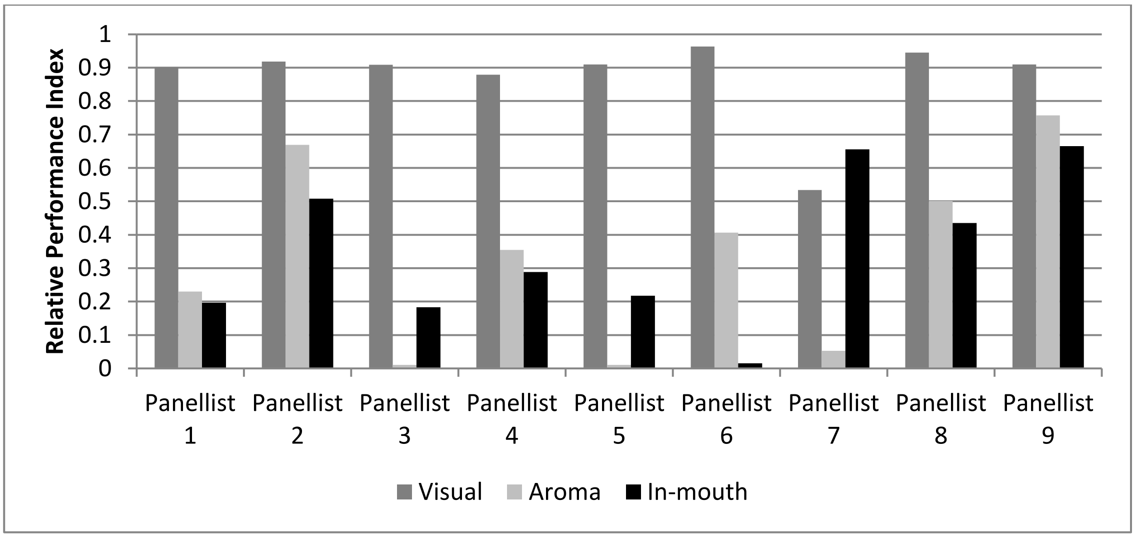

Figure 4.3 Relative performance index showing the performance of nine trained panellists' evaluations of visual, aroma and in-mouth perceptions of ten brandies during partial napping. Error bars are not shown since only one observation is collected.

Table 4.2 Panellists' accuracy in the projective mapping task expressed as a ratio of the Euclidean distance between two duplicated samples and the maximum inter-sample distance (D_r). Results are shown for assessments of two sets of brandies, consisting of six and ten brandies each, based on either visual, aroma or in-mouth sensations.

Panellist	p-value	Visual	Aroma	In-mouth
Panellist 1	0.043	0.09 b[1]	0.52 a	0.51 a
Panellist 2	0.042	0.07 b	0.19 ab	0.36 a
Panellist 3	0.465	0.34 a	0.40 a	0.57 a
Panellist 4	0.365	0.16 a	0.40 a	0.44 a
Panellist 5	0.121	0.15 a	0.47 a	0.59 a
Panellist 6	0.002	0.09 b	0.33 b	0.76 a
Panellist 7	0.026	0.21 b	0.60 a	0.76 a
Panellist 8	0.223	0.72 a	0.60 a	0.74 a

1: Letters denote Fisher LSD values

Table 4.2 compares eight panellists' accuracy in positioning duplicate samples close together in visual aroma and in-mouth assessments as measured by the D_r ratio. All the panellists were the most accurate in their visual assessments. Significant differences were observed for half the panellists. Most panellists were equally accurate in their aroma and in-mouth assessments, with the exception of Panellist 6 who was significantly less accurate in her in-mouth assessments.

4.3.1.5 Contribution of in-mouth perception to the overall product configuration

For each panellist, product configurations were generated based on all three modalities and another set of configurations based only on aroma and appearance. RV coefficients were calculated between these configurations to determine the contribution of the in-mouth modality to the overall product configuration obtained from each panellist. Very high RV coefficients (0.880-0.963) were observed throughout, indicating that in-mouth perceptions had limited influence on the final product configurations. This also translated to the overall consensus configuration where the inclusion of in-mouth perceptions had a negligible on effect on the product configuration. This is consistent with previous reports of strong correlations between ortho-nasal and retronasal odour perceptions in wines and Pisco brandies (Aubry et al., 1999; Peña y Lilo et al., 1995). It has been suggested that nosing is a sufficient proxy for retronasal flavour evaluation of Pisco brandies (Peña y Lilo et al., 1995). The panellists based their in-mouth napping evaluation primarily on retronasal flavours (Louw et al., 2013), which may explain why the addition of the in-mouth modality did not significantly influence the overall product configuration. Arguably, aroma perception is most commonly used, and possibly sufficient, to describe differentiation between spirit products such as brandy and whisky (Peña y Lilo et al., 2005; Jack, 2003). However, mouthfeel perceptions such as smoothness and alcohol burn can be an important driver of liking for consumers and it is important to also gain understanding of these product attributes. It has been suggested that panellists do not use the same perceptual approach to sample evaluation when they are required to treat different aspects of flavour (taste and retronasal odour) holistically, as was done in this study, compared to when they are tasked to evaluate different in-mouth sensations individually (Auvray & Spence, 2008). Although the holistic "in-mouth" experience is representative of the way in which consumers interact with brandy, it is possible that more useful information regarding differentiation on taste and mouthfeel if the partial napping task is restricted to these sensations alone (i.e. retronasal odours are excluded).

4.3.2 Partial napping restricted to taste and mouthfeel perception

Considering the influence of visual cues on taste and mouthfeel perceptions as well as cross-modal interactions between odour, taste, somatosensory and trigeminal perceptions (Bult, 2007; Auvray & Spence, 2008) it was deemed necessary to understand to what extent panellists should be restricted during taste and mouthfeel partial napping exercises of brandy. To this end, panellists were asked to evaluate the taste and mouthfeel properties of ten brandies (Sample Set 2 in Table 1) under three conditions. In the first condition (CG) panellists received the brandy samples in clear glasses and were asked to perform partial napping by first evaluating the aroma of the products and then the taste and mouthfeel, without regard for retronasal odour perception. In the second strategy (DG) they were given the same task, but the brandies were served in black glasses to avoid visual bias. The final condition (NC) involved the panellists evaluating the samples for taste and mouthfeel completely independent of appearance, aroma and retronasal flavours by tasting the samples in black glasses, wearing nose clips, before they have smelled the brandies. Afterwards, they evaluated the samples based on smell.

4.3.2.1 Comparison of the overall product configurations

The overall product configurations obtained with CG, DG and NC were similar (RV 0.849-0.867 between conditions). Figure 4.4 shows the product configurations with its associated descriptors. In all three configurations, C1, C4, C3 and to some extent, C6 are situated opposite C9, C9R, C8 and C7. In general, the former group was associated with fruitcake, spicy aromas and bitterness while the latter group was associated soapy and fresh fruit aromas as well as sweetness and/or viscosity. The positions of C2 and C5 differed between the conditions. In the CG configuration, C5 clearly separates from the other samples whereas it clusters with samples C7-C9 in the DG and NC configurations. Sample C2 is positioned closer to C3 in the DG configuration than in the other configurations.

Based on the distance between the blind duplicate samples, C9 and C9R, relative to the maximum inter-sample distances in the product configurations, the CG strategy delivered the most accurate results (D_r=0.02), followed by the NC (D_r=0.32) and DG (D_r=0.33) conditions which delivered comparably accurate results. In terms of repeatability, the CG condition appeared to be very repeatable (RV=0.873). The RV coefficients between the DG sessions (0.719) and the NC sessions (0.718) were much lower. These values are more comparable to what has been observed from the global napping method where product differentiation is evaluated holistically and for which repeated measurements has been recommended (Louw et al., 2013).

4.3.2.2. Comparison of the taste/mouthfeel configurations

When comparing only the taste and mouthfeel-based configurations, the differences between the three conditions were more apparent (RV 0.740-0.756 between conditions). It appears as if the panellists' taste/mouthfeel assessments were less robust than their aroma assessments (RV 0.813-0.864 between conditions). In the CG configuration (Figure 4.5), the first dimension, F1, was driven primarily by basic taste attributes, with sweetness associated with the negative end of F1 and bitterness, sourness and saltiness associated with the positive end. The second dimension, F2, highlighted differences in mouthfeel, with the positive end of F2 being associated with smoothness and viscosity, opposed to an oily mouthfeel at the negative end. The DG configuration appeared to be driven more by mouthfeel attributes. F1 seemed to represent an oily vs. viscous dimension, with bitterness and sourness being associated with oiliness at the negative end of F1. F2 highlighted differences in alcohol burn, astringency and sweetness. In the NC configurations, the samples did not separate clearly from each other, with the exception of C4. F1 was driven by both taste and mouthfeel attributes; the negative end was associated with bitter, salty, sour tastes and astringency while the positive end was strongly associated with smoothness, viscosity and sweetness. It appears as if the attributes used to describe differentiation on the first two dimensions of the CG configuration was compressed into the first dimension of the NC configuration.

The duplicated samples were positioned much closer to one another in the CG configuration ($D_{r\%}$ =0.14) than in the other configurations. Using nose clips, the panel as a group fared better at positioning the duplicated samples close to each other (D_r=0.41) than without (D_r=0.67). Although the DG condition provided the least accurate taste/mouthfeel configuration in terms of the duplicated samples' positions, it delivered the most repeatable results. The between-rep RV coefficient for DG was 0.777 compared to 0.725 for CG and 0.659 for NC.

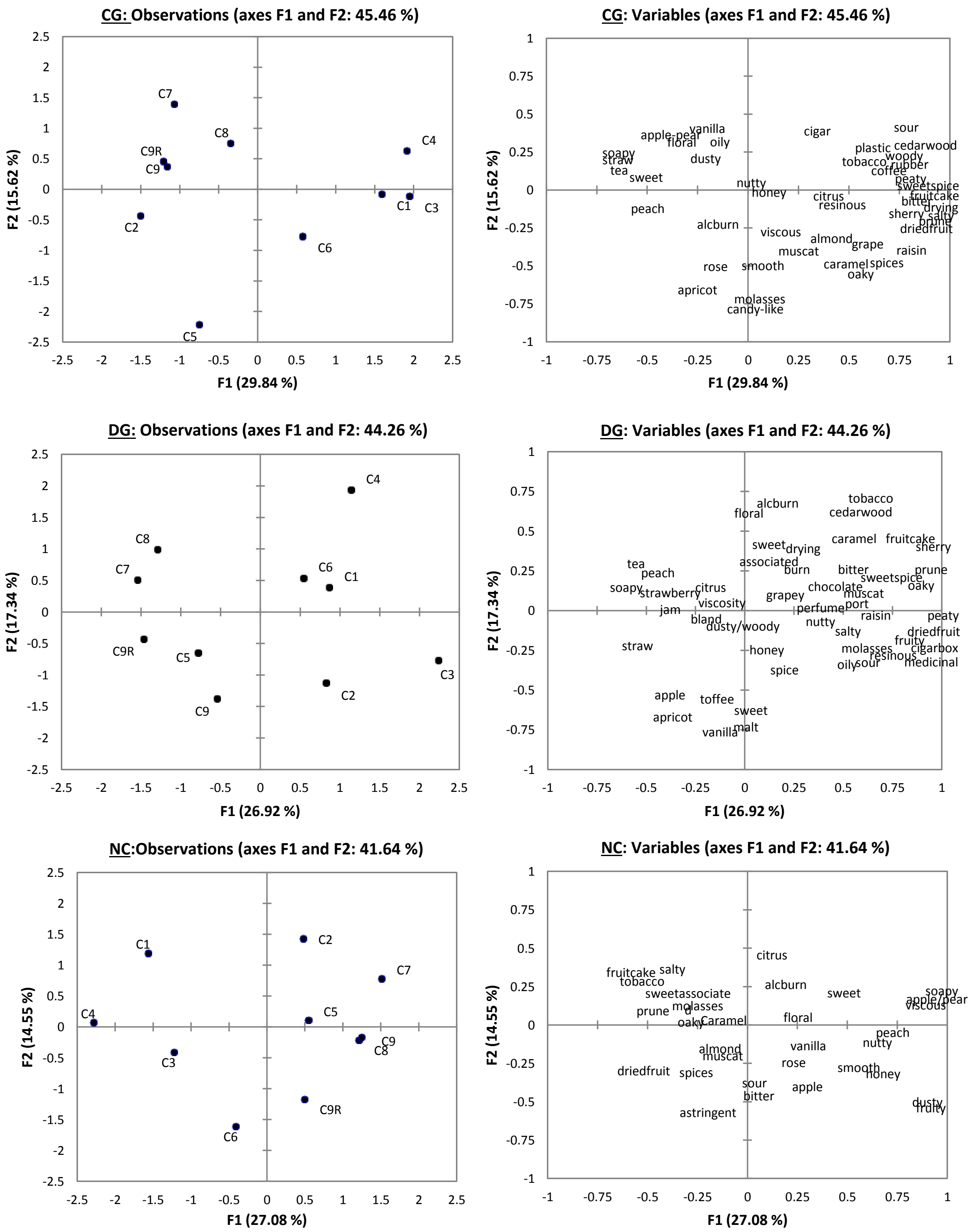

Figure 4.4 Consensus configurations and corresponding variable plots from partial napping of ten brandies served in clear glasses (CG), dark glasses (DG) and in dark glasses with nose clips provided to exclude retronasal odour perception (NC). Samples C9 and C9R are duplicates of the same brand. The partial napping was structured around two evaluations focussing on aroma and taste/mouthfeel respectively.

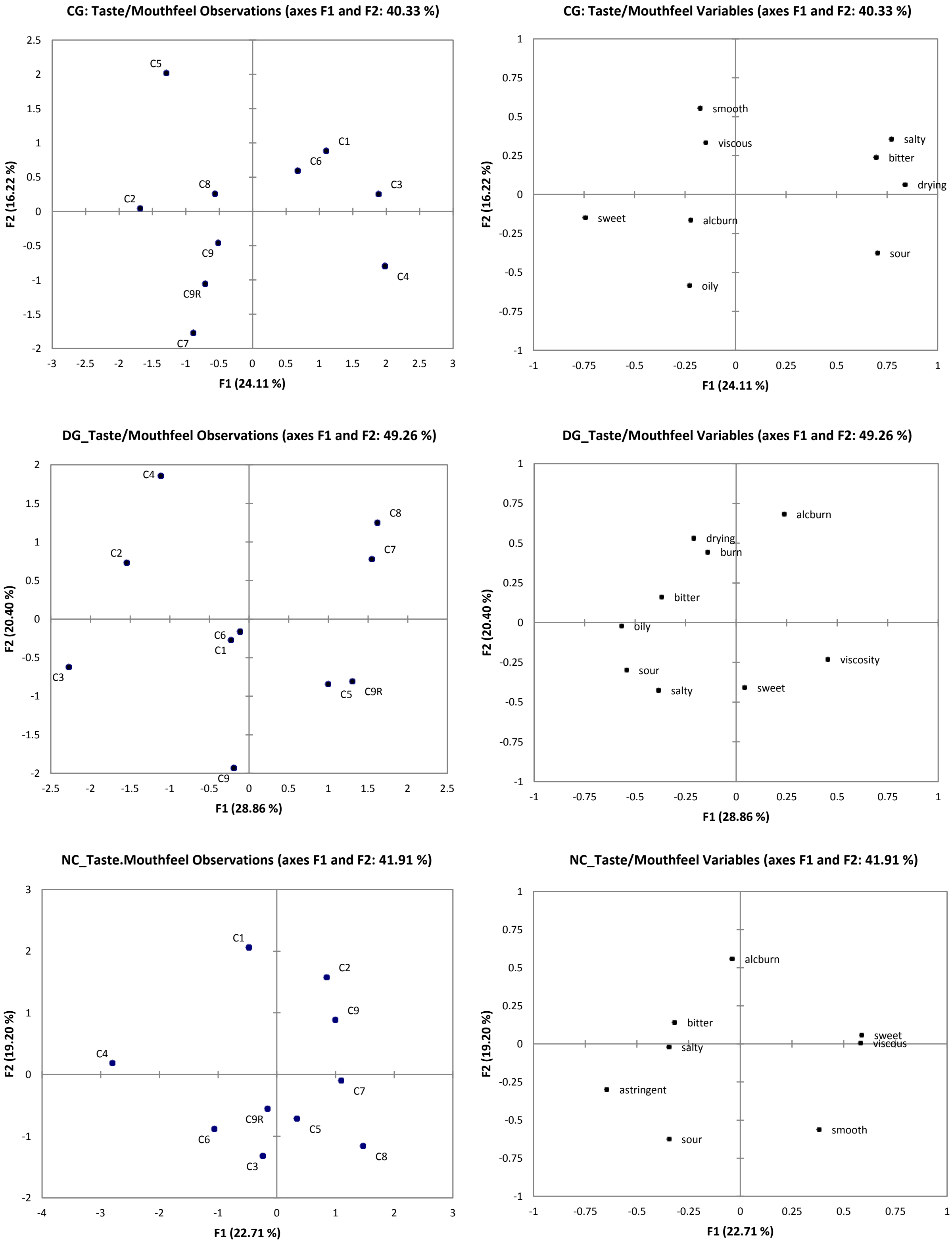

Figure 4.5 Consensus configurations and corresponding variable plots from partial napping (focussing on taste and mouthfeel) of ten brandies that were served in clear glasses (CG), dark glasses (DG) and in dark glasses with nose clips provided to exclude retronasal odour perception (NC). Samples C9 and C9R are duplicates of the same brand.

4.3.2.3 Panellist performance in taste/mouthfeel evaluation

No statistically significant differences in panellist performance were observed between the three conditions (Figure 4.6), although they seemed to perform slightly better in evaluating taste and mouthfeel when these sensations were isolated from appearance, aroma and retronasal odour perceptions (NC condition). It is tempting to compare the RPI and D_r values obtained with these restrictive approaches to taste and mouthfeel evaluation to those obtained with the more integrated "in-mouth perceptions" evaluations from the first section of this study. Although the three restricted approaches tested did generally deliver more consistent (higher RPI) and more accurate (lower D_r) results, it must be considered that besides the fact that the sample set was different, the panel that evaluated Sample Set 1 in the first part of the study were less experienced in partial napping compared to the panel who performed the restricted partial napping evaluations in the second part of the study.

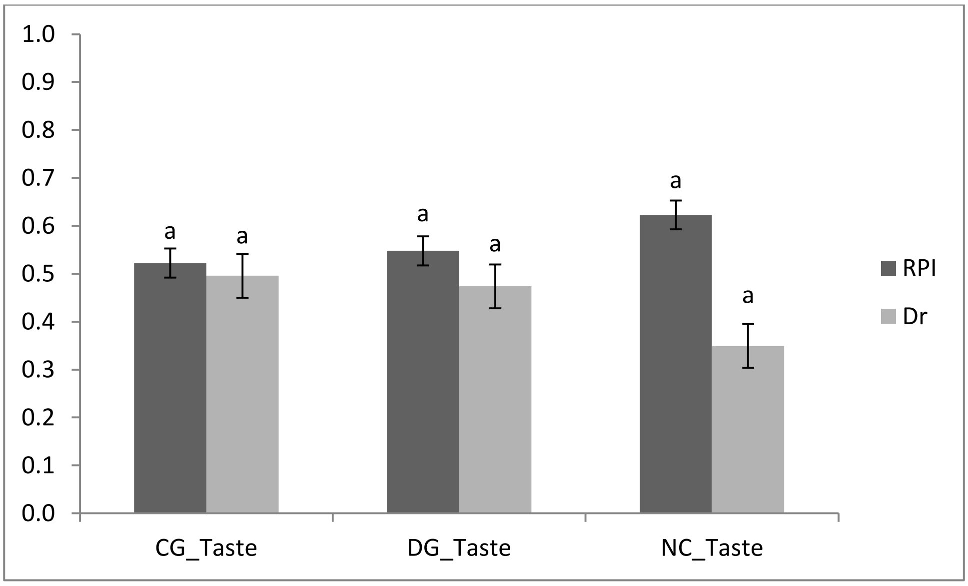

Figure 4.6 Average performance of twelve trained panellists in basic taste and mouthfeel (abbreviated as *taste*) projective mapping of ten brandies served in clear glasses (CG), dark glasses (DG) and in dark glasses with nose clips provided to exclude retronasal odour perception (NC). The relative performance index (RPI) indicates consistency over two replicates with values closer to 1 being better, while D_r indicates the proximity between duplicated samples relative to the rest of the samples as perceived by the panellist, with values closer to 0 being better. Error bars denote standard errors and the letters above each bar denotes Fisher LSD values.

4.4 Conclusion

The results from this study showed that panellists provided the most reliable projective mapping results based on visual perceptions, in comparison to aroma and mouthfeel, and that the perception of colour likely influenced their assessment of in-mouth sensations. This is not surprising as the dominance of the sense of sight relative to the other senses in sensory evaluation of food is well documented (Auvray & Spence, 2008; Lelièvre et al., 2009; Morrot et al., 2001; Parr et al., 2003). It does not appear as if panellists performed significantly better or worse when evaluating samples by aroma or by in-mouth sensations. Based on the results from this study, evaluating the in-mouth perception of brandy as a single construct or modality, does not add sufficient value in a partial napping exercise. This creates an opportunity for using a restrictive "aroma napping" approach for the evaluation of larger numbers of spirit samples to obtain an overview of the most distinctive differentiating features in a product set. However, useful information regarding the taste and mouthfeel properties of brandies can be collected by instructing panellists not to consider retronasal flavours. The results indicated that, as individuals, the panellists performed equally well in their taste and mouthfeel assessments regardless of whether colour and retronasal flavours were perceivable. However, the consensus configuration obtained when the products were served in clear glasses provided accurate and clearly interpretable results compared to the other conditions. Based on the results from this study, the CG approach is recommended, especially since it is also a more natural approach to tasting brandy and more relevant to the consumer's product experience. However, the findings of this study should ideally be validated through the evaluation of other brandy sample sets.

4.5 References:

Abdi, H. (2007). Rv and congruence coefficients. IN: *Encyclopedia of Measurement & Statistics*. (Ed. N.J. Salkind). Thousand Oaks, USA: SAGE Publications Inc., 849-853.

Aubry, V., Etievant, P., Sauvageot, F., & Issanchou, S. (1999). Sensory analysis of Burgundy Pinot noir wines: A comparison of orthonasal and retronasal profiling. *Journal of Sensory Studies, 14,* 97-117.

Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and Cognition*, *17,* 1016-1031.

Bult, J.H.F., de Wijk, R.A., & Hummel, T. (2007). Investigations on multimodal sensory integration: texture, taste and ortho- and retronasal olfactory stimuli in concert. *Neuroscience Letters*, *411,* 6-10.

Caldeira, I., Mateus, A. M., & Belchior, A.P. (2006) Flavour and odour profile modifications during the first five years of Lourinhã brandy maturation on different wooden barrels. *Analitica Chimica Acta*, *563,* 264-273.

Caldeira, I., Anjos, O., Portal, V., Belchior, A., & Canas, S. (2010). Sensory and chemical modifications of wine-brandy aged with chestnut and oak wood fragments in comparison to wooden barrels. *Analytica Chimica Acta*, *660,* 43-52.

Canas, S., Caldeira, I., & Belchior, P. (2009). Comparison of alternative systems for the ageing of wine brandy, wood shape and wood botanical species effect. *Ciência E Técnica Vitivinicola, 24 (2),* 91-99.

Cartier, R., Rytz, A., Lecomte, A., Poblete, F., Krystlik, J., Belin, E., & Martin, N. (2006). Sorting procedure as an alternative to quantitative descriptive analysis to obtain a product sensory map. *Food Quality and Preference, 17,* 562-571.

Dehlholm, C., Brockhoff, P.B., Meinert, L., Aaslung, M.D., & Bredie, W.L.P. (2012). Rapid descriptive sensory methods – comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference, 26,* 267-277.

European Union. (2008). Regulation (EC) No 110/2008 of the European Parliament and of the Council of 15 January 2008. *Official Journal of the European Union,* L 39/30-31.

Giacalone, D., Ribeiro, L.M., & Frøst, M.B. (2013). Consumer-based product profiling: Application of Partial Napping® for sensory characterisation of speciality beers by novices and experts. *Journal of Food Products Marketing, 19:3,* 201-218.

Grygorczyk, A., Lesschaeve, I., Corredig, M., Duizer, L. (2013). Impact of structure modification on texture of a soymilk and cow's milk gel assessed using the napping procedure. *Journal of Texture Studies*, *44,* 236-246.

Jack, F. (2003). Development of guidelines for the preparation and handling of sensory samples in the Scotch whisky industry. *Journal of the Institute of Brewing*, *109,* 114-119.

Lawless, H.T., & Heymann, H. (1998). *Sensory evaluation of food: Principles and practices.* (1st ed). New York, NY: Chapman & Hall.

Le Dien, S., & Pagès, J. (2003). Hierarchical multiple factor analysis: application to the comparison of sensory profiles. *Food Quality and Preference, 6,* 91-98.

Lelièvre, M., Chollet, S., Abdi, H. & Valentin, D. (2009). Beer-trained and untrained assessors rely more on vision than on taste when they categorize beers. *Chemosensory perception, 2,* 143-153.

Louw, L., Malherbe, S., Næs, T., Lambrechts, M., Van Rensburg, P., & Nieuwoudt, H. (2013). Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products. *Food Quality and Preference, 30,* 192-201.

Louw, L., & Lambrechts, M.G. (2012). Grape-based brandies: production, sensory properties and sensory evaluation. IN: *Alcoholic beverages: sensory evaluation and consumer research* (Ed J Piggot). Cambridge: Woodhead Publishing Limited.

Morrot, G., Brochet, F., & Dubourdieu, D. (2001). *Brain and Language*, *79,* 309-320.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire Valley. *Food Quality and Preference, 16,* 642-649.

Parr, W.V., White, K.G., & Heatherbell, D.A. (2003). The nose knows: influence of colour on perception of wine aroma. *Journal of Wine Research, 14,* 79-101.

Peña Y Lilo M., Latrille E., Casaubon, G., Agosin, E., Bordeu E, & Martin, N.(2005). Comparison between odour and aroma profiles of Chilean Pisco spirit. *Food Quality and Preference, 16,* 59-70.

Perrin, L., Symoneaux, R., Maitre, I., Asselin, C., Jourjon, F., & Pages, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: case of wines from Loire Valley. *Food Quality and Preference, 19,* 1-11.

Pfeiffer, J.C. & Gilbert, C.C. (2008). Napping by Modality: a happy medium between analytical and holistic approaches. *Proceedings of Sensometrics Conference 2008, Canada*.

Robert, P., & Escoufier, Y. (1976). A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Applied Statistics, 25,* 257-265.

South African Department of Agriculture (1989). Table 6 (Substances which may be added to liquor products) of the South African Liquor Products Act 60 of 1989, *Government Gazette, 35509*.

Stone, H., & Sidel, J.L. (1992). *Sensory Evaluation Practices.* (2nd Ed.) Orlando, FL: Academic Press.

Toerien, W. (2008). *Fire Water*. Cape Town, Quivertree publication in association with the South African brandy foundation.

Torri, L, Dinella, C., Recchia, A., Næs, T., Tuorila, H., & Monteleone, E. (2013). Projective mapping for interpreting wine aroma differences as perceived by naïve and experienced assessors. *Food Quality and Preference, 29,* 6-15.

Villamor, R.R., Evans, M.A., Mattinson, D.S., & Ross, C.F. (2013). Effects of ethanol, tannin and fructose on the headspace concentration and potential sensory significance of odorants in a model wine. *Food Research International*, *50,* 38-45.

Chapter 5

Research results

Rapid sensory evaluation of brandy: Can sorting be used as a one session screening tool?

Chapter 5: Rapid sensory evaluation of brandy: Can sorting be used as a one session screening tool?

Abstract

The Napping® method has been validated for brandy sensory evaluation, but requires several repeated sessions in order to obtain reliable results. Sorting has been identified as a method that is more stable over repeated sessions and a potentially more appropriate tool for rapid sensory screening of large sample sets of brandy, without the need to use repeated sessions. In this study, ten brandies were evaluated with conventional profiling, Napping® and sorting. The overall product configurations, method repeatability and accuracy were compared and sorting was found to deliver results sufficiently similar to the currently validated methods, if less accurate and descriptive than Napping®. It was further found that representative sample subsets could be selected from the product configuration obtained from a single sorting session, making it more appropriate as a sample screening tool than Napping®.

5.1 Introduction

Brandy is a grape-based distilled beverage with a minimum alcohol content of 36% alcohol by volume (ABV) (European Union, 2008). It is produced in many countries. In South Africa, the two key brandy styles are potstill and blended brandies which differ in distillation method and wood maturation regimes. Potstill brandies are distilled in copper pots; in contrast, blended brandies can contain up to 70% brandy spirit from column distillation. Column distillation produces a more neutral spirit than pot distillation, resulting in a less flavourful beverage (Toerien, 2008). Both potstill and blended brandies are required to undergo wood maturation for at least three years, although potstill brandies are often matured for as much as twenty years.

Brandy flavour can roughly be divided into two categories: those originating from the base wine and concentrated during distillation and those extracted during oak maturation. The base wine flavour components are typically fruity and floral, while the oak-derived volatile components are typically sweet-associated, nutty and spicy (Louw & Lambrechts, 2012). Phenolic compounds extracted during oak maturation can impact on bitterness, mouthfeel and also the perception of volatile flavour compounds (Caldeira et al., 2006; Caldeira et al., 2010; Canas et al., 2009).

The high alcohol content and complex sensory character of brandy make it challenging to evaluate with sensory profiling. Until recently, time consuming and expensive conventional profiling (CP) methods were the only reliable tools available for brandy sensory evaluation, since the effectiveness of rapid sensory profiling methods has not been tested on high alcohol products. In our previous work (Louw et al., 2013), two projective mapping techniques were validated for brandy: Napping®, also referred to as global napping, (GN) (Pagès, 2005) and partial napping (PN) (Pfeiffer & Gilbert, 2008). GN is a holistic tool that provides a snapshot of sensory differentiation in a product set. Panellists are required to place samples on a sheet of paper according to perceived similarity in addition to providing words for each of the samples. Partial napping is a more analytical version of GN where panellists evaluate a specific sensory modality e.g. smell. The aim is to obtain a more focussed account of sample differentiation. Readers can refer to the work of Dehlholm et al. (2012) for a detailed comparison of GN and PN.

Our results showed that PN was more robust and better suited for evaluation of a reasonably large sample set of ten brandies than GN (Louw et al., 2013). However, PN is more time intensive to perform, since it generates more data that have to be processed and interpreted. It was suggested that repeated

GN evaluations can be collected for larger sample sets to ensure stable results; RV coefficients between CP and GN results increased with up to 35% when three GN sessions were captured instead of just one session. Other researchers have also encouraged the use of repeated measurements for projective mapping of red wines (Hopfer & Heymann, 2013; Ross et al., 2012).

Brandy can be considerably more expensive per volume compared to alcoholic beverages with lower alcohol content like wine and beer. The cost-saving impact of a rapid sensory screening method that requires lower sample volumes compared to conventional profiling can be more significant for brandy than for low alcohol products. In the case of brandy, where repeated measurements for GN is essential, it will be useful to have an alternative method that deliver reliable results in one session despite the fatigue associated with high alcohol content and without the more complex data analyses associated with PN.

The sorting task has been advocated as a method for conducting preliminary investigative studies (Deegan et al., 2010) and has been applied to a wide range of products including extensive application to beer (Chollet & Valentin, 2001; Chollet et al., 2011) and wine (Ballester et al., 2005; Campo et al., 2008; Gawel et al., 2001; Johnson et al., 2013; Piombino et al., 2004; Sáenz-Navajas et al., 2012). Sorting has already shown to deliver results similar to CP when applied to cereals and waters (Cartier et al., 2006; Falahee et al., 1997) and has also been reported as stable over repeated measurements when applied to cereals and beer (Cartier et al., 2006; Chollet et al., 2011). Both GN and sorting involve the consideration of overall similarity between samples, but, unlike GN, the sorting task does not require panellists to consider degree of difference or to transfer multidimensional differentiation on a two-dimensional space. Sorting has further been reported to be well suited to larger sample set sizes (Chollet et al., 2011). However, sorting has the same disadvantage as GN: the panellist is presented with all the samples at the same time. This places a limitation on sample set size, especially with high fatigue products such as brandy. Nevertheless, this limitation is unavoidable as most other rapid sensory profiling methods also require simultaneous sample presentation (Valentin et al., 2012; Varela & Ares, 2012).

The aim of this study is to evaluate the sorting task's efficiency as a one session sample screening tool for brandy in order to increase the flexibility and cost-efficiency of the brandy sensory evaluation toolbox. This was done by comparing sorting results to CP and GN in terms of overall product configuration to determine the reliability of the sorting method. Differences in product descriptions will be explored to determine whether the same conclusions regarding the sensory attributes driving product differentiation can be drawn. Accuracy and repeatability of sorting over sessions will be evaluated and compared to GN and CP. Finally, three single session GN and sorting evaluations will be put to the test as screening tools in hypothetical sample selection scenario. Four brandy sensory styles were pre-determined with conventional profiling. Through a cluster analysis based sample selection procedure it was attempted to select samples from each of the sorting and GN configuration that represents each of the four styles. For each method it will also be assessed whether more representative samples can be selected from a single session or from the consensus configuration obtained from multiple sessions. The outcomes of this hypothetical scenario test will be used to illustrate the value of GN and sorting as sample screening tools.

5.2 Materials and methods

5.2.1 Products

A set of ten South African brandies were tested (Table 5.1). The same sample set was also used in our previous work (Louw *et al.*, 2013). Brandy B6 was presented twice as a blind duplicate sample. The sample set included blended and potstill brandies to ensure a wide variety of sensory styles. The samples were diluted to 20% ABV with odourless water one hour prior to evaluation. The samples were served at room temperature in clear standard 250 ml wine tasting glasses.

Table 5.1: List of brandies evaluated.

Sample Number[a]	Style Code	South African Brandy Style	Colour [b]	Key aroma trait [b]
B1	B_LF1	Blended (B)	Light (L)	Fruity (F)
B2	B_LF2	Blended (B)	Light (L)	Fruity (F)
B3	P_LT3	Potstill (P)	Light (L)	Tobacco (T)
B4	B_DW4	Blended (B)	Dark (D)	Woody (W)
B5	B_DW5	Blended (B)	Dark (D)	Woody (W)
B6	B_DW6	Blended (B)	Dark (D)	Woody (W)
B6R	B_DW6R	Blended (B)	Dark (D)	Woody (W)
B7	P_DW7	Potstill (P)	Dark (D)	Woody (W)
B8	P_DW8	Potstill (P)	Dark (D)	Woody (W)
B9	B_LT9	Blended (B)	Light (L)	Tobacco (T)

[a] Sample numbers can be cross referenced to Louw *et al.*, 2013; [b] As determined by conventional descriptive profiling

5.2.2 Sensory methods

Three sensory methods were used, conventional profiling (CP), global napping (GN) and sorting. The results from the three methods were collected independently from each other, in other words, there were no relation between the data collection of the GN and sorting sessions. All tasting sessions took place in white tasting booth with controlled air conditioning and lighting.

5.2.2.1 Conventional profiling

Conventional profiling (CP) was performed according to the guidelines provided in Lawless and Heymann (1998). Specific details regarding the language development and training procedures and data capturing have been described in our previous work (Louw et al. 2013).

5.2.2.2 Napping®

The Napping® task (GN) was performed as proposed by Perrin et al. (2008). The panellists had to evaluate the samples presented and position them on an A3 sheet of paper according to overall sensory similarity and write down words to describe each sample. Data from three sessions were captured which will be referred to as GN 1 to GN 3.

5.2.2.3 Sorting

The panellists were instructed to evaluate the samples and to group them according to sensory similarity. They had to make at least two groups and no more than nine groups. They were asked to provide descriptors for each group of samples. Three sessions were captured which will be referred to as Sort 1 to Sort 3.

5.2.3 Panellists

Ten panellists took part in the GN and sorting evaluations while 12 panellists participated in the CP evaluations. All panellists were recruited from a pool of trained panellists employed by Distell Ltd. The panellists were screened for sensory acuity according to the guidelines of Stone and Sidel (1992) and as described in our previous work (Louw et al., 2013).

5.2.4 Statistical analyses

5.2.4.1 Generating the product configurations

The CP data were evaluated with principal component analysis (PCA) using the panel average. The GN data was captured by measuring the coordinates of each sample relative to the centre of the tasting sheet. The descriptive terms were captured on a citation basis; if a panellist used a descriptor to describe a specific sample, that sample was allocated a value of 1, while the samples for which that descriptor was not used received a value of 0. Where panellists used similar descriptors, the scores for those descriptors were summed. The data were analysed with multiple factor analysis (MFA). Each panellist's coordinates were used as a separate data block. The citation data from all the panellists were aggregated into a single data matrix which was included as passive data i.e. the variables were given very low weights so as not to contribute to the overall MFA configuration. The variables were centred but not scaled.

The descriptors for the sorted groups, or classes, were also captured on a citation basis with the assumption that the descriptors used to describe the group also describe every sample in the group. Attributes with similar meanings were pooled. Attributes cited less than 5% of the possible number of citations or did not have more than one citation for any of the samples were also removed. The sorting class data were captured by assigning a class to each sample for each panellist's sorting. The data were arranged into a matrix with samples as columns and the panellists sorting classes as rows while the aggregated attribute data were similarly arranged but with the attributes as rows (Figure 5.1). MFA was used to combine the sorting classes and the attribute data. The data were centred and scaled. The attribute data were given very low weights so that the final configuration is only based on the sorting class data.

	B_LF1	B_LF2	P_LT3	B_DW4	B_DW5	B_DW6	B_DW6R	B_LT9	P_DW7	P_DW8
Panellist 1	Class 1	Class 2	Class 1	Class 1	Class 2	Class 2	Class 2	Class 1	Class 3	Class 4
Panellist 2	Class 1	Class 1	Class 2	Class 1	Class 3	Class 4	Class 3	Class 1	Class 5	Class 5
.										
.										
.										
Panellist n	Class 1	Class 1	Class 1	Class 2	Class 3	Class 4	Class 1	Class 2	Class 5	Class 6
Descriptor 1	6	6	7	5	7	4	7	7	6	6
.										
.										
.										
Descriptor n	4	2	3	2	4	4	4	2	1	7

Figure 5.1 Condensed illustration of the sorting data matrix with sample names in the first row, followed by classification data and the descriptive data.

Agglomerative hierarchical cluster analysis was applied to each product configuration's scores to identify sample clusters. The Ward method was used to calculate dissimilarity based on Euclidean distances with truncation pre-defined to four classes. The four clusters from each product configuration were used to identify four differentiating samples in a hypothetical sample screening exercise. All the data analyses were performed in XLStat 2010 (Addinsoft, www.xlstat.com).

5.2.4.2 Method performance indicators

Three indicators were used to test the effectiveness of sorting as a screening tool for brandy: reliability, accuracy and repeatability. Reliability is defined as the similarity between the product configuration obtained with sorting and the configuration obtained with CP. The descriptors cited during sorting were compared to the CP attributes. Repeatability is the similarity between the results obtained over three sessions. RV coefficients were used to measure reliability and repeatability (Abdi, 2007; Robert & Escoufier, 1976). An RV coefficient above 0.700 indicates good similarity (Cartier et al., 2006). Accuracy is defined as the panellists' ability to position two duplicated samples close together and sort them in the same group. Accuracy was tested using a distance ratio (D_r) based on the ratio between the Euclidean distance between the duplicated samples and the maximum inter-sample distance (Torri et al., 2013).

5.3 Results and discussion

5.3.1 Overview of product configurations

Table 5.2 shows the RV coefficients between CP, each sorting session and the results from the sorting consensus configuration after MFA. The same is shown for GN vs. CP. All three sorting sessions had a RV coefficient higher than 0.700 compared to CP, whereas Rep 1 for GN was below 0.700, indicating a poorer reliability. The average RV coefficient for the GN method was 0.708 and 0.715 for sorting. When the three sessions were combined using MFA to generate an overall consensus configuration, the RV for sorting was 0.741, a relatively low increase from the average, whereas the RV for GN increased to 0.795, which can be viewed as a considerable increase from the average. It seems that employing repeated measurements does not hold much benefit for sorting, whereas it is almost a necessity for GN.

Table 5.2: A comparison of sensory configurations obtained with conventional profiling (CP, global napping (GN) and sorting, using the RV coefficient

RV relative to CP	Sorting	Global Napping (GN)
Rep 1	0.724	0.658
Rep 2	0.700	0.765
Rep 3	0.720	0.703
Consensus[a]	0.741	0.795
Average	0.715	0.708

[a]Consensus configuration from Rep1-Rep3 resulting from multiple factor analysis

Figures 5.2-5.4 respectively show the CP, GN and sorting product configurations. Overall, the GN (Fig 5.3) and sorting (Fig 5.4) configurations appear to be rotations of the CP configuration. Irrespective of direction, samples B_LT9 and P_LT3 are positioned furthest from B_DW5 and B_DW6/R, while P_DW7 and P_DW8 are positioned on their own. The first two dimensions of the CP configuration (Fig 5.2), F1 and F2, explain 70.18% of the variance in the data, with the majority of the variance being explained by F1. The samples also appear to differentiate more strongly on F1. In contrast, the sorting configuration (Figure 5.4) appears to be roughly a 90° rotation of the CP configuration (Figure 5.2). The rotation can be justified by the very small difference in the amount of variance explained by F1 and F2 of the sorting configuration (5.92%) in addition to the fact that the degree of difference between the groups in the sorting task are not captured. In terms of rotation, the GN configuration (Fig 5.3) appears to be a compromise between the CP (Fig 5.2) and sorting (Fig.5.4) configurations. In essence, GN is a sorting exercise, but with an added dimension of distances between the samples and between groups. This captures more information on the differentiation in the sample set, which is also evident from the higher percentage of variance explained by F1 and F2 (53.67% with 36.70% explained on F1).

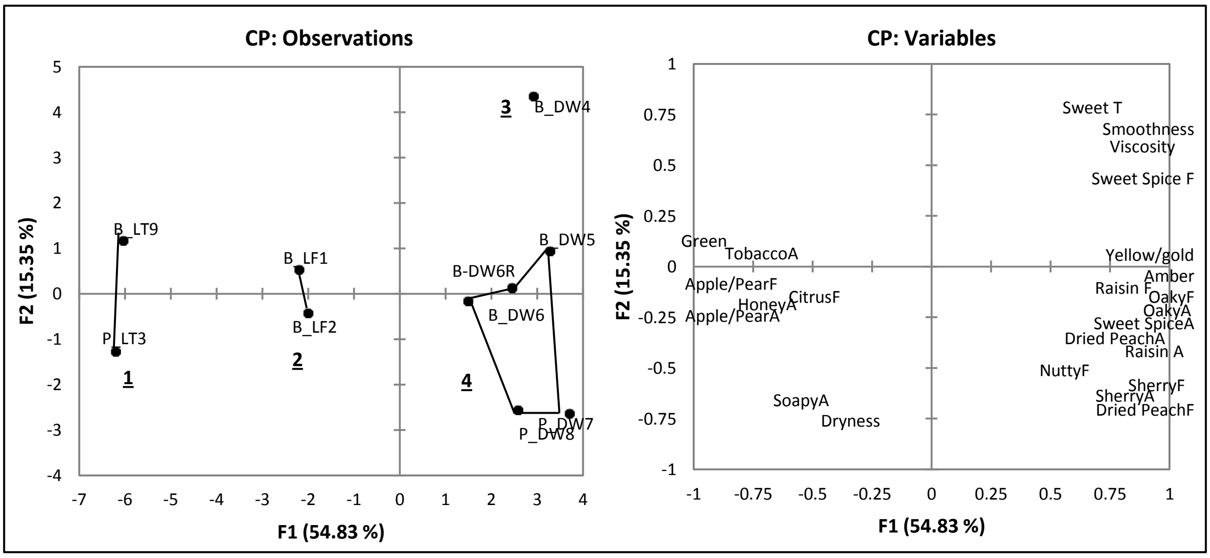

Figure 5.2 PCA configuration obtained from conventional profiling of 10 brandies with observations showed on the left and variables shown on the right. Connected samples represent four clusters (numbered 1-4) as identified with agglomerative hierarchical cluster analysis of the samples scores on F1 and F2**.** Sample B_DW6R is a blind duplicate of sample B_DW6.

The nature of the CP data allows for highlighting differences among samples that do not come across as strongly in GN and sorting. For example, B_DW4 differentiate on the CP configuration, due to its smoothness and viscosity (Figure 5.3) while it does not differentiate on the GN and sorting configurations (Figures 5.3 and 5.4). A similar phenomenon was noted by Pagès in 2005. In his study on the napping of wine, it was also noted that the panellists mainly positioned the wines based on olfaction. Although one of the wines was perceived as significantly sweeter in a profiling exercise using line scales, the panellists did not take this into account during the napping exercise (Pagès, 2005).

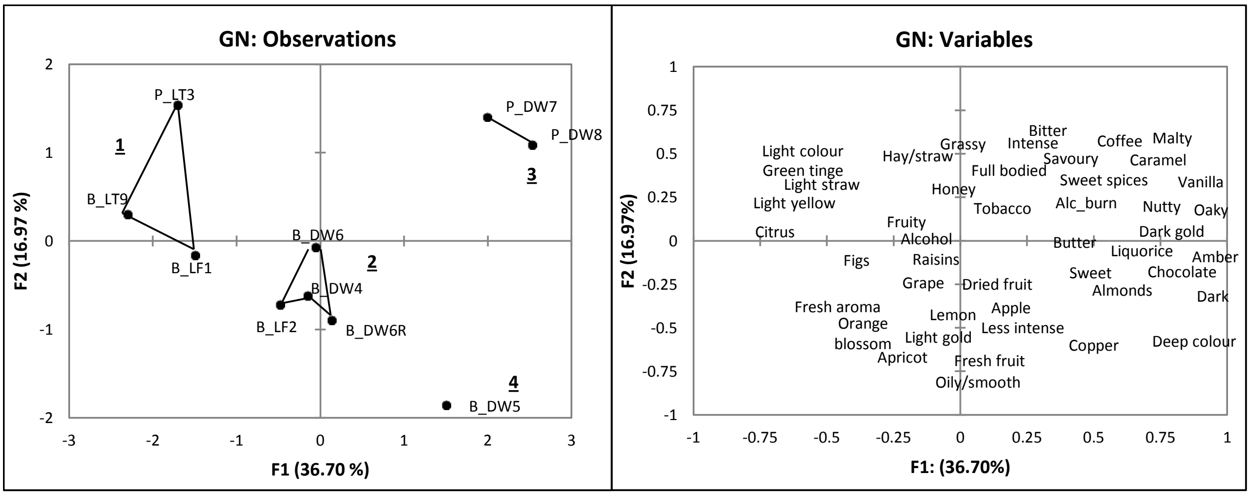

Figure 5.3 Sample configuration obtained from Napping® (GN) of 10 brandies with observations shown on the left and attributes shown on the right. Connected samples represent four clusters (numbered 1-4) as identified with agglomerative hierarchical cluster analysis of the samples scores on F1 and F2. Sample B6_DWR is a blind duplicate of sample B_DW6.

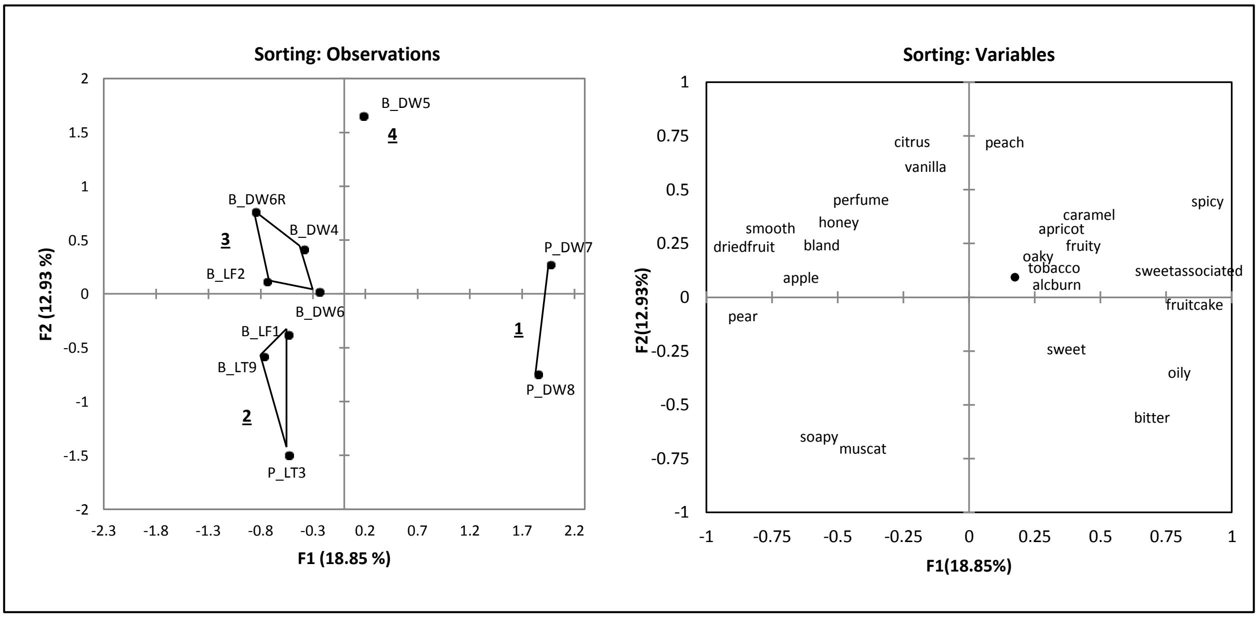

Figure 5.4 Sample configuration obtained from free sorting of 10 brandies with observations showed on the left and attributes shown on the right. Connected samples represent four clusters (numbered 1-4) identified with agglomerative hierarchical cluster analysis of the samples scores on F1 and F2. Sample B6_DWR is a blind duplicate of sample B_DW6.

5.3.2. Comparison of descriptive terms used

As expected, some differences were observed in the descriptive terms used for the three methods. The CP configuration (Figure 5.1) shows B_LT9 and P_LT3 being characterised by more intense green colour, citrus, honey, apple/pear and tobacco notes. This contrasts with B_DW6/R, B_DW5 and also B_DW4 and P_DW7 and P_DW8 that are collectively darker in colour and with dried fruit, sweet spice and oaky notes. P_DW7 and P_DW8 differentiate due to sherry and nutty flavours, while B_DW4 differentiates on mouthfeel as previously mentioned. In the GN configuration (Figure 5.3), P_LT3 and B_LT9 were still characterised by light colour, citrus and hay/straw flavours. They still contrasted with B_DW5 and B_DW6/R which are darker in colour. It appears as if B_DW5 and B_DW6/R were perceived as overall fruitier than B_LT9 and P_LT3, but no distinction was made in terms of fresh or dried fruit. Nevertheless, P_DW7 and P_DW8 were still characterised by spicy, nutty and oaky flavours. In the sorting configuration (Figure 5.4), P_DW7 and P_DW8 were associated with spicy, sweet associated and fruitcake notes, bearing some similarity to the terms associated with these samples in CP and GN. B_DW5 and B_DW6/R were characterised by fresh fruit notes, notably citrus and peach, which is more similar to the GN configuration than to the CP configuration (Figure 5.2). Finally, P_LT3 and B_LT9 were associated with soapy, muscat flavours and not the hay/straw or tobacco characters that were observed in the GN and CP configurations.

Though colour attributes were common in the CP and GN data, the panellists did not cite colour as a sorting criteria and in fact commented after the sorting sessions that their main criterion was aroma. However, the CP and GN results show that B_DW5 is characterised by a deep amber colour which is its main source of differentiation (Figure 5.1 and Figure 5.2). B_DW5 is also a differentiating sample in the sorting configuration (Figure 5.3). This suggests that although the panellists did not cite colour as a sorting criterion, the samples' appearance did influence the way in which they sorted the samples. This relates to the congruence-seeking hypothesis proposed by Lelièvre et al. (2009) where beer tasters were heavily influenced by vision when sorting beer samples. The authors hypothesised that the tasters in their study subconsciously turned to visual cues to confirm their chemosensory perceptions, explaining why the panellists in their study did not feel that colour affected their sorting of the beer samples. Visual perception is driven more by cognitive processing while olfaction and gustation, is governed more by emotions (Bredie & Møller, 2012). Our results support the idea that the subconscious influence of visual perceptions on the panellists sorting exercise may be linked to their need to rationalise their decisions.

Although the panellists admitted that they sorted the samples primarily on aroma, they did use mouthfeel related words to describe their groups. The panellists were trained to consider these characteristics and may have provided information on these perceptions out of habit, even if they did not consider it as an important sorting or napping criterion. In the case of GN where descriptors are provided for each sample, this provides an interesting distinction between which characters are present in the sample and which are important on a holistic level. The similarity between the sorting and GN configurations suggest that the panellists used similar criteria to differentiate between the samples. The nature of the GN task allows panellists to describe each sample, instead of just the group of samples, which gives them the opportunity to add sample specific information that is not captured with sorting. This may be why the role of colour is highlighted better on the GN map.

5.3.3 Repeatability

The three sorting sessions were more similar to each other (RV: 0.863-0.927) than the three GN sessions (RV: 0.611-0.583) (Table 5.3). The high similarity between the three sorting sessions partially explains why combining the three sessions do not improve the similarity of the consensus configuration relative to CP. The same was observed in our previous work on partial napping and brandy, where combining three highly similar partial napping sessions did not result in a consensus configuration more similar to CP than the three individual sessions (Louw et al., 2013).

Table 5.3 The RV coefficients between three repeated sorting evaluations and three repeated Napping® (GN) evaluations.

	Sort 1	Sort 2	Sort 3	GN 1	GN 2	GN 3
Sort 1	1.000					
Sort 2	0.927	1.000				
Sort 3	0.863	0.896	1.000			
GN 1				1.00		
GN 2				0.853	1.00	
GN 3				0.611	0.596	1.00

5.3.4 Positioning of duplicate samples as a measure of accuracy

On average, the accuracy of GN was similar to CP, whereas the panels' accuracy in the sorting task was somewhat lower (Table 5.4). Between-session variability was observed; in Rep 2 of the sorting exercise, the panel's accuracy was very poor (D_r=0.760). Over the three sessions, a maximum of 40% of the panellists were able to group the duplicate samples into the same group. The panellists who sorted the duplicate samples into the same group were not the same over the three sessions, suggesting that it was not a case of some panellists simply being better at the task. In a study on beer only around half of trained panellists were able to sort duplicate samples into the same group (Chollet et al., 2011). The authors suggested that since the beers were served in randomised order, the carry-over effect from stronger beers to the next would have been different for the different panellists, which may have had an effect on their choices in the sorting exercise. This may also explain our results, since the alcohol burn associated with brandy, even when diluted to 20% ABV, also has a strong carry-over effect. A study on reconstituted wines reported that a maximum of 30% of panellists were able to sort duplicate samples into the same group (Sáenz-Navajas et al., 2012). The authors hypothesised that the difficulty that panellists experienced in sorting duplicate samples into the same group was related to the amount of sensory information that panellists were exposed to. Nevertheless, with the exception of Rep 2, the PPI scores for the sorting task, although higher than CP and GN, were not particularly poor.

Table 5.4 Accuracy in positioning duplicate samples close together for conventional profiling (CP), sorting and Napping® (GN) for three repeated evaluations, a combination of three sessions and average over three sessions.

Distance ratio (D_r)[a]	CP	Sorting	Global Napping (GN)
Rep 1	0.201	0.299 (33%)[b]	0.115
Rep 2	0.282	0.760 (20%)	0.171
Rep 3	0.175	0.218 (40%)	0.576
Consensus	0.093	0.280	0.174
Average	0.219	0.426	0.287

[a] Ratio between the Euclidean distance between two duplicated samples and the maximum inter-sample distance on a product configuration; [b]The percentage of panellists who put the duplicate samples in the same group

5.3.5 Effectiveness in a hypothetical sample screening exercise

As the purpose of this study is to investigate sorting as a one session sample screening tool, let's assume a hypothetical scenario where the aim is to select four samples that are perceived as very different out of a set of ten brandies for further research on consumer product acceptance. As the GN and sorting results from this study mainly highlighted olfactory differences, the sample selection will be focussed on aroma. A good strategy is to select samples from different parts of a sensory product configuration to cover as much of the differentiation in the configuration as possible, either visually or with the aid of cluster analysis (Næs, Brockhoff & Tomic, 2010). We have applied agglomerative hierarchical cluster analysis to identify four clusters for each product map. In this case, a group containing only one sample is also considered as a cluster, as this sample may represent a unique style. The final sample selection should also take other product information into account; we know that B_DW6 and B_DW6R, are replicates and that the sample set also contains two styles of brandies, namely blended and potstill (Table 5.1).

Four brandy styles, based on type, colour and main aroma attributes, could be identified from the CP results (Figure 5.1). These are *light with tobacco aromas* (LT), *light and fruity* (LF), *dark and woody* (DW) blended brandies and *dark and woody* potstill brandies. These are summarised in Table 5.5. Using the sample selection procedure described above, four differentiating brandies will be selected from the product configurations of each of the sorting and GN sessions by selecting one sample from each cluster. The selections from each product configuration will be compared to the styles identified with CP to validate the representativeness of each sample screening. Figure 5.5 shows the product configurations from each of the GN and sorting replications with four clusters indicated on each plot.

Table 5.5 Comparison of sample subsets chosen from sorting and global napping (GN) product configurations to brandy type and style categories identified with conventional sensory profiling (CP). Sort 1-3 refers to individual sessions while Sort all refers to consensus configuration derived from Sort 1-3 through MFA. The same annotation is used for GN.

CP Cluster	Brandy style	Sensory style	Sample	Sort all	Sort 1	Sort 2	Sort 3	GN all	GN 1	GN 2	GN 3
1	Potstill Blended	Light/Tobacco (LT)	P_LT3 B_LT9	P_LT3	P_LT3	NONE	B_LT9	B_LT9	P_LT3	P_LT3 B_LT9	B_LF1
4	Potstill	Dark/Woody (DW)	P_DW7 P_DW8	P_DW8	P_DW7	P_DW7 P_DW8	P_DW7	P_DW7	P_DW8	P_DW8	NONE
2	Blended	Light/Fruity (LF)	B_LF1 B_LF2	NONE	B_LF2	B_LF2	B_LF2	B_LF2	NONE	NONE	B_LF1
3,4	Blended	Dark/Woody (DW)	B_DW4 B_DW5 B_DW6 B_DW6R	B_DW5 B_DW6R	B_DW5	B_DW6	B_DW4	B_DW5	B_DW4 B_DW6	B_DW5	B_DW4 B_DW5

In Sort 1 (Figure 5.5, top right), samples P_DW7, P_LT3 and B_DW5 were selected from their respective clusters based on their extreme positions on F1 and F2. B_LF2 was selected based on its position at the centre of the plot. This subset represents all four styles summarised in Table 5.5.

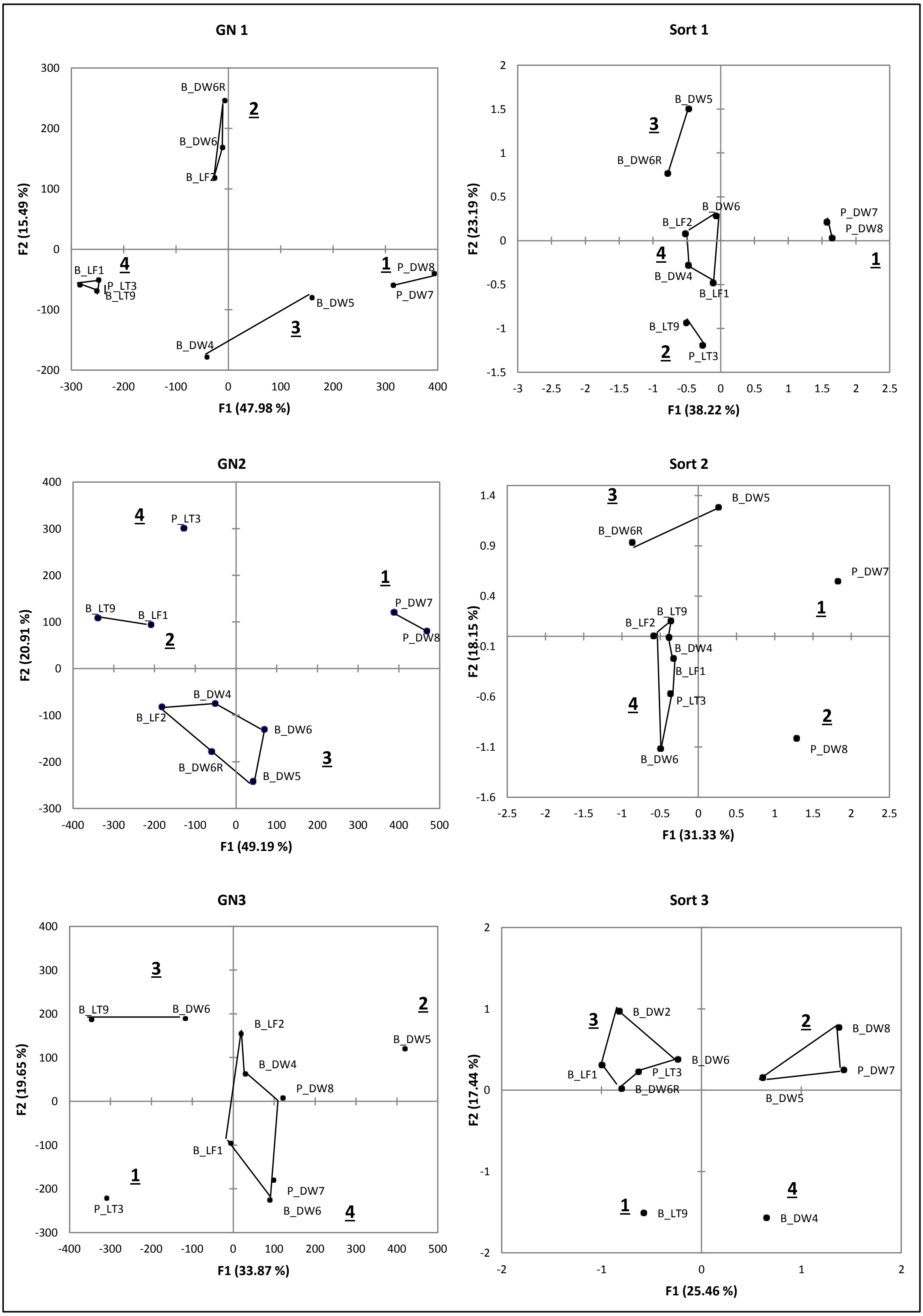

Figure 5.5 Sample configurations obtained from three Napping® (GN 1 –GN 3) evaluations and three sorting evaluations (Sort 1 – Sort 3) of ten brandies. Connected samples represent four clusters (numbered 1-4) identified with agglomerative hierarchical cluster analysis of the samples scores on F1 and F2. Sample B_DW6R is a duplicate of B_DW6.

In Sort 2 (Figure 5.5, middle right), the opposing positions of B_DW6 and B_DW6R flag a warning sign that the configuration is not accurate. Ideally one would repeat the evaluation to create a more accurate map. This would avoid investing in expensive consumer research with a sample set based on an inaccurate sample selection procedure. However, if forced to work with these results, one could refer to differentiation revealed in the third dimension (F3), which adds a further 12.04% to the 49.58% variance already explained by the first two dimensions, F1 (31.33 %) and F2 (18.15 %), respectively (Figure 5.6, left). The sample configuration based on F1 and F3 is shown in Figure 5.6 (left). From this plot, B_DW6 was selected as it, and its duplicate B_DW6R were the only samples present lower left quadrant. The three remaining samples were selected from Cluster 1 and Cluster 2 in order to cover a larger area of the product space. Cluster 1 is the most heterogeneous cluster and the inter-sample distance on the cluster analysis dendogram is very close to the truncation threshold (Figure 5.6, right). Therefor the two samples in Cluster 1 (P_DW7 and P_DW8) were selected. B_LF2 was selected as a sample from Cluster 2 as it had the highest negative score on F1. This selection contains two *dark/woody* potstill brandies, one *dark/woody* blended brandy and one *light/fruity* blended brandy, but no *light/tobacco* brandies.

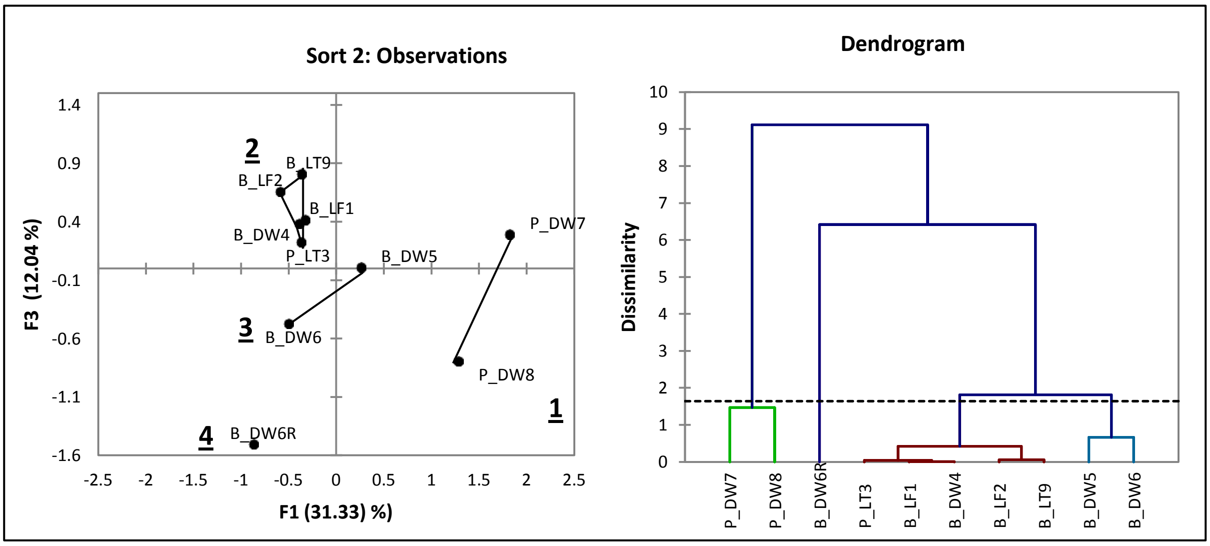

Figure 5.6 Left Sample configuration obtained from a single sorting session of 10 brandies. Connected samples represent four clusters (numbered 1-4) identified with agglomerative hierarchical cluster analysis of the samples scores on F1 and F3. Sample B_DW6R is a blind duplicate of Sample B_DW6. **Right:** A dendogram showing distances between the clusters.

In Sort 3 (Figure 5.5, bottom right), the four clusters also represent the four quadrants of the product configuration and one sample was selected from each: B_LF2, P_DW7, B_DW4 and B_LT9. This selection coincidently represents each of the styles identified with CP in Table 5.5.

We can also consider the sorting consensus configuration, with the data from all three sessions combined; Sort All (Figure 5.4). B_DW5, P_LT3 and P_DW8 represent three clusters as well as the extreme ends of F1 and F2. As for the final sample, B_DW6 and B_DW6R are positioned on opposing ends of the remaining cluster, suggesting that the cluster is fairly homogeneous and that B_DW6 may be a representative choice. These samples represent the *light/tobacco, dark/woody* potstill and *dark/woody* blended brandies styles but not the *light/fruity* style. More representative sample sets were selected from the single sorting sessions than from the sorting consensus configuration based on three repeated session.

The same approach was followed for the GN configurations. In GN 1 (Figure 5.5, top left) P_DW7, B_DW4 and B_DW6 were selected based on their high scores on F1 and F2. As the selection now contains two blended brandies and one potstill brandy, potstill brandy P_LT3 could be the most interesting sample from Cluster 4. The final selection is summarised in Table 5.5.

From GN 2 (Figure 5.5, middle left), P_LT3 would be an obvious choice as it is positioned in its own cluster. Either P_DW7 or P_DW8 could be selected from Cluster 3. B_DW5 represents the extreme negative end of F2 which makes it a good choice from Cluster 4. B_LT9 has a higher negative score on F1 and may be a good selection. With this selection, three of the four styles identified with CP are covered.

In GN 3 (Figure 5.5, bottom left), the opposing positions of B_DW6 and B_DW6R again raise concerns about the accuracy of the configuration. As mentioned during the discussion of Sort 2, it would be best to repeat the exercise to obtain more reliable results. For the purpose of discussion, the same approach will be followed as for Sort 2. As the largest difference between B_DW6 and B_DW6R is explained by F2, the configuration from F1 and F3 is compared. In this configuration (Figure 5.7), B_DW6 and B_DW6R still clusters in separate clusters, but they are closer on the product map. B_LF1, B_LT9, B_DW5 and B_DW4 would seem like the best choices from this configuration. This subset contains one light/fresh blended brandy, one light/tobacco potstill brandy and two dark/oaky blended brandies.

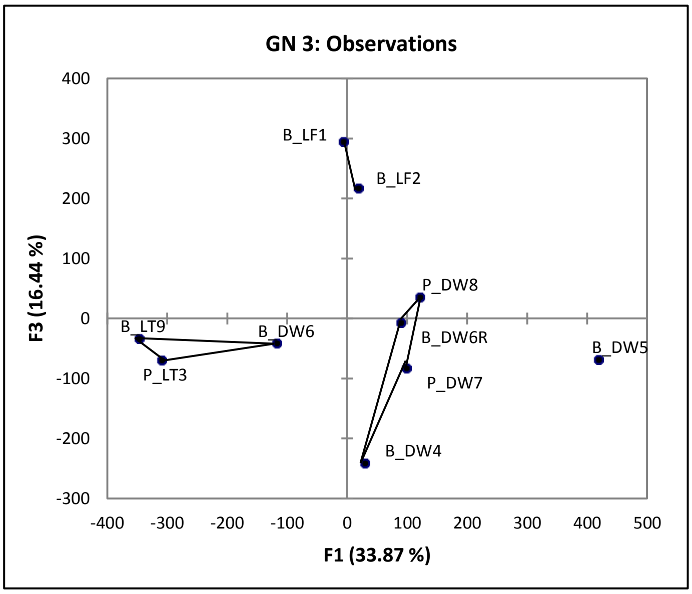

Figure 5.7 Sample configuration obtained from projective mapping of 10 brandies showing samples' scores on the first and third dimensions. Connected samples represent four clusters (numbered 1-4) identified with agglomerative hierarchical cluster analysis of the samples scores on F1 and F3. Sample B6R is a blind duplicate of Sample B6.

Finally, the GN consensus configuration (GN All) (Figure 5.3) B_DW5, P_DW7 and B_LT9 were chosen based on their high scores on F1 and F2. The remaining sample is to be selected from Cluster 2. The choice is not obvious, but B_LF2 could be selected since it is furthest from the centre of the plot. With this selection, each of the four CP styles is selected, although it must be noted that the light/fresh blended brandy group could easily have been replaced by another dark/oaky blended brandy from Cluster 2.

In this study, none of the individual GN replications delivered a subset of samples representative of the differentiating styles identified by the more analytical and detailed CP method. However, when the data from the three GN replicates were combined, a representative sample set could be selected. In contrast, two of the individual sorting replicates delivered representative subsets. The third was based on an inaccurate dataset, as highlighted by the duplicate samples, and it is not surprising that the sample selection was not representative. The subset selected from the consensus sorting configuration was actually less representative than the selections based on the individual sessions. These observations support the idea that one sorting session is sufficient as a screening method, while it is best to perform repeated measurements with GN.

GN 3 and Sort 2 illustrate the benefit of including a blind control sample as an alert for inaccurate results. If it were unknown that B_DW6 and B_DW6R were the same sample, they could both easily have been included in the subset, which would result in testing two samples with consumers that are virtually the same.

5.4 Conclusion

The aim of this study was to establish sorting as a one-session screening tool for brandy sample selection. The results indicated that the individual sorting sessions tended to be more reliable than the individual GN sessions. Incorporating repeated measurements increased the reliability of GN, but not sorting. In this case, sorting was more repeatable than global napping, but less accurate. GN provided more informative descriptive information, as the attributes are captured by sample instead of by group. This makes it a more appropriate tool for studies where a sensory snapshot is required without any further actions than sorting. As hoped, the results showed that a single sorting session was sufficient to identify subsets of samples representative of what was found with CP results. Including a blind duplicate sample is highly recommended as a quality control point for both sorting and GN, especially for complex products such as brandy. Although the results are based on only one sample set, sorting promises to be an excellent screening tool for brandy for studies where further steps are to be taken on a subset of samples.

As the panellists tended to focus mainly on olfaction, one may consider sorting brandy products only on aroma. Previous results showed a high correlation between orthonasal and retronasal odour quantification of Pisco brandy; aroma may well be used as a proxy for retronasal flavours (Peña Y Lilo et al., 2005). Although information on taste and mouthfeel would not be captured in an aroma sort, it seems that it is also unlikely to be captured nor accurately described with holistic sorting. Between 9 and 20 beers are recommended for sorting tasks as sample sets that are too small or too large sample set sizes affects task efficiency (Chollet *l* 2011). The maximum number of brandies should be determined in future studies as brandies are much more volatile than beer and 20 samples may be too many.

5.5 References

Abdi, H.(2007). RV and congruence coefficients. IN: Encyclopedia of Measurement & Statistics. (Ed (N.J. Salkind). SAGE publications Inc., 849-853.

Ballester, J., Dacremont, C., Le Fur, Y.,& Etievant, P. (2005). The role of olfaction in the elaboration and the use of the Chardonnay wine concept. *Food Quality and Preference, 16,* 351-359.

Bredie, W.L.P., & Møller, P. (2012). Overview of sensory perception. IN: *Alcoholic beverages: sensory evaluation and consumer research* (Ed J Piggot). Cambridge: Woodhead Publishing Limited. p 18-19.

Caldeira, I., Mateus, A.M., & Belchior, A.P. (2006). Flavour and odour profile modifications during the first five years of Lourinhã brandy maturation on different wooden barrels. *Analytica Chimica Acta, 563*, 264-273.

Caldeira, I., Anjos, O., Portal, V., Belchior, A., & Canas, S. (2010). Sensory and chemical modifications of wine brandy aged with chestnut and oak wood fragments in comparison to wooden barrels. *Analytica Chimica Acta, 660,* 43-52.

Campo, E., Do, B.V., Ferreira, V., & Valentin, D. (2008). Aroma properties of young Spanish monovarietal white wines: a study using sorting task, list of terms and frequency of citation. *Australian Journal of Grape and Wine Research, 14,* 104-115.

Canas, S., Caldeira, I., & Belchior, P. (2009). Comparison of alternative systems for the ageing of wine brandy. Wood shape and wood botanical species effect. *Ciênca E Técnica Vitivinicola 24 (2),* 91-99.

Cartier, R., Rytz, A., Lecomte, A., Poblete, F., Krystlik, J., Belin, E., & Martin, N. (2006). Sorting procedure as an alternative to quantitative descriptive analysis to obtain a product sensory map. *Food Quality and Preference, 17,* 562-571.

Chollet, S., & Valentin, D. (2001). Impact of training on beer flavour perception and description: Are trained and untrained subjects really different. *Journal of Sensory Studies*, *16,* 601-618.

Chollet, S., Lelièvre, M., Abdi, H., & Valentin, D. (2011). Sort and beer: Everything you wanted to know about the sorting task but did not dare to ask. *Food Quality and Preference*, *22*, 507-520.

Deegan, K.C., Koivisto, L., Näkkilä, J., Hyvönen, L., & Tuorila, H. 2010. Application of a sorting procedure to greenhouse-grown cucumbers and tomatoes. *LWT Food Science and Technology*, *43*, 393-400.

Dehlholm, C., Brockhoff, P.B., Meinert, L., Aaslung, M.D., & Bredie, W.L.P. (2012). Rapid descriptive sensory methods – comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference, 26,* 267-277.

European Union (2008). Regulation (EC) No 110/2008 of the European Parliament and of the Council of 15 January 2008. *Official Journal of the European Union*, L 39/30-21.

Falahee, M., & MacRae, A.W. (1997) Perceptual variation among drinking waters: The reliability of sorting and ranking data for multidimensional scaling. *Food Quality and Preference, 8,* 389.

Gawel, R., Iland, P.G., & Francis, I.L. (2001). Characterizing the astringency of red wine: A case study. *Food Quality and Preference*, *12,* 83-94.

Hopfer, H., & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual measurements. *Food Quality and Preference*, *28,* 164-184.

Johnson, T.E., Hasted, A., Ristic, R., & Bastian, S.E.P. (2013). Multidimensional scaling (MDS), cluster analysis and descriptive analyses provide preliminary insights into Austrialian Shiraz wine regional characteristics. *Food Quality and Preference, 29,* 174-185.

Lawless, H.T., & Heymann, H. (1998). *Sensory evaluation of food: Principles and practices.* (1st ed). New York, NY: Chapman & Hall.

Lelièvre, M., Chollet, S., Abdi, H. & Valentin, D. (2009). Beer-trained and untrained assessors rely more on vision than on taste when they categorize beers. *Chemosensory perception, 2,* 143-153.

Louw, L., & Lambrechts, M.G. (2012). Grape-based brandies: production, sensory properties and sensory evaluation. IN: *Alcoholic beverages: sensory evaluation and consumer research* (Ed J Piggot). Cambridge: Woodhead Publishing Limited, p 281-197.

Louw, L., Malherbe, S., Næs, T., Lambrechts, M., Van Rensburg, P., & Nieuwoudt, H. (2013). Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products. *Food Quality and Preference, 30,* 192-201.

Næs, T., Brockhoff, P.B., & Tomic, O. (2010). *Statistics for sensory and consumer science*. United Kingdom: John Wiley & Sons Ltd: p 146-147.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire Valley. *Food Quality and Preference, 16,* 642-649.

Peña Y Lilo, M., Latrille, E., Casaubon, G., Agosin, E., & Bordeu E. (2005). Comparison between odour and aroma profiles of Chilean Pisco spirit. *Food Quality and Preference,* 16, 59-70.

Perrin, L., Symoneaux, R., Maitre, I., Asselin, C., Jourjon, F., & Pages, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: case of wines from Loire Valley. *Food Quality and Preference, 19,* 1-11.

Pfeiffer, J.C., & Gilbert, C.C. (2008). Napping by Modality: a happy medium between analytical and holistic approaches. *Proceedings of Sensometrics Conference 2008,* Canada.

Piombino, P., Nicklaus, S., Le Fur, Y., Moio, L., & Le Quere, J. (2004). Selection of products presenting given flavor characteristics: An application to wine. *American Journal of Enology and Viticulture, 55:1,* 27-34.

Robert, P. & Escoufier, Y.(1976). A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Applied Statistics, 25,* 257-265.

Ross, C.F., Weller, K.M., & Alldredge, J.R. (2012). Impact of serving temperature on sensory properties of red wine as evaluated using projective mapping by a trained panel. *Journal of Sensory Studies, 27,* 463-470.

Sáenz-Navajas, M., Campo, E., Avizcuri, J.M., Valentin, D., Fernández-Zurbano, P., & Ferreira, V. (2012). Contribution of non-volatile and aroma fractions to in-mouth sensory properties of red wines: Wine reconstitution strategies and sensory sorting task. *Analytica Chimica Acta, 732,* 64-72.

Stone, H., & Sidel, J.L. 1992. *Sensory Evaluation Practices.* (2nd ed). Orlando, FL: Academic Press.

Toerien, W. (2008). *Fire Water.* Cape town: Quivertree publication in association with the South African Brandy Foundation.

Torri, L, Dinella, C., Recchia, A., Næs, T., Tuorila, H., & Monteleone, E. (2013). Projective mapping for interpreting wine aroma differences as perceived by naïve and experienced assessors. *Food Quality and Preference, 29,* 6-15.

Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012). Quick and dirty but still pretty good: a review of new descriptive methods in food science. *International Journal of Food Science and Technology, 47(8),* 1563-1578.

Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International, 48,* 893-908.

Chapter 6

Research results

The effect of tasting sheet shape on product configurations and panellists' performance in sensory projective mapping tasks of brandy products

Chapter 6: The effect of tasting sheet shape on product configurations and panellists' performance in sensory projective mapping tasks of brandy products

Abstract

Projective mapping is a rapid sensory profiling method used to obtain overviews of the sensory differentiation in product sets. Elongated projective mapping tasting sheets, i.e. rectangles, have been hypothesised to bring forth more prominent sample differences, while shapes with equal perpendicular bisectors, such as circles or squares, could reportedly be used to visualise more subtle sample differences. This hypothesis was further explored in the present study in order to gain a better understanding of the practical implications of using different tasting sheet shapes to obtain different project goals. The results showed that very similar product configurations could be obtained with square, rectangular and round tasting sheets. Panellists performed better on round tasting sheets, leading to more accurate results. Square tasting sheet delivered the most different results. The practical significance of using different tasting sheet shapes to elicit either broader or more subtle sample differences could not be established, although there were some indications that this proposed effect may be sample set dependent.

6.1 Introduction

Projective mapping, and its well-known sub-form Napping®, are rapid profiling methodologies designed to obtain a broad overview of the sensory differentiation in a product set as perceived by either trained (Risvik et al., 1994), expert (Pagès, 2005) or naïve assessors (Risvik et al., 1997). As with other rapid sensory profiling methods, such as sorting (Lawless et al., 1995), the aim is to reduce the time- and cost impact of conventional descriptive profiling by excluding the panel consensus training phase (Cartier et al., 2006; Dehlholm et al., 2012.). Although first used as a market research tool, its application to understand perceptual differences among food products was first proposed by Risvik et al. in 1994. Projective mapping requires panellists to place samples on a sheet of paper, based on perceived sensory similarity between the samples being evaluated. Samples that are perceived as very similar are placed close together, while samples that are perceived as very different, are placed far apart. More recent variations of the projective mapping task also include a descriptive step, known as ultra-flash profiling, where panellists provide terms to describe each of the samples (Perrin et al., 2008).

The validity of projective mapping relative to conventional descriptive profiling is well-documented and has been tested on a variety of product types across food categories: chocolates (Kennedy & Heymann, 2009; Risvik et al., 1994), dried soups (Risvik et al., 1997), dairy products (Barcenas et al., 2004, Drake et al., 2009, Nestrud & Lawless, 2010), hot and cold non-alcoholic beverages (Moussaoui & Varela, 2010; Nestrud & Lawless, 2008; Veinand et al., 2011), wine (Pagès, 2005; Perrin et al., 2008), and brandy (Louw et al., 2013), among others. In order for projective mapping to become a work-horse methodology for projects with considerable cost or time constraints, it is necessary to understand the factors that affect the final consensus configuration and the quality of the results. Past research has already clarified some practical aspects relating to the sensory testing phase. In terms of panel expertise, untrained, trained and product experts are able to perform the task and deliver interpretable and actionable results, although trained and expert panels tend to deliver more reliable results than consumers (Barcenas et al., 2004; Pagès, 2005; Perrin et al., 2008; Veinand et al., 2011). Projective mapping studies with trained or expert panels typically involve panels consisting of 8-22 tasters

(Barcenas et al., 2004; Hopfer & Heymann, 2013) which are similar to the panel size for conventional profiling studies (Heymann et al., 2012). Larger numbers of consumers can be used if the study forms part of a more comprehensive consumer research study (Torri et al., 2013). The question of whether repeated measurements should be used is often asked in sensory science and some recent publications have strongly recommended the use of repeated measurements for projective mapping (Hopfer & Heymann, 2013; Louw et al., 2013).

Despite these clarifications, there are still some questions remaining that pertain to the optimal way to execute the methodology in the context of a specific research question or practical application. One of the factors which have received very little attention to date is the shape of the tasting sheet and the effect this could have on the projective strategies of panellists. Square (Kennedy & Heymann, 2009; Nestrud & Lawless, 2010) and rectangular tasting sheets (Barcenas et al., 2004; Louw et al., 2013; Pagès, 2005; Perrin et al., 2008) of various dimensions, have been used in projective mapping studies; presumably dictated by practical constraints such as availability and booth size. However, these studies did not focus on investigating the influence of paper shape on the outcomes of the results. In two recent projective mapping studies, the effect of square vs. rectangular (Hofper & Heymann, 2013) and rectangular vs. circular (Dehlholm, 2013) tasting sheets, on panellists' use of the X- and Y-coordinates and the consensus product configuration have been investigated.

The elongated axis of the rectangular tasting sheet has in both cases been implicated to provide a dominant direction of differentiation, which panellists use to project the main product differences on. In contrast, shapes of equal length and width were hypothesised to highlight more subtle sample differences. The first study reported that the overall consensus configuration obtained from square and rectangular sheets differed significantly (Hopfer & Heymann, 2013), while the second study found that although the panellists treated the round and rectangular tasting sheets differently, the overall consensus configurations were very similar (Dehlholm, 2013). Although panellists are clearly influenced by the shape of the tasting sheet, it is still uncertain whether any effect of tasting sheet shape on overall results is substantial enough to make a practical impact.

On an individual panellist level, both studies mainly focussed on the panellists' use of the horizontal and vertical axes and not on the panellists' internal consistency in performing the task (Dehlholm, 2013; Hopfer & Heymann, 2013). Dehlholm suggested that the horizontal-vertical aspect ratios of rectangular sheets are more similar to humans' visual field that that of a round tasting sheet and that a rectangular tasting sheet can be a more natural projective space than a round one. Although this implies that panellists may find it easier to use a rectangular tasting sheet than a round one, differences in individual panellists' performance between the two tasting sheet geometries were not investigated.

This study aims at further exploring the potential use of using tasting sheet shape to extract either gross or subtle perceived sample differences. If tasting sheet shape is found to affect projective mapping results in such a way that different conclusions can be drawn, then adjusting the tasting sheet shape could potentially be used a tool to create a better fit between sensory profiling method and project goal. A second aim is to determine the effect of sheet shape on panellist performance in the projective mapping task. Improved panellist performance translates to more reliable results; adjusting the tasting sheet shape may possibly be an easy and practical way to get the best out of the projective mapping method. Three sheet shapes are compared: rectangular, square and round. The study is conducted on South African brandies, a grape-based distilled beverage.

6.2. Materials and Methods

6.2.1 Products

Eight South African brandy samples were evaluated in this study (Table 6.1). The samples represent two South African brandy styles, namely blended brandies and potstill brandies. According to South African legislation, blended brandies must contain at least 43% ABV (alcohol by volume) and a minimum of 30% of the brandy must be distilled in a copper potstill while the rest can be produced with column distillation (South African Department of Agriculture, 1989). Potstill brandies contain at least 38% ABV and must be 100% copper pot distilled, resulting in more flavourful brandies compared to blended brandies. Blended brandies encompass the majority of the lower priced brandy market in South Africa while potstill brandies are marketed as super premium products. Samples OR1 and OR2 represent blind replicates of the same product. The samples were stored at room temperature in their original packaging. One hour prior to tasting, the brandies were diluted to 20% ABV with odourless distilled water. The brandies were served at room temperature in standard 250 ml clear tulip-shaped tasting glasses in 30 ml units. The yellowness index of the brandies was determined by using LICO 500 as described by the supplier (León et al., 2006; Sharpe et al., 1992).

Table 6.1: Brandies evaluated in this study

Sample code	South African brandy style	Yellowness Index[b] (Gold colour intensity)	Alcohol content at full strength (ABV)[c]
OR1[a]	Potstill	83	38%
OR2	Potstill	83	38%
KG	Potstill	61	38%
CS	Potstill	14.5	38%
CM	Blended	68	43%
OV	Blended brandy	46.5	43%
VR	Blended brandy	73	43%
KX	Blended brandy	48	43%

[a]OR1 and OR2 are duplicates of the same brandy brand; [b]Yellowness index measured with LICO 500 tests.
[c]Alcohol by volume

6.2.2 Panellists

Ten people were recruited for this study from a pool of trained panellists employed at Distell Ltd. The panellists were screened for sensory acuity as described in our previous work (Louw et al., 2013). The panel consisted of women between the ages of 40 and 60 years. The panellists were experienced in brandy evaluation and projective mapping.

6.2.3. Experimental design

Three white paper tasting sheets were used namely, a rectangular shape of standard A3 paper dimensions (420mm x 297 mm), a square shape (297mm x 297 mm) and a round shape with a diameter of 295 mm. The panellists completed two replicate projective mapping evaluations on each paper shape over six sessions. The sessions were captured over a period of three weeks. The paper shapes were presented in a randomised order to minimise possible learning effects. The presentation order, as shown in Table 6.2, was generated using the "DOE for sensory data analysis" function in XLStat 2012.

Table 6.2: Experimental design

	Session 1	Session 2	Session 3	Session 4	Session 5	Session 6
Panellist 1	Square	Rectangle	Round	Round	Square	Rectangle
Panellist 2	Rectangle	Round	Square	Square	Rectangle	Round
Panellist 3	Rectangle	Square	Rectangle	Round	Square	Round
Panellist 4	Round	Rectangle	Square	Square	Round	Rectangle
Panellist 5	Square	Round	Rectangle	Rectangle	Round	Square
Panellist 6	Round	Round	Rectangle	Square	Rectangle	Square
Panellist 7	Rectangle	Square	Round	Square	Round	Rectangle
Panellist 8	Round	Square	Round	Rectangle	Rectangle	Square
Panellist 9	Square	Round	Rectangle	Round	Rectangle	Square
Panellist 10	Square	Rectangle	Square	Rectangle	Round	Round

6.2.4 The projective mapping task

The projective mapping task was performed in conjunction with ultra-flash profiling as described by Perrin et al., (2008). Panellists were instructed to smell and taste all the samples in the order received and to position them on a tasting sheet according to sensory similarity. Similar samples were to be positioned close together and very different samples, far apart. For the ultra-flash profiling stage, the panellists had to provide sensory descriptors for each sample.

6.2.5 Sensory evaluation conditions

All the evaluations were conducted in white tasting booths with controlled air conditioning and lighting to ensure unbiased responses. The panellists received all the brandies simultaneously, although the brandies were arranged in different, randomised orders on the sample trays, to reduce serving order effects. The samples were coded with 3-digit codes to eliminate bias. Distilled water and cream crackers were available to the panellists for the purpose of palate cleansing.

6.2.6 Statistical analyses

The projective mapping data were captured by measuring the coordinates of each sample relative to the centre of the tasting sheet. The descriptive terms were captured on a citation basis; if a panellist used a descriptor to describe a specific sample, that sample was allocated a value of 1, while the samples for which that descriptor was not used received a value of 0. Where panellists used similar descriptors, the scores for those descriptors were summed to create a single table of descriptive attributes. The data were analysed with multiple factor analysis (MFA) to obtain a consensus configuration. Each panellist's coordinates were used as a separate block of data in the MFA calculations, resulting in ten data blocks of two variables each. The descriptive data were included as passive or supplementary data i.e. they were given very low weights so that it could be superimposed onto the overall consensus configuration without influencing the sample scores generated from the coordinate data. Agglomerative hierarchical cluster analysis, using the Ward method, was used to identify sample clusters for each consensus configuration.

RV coefficients were used to compare the results from the three tasting sheet conditions (Abdi, 2007; Robert & Escoufier, 1976). An RV coefficient above 0.700 has been reported to indicate good similarity (Cartier et al., 2006). The distance (i.e. difference) between the two duplicated samples on the sheet, relative to the maximum inter-sample distance, expressed as a ratio D_r, was taken as a measure of accuracy (Torri *et al.*, 2013). Low values are interpreted as more accurate. For the overall consensus configurations, this measure was based on the Euclidean distances between the samples on the MFA consensus configuration. The between treatment differences were tested for each performance indicator with ANOVA where both shape and panellist were used as fixed effects. Interactions were tested where relevant. Significant differences were based on Type III Sum of Squares while individual differences were evaluated with the Fisher LSD post-hoc test.

Panellist performance was measured using the relative performance index (RPI). This indicator is based on explained variance after procrustes data transformation and provides a comparison of the product maps generated from each panellist's data. The RPI is calculated using Equation 1

EQ 1: $$Relative\ performance\ index = \frac{\left(sum\ of\ variances - \frac{SSQ}{n_{samples} \times n_{configurations}}\right)}{sum\ of\ variances}$$

Where SSQ = Residual Sum of Squares from procrustes ANOVA after compensating for rotation, translation and scaling done during procrustes analyses and n = number of samples or configurations in the data set, as annotated.

6.3 Results

6.3.1 Comparison of the overall consensus configurations

Table 6.3 shows the variance explained by the first two dimensions, F1 and F2, for each tasting sheet shape. For all three shapes, more than 54% of the total variance in the data was explained by F1 and F2. The highest amount of explained variance on each of the three shapes was observed for the square tasting sheet, most of which was explained on F1. On F1 and F2, similar levels of explained variance were observed for the rectangular and round sheets.

Table 6.3: Variance explained by the first two dimensions resulting from multiple factor analysis of projective mapping coordinates obtained from rectangular, square and round tasting sheets.

	Eigenvalue	Variability (%)	Cumulative %
Rectangle F1	9.513	36.196	36.196
Rectangle F2	4.944	18.810	55.006
Square F1	10.755	40.554	40.554
Square F2	4.584	17.283	57.837
Round F1	10.173	35.400	35.400
Round F2	5.458	18.991	54.390

6.3.1.1. Comparison of descriptors

After consolidating the descriptors cited by each panellist, a similar number of descriptors were obtained for each of the three sheet shapes (Table 6.4). The descriptor lists were dominated by aroma characteristics. Most of the descriptors were common to all three sheet shapes. On a panellist level, no significant shape effect, or shape*panellist interactions were observed (not shown) in terms of the number of descriptors used.

Table 6.4 Descriptors generated during the projective mapping evaluation of eight brandies based using three different tasting sheet shapes. Abbreviations are shown in brackets.

Attribute	Shape[a]	Attribute	Shape
alcohol burn (alc burn)	R,S,C	oaky	R,S,C
apple	R,S,C	oily mouthfeel (oily)	R,S,C
apple/pear	R	peach	R,S,C
bitter taste (bitter)	R,S,C	pear	S,C
burnt	R	prune	R,S,C
caramel	R,S,C	raisin	R,S,C
chocolate	R,S,C	sherry	R,S,C
citrus	R,S,C	smoky	R,S,C
coffee	R,S,C	soapy	R,S
dried fruit	R,C	sour taste (sour)	R,S,C
dried peach	S	spices	R,S,C
drying	S,C	sweet taste (sweet)	R,S,C
floral	R,S,C	general sweet flavours (sweetassociated)	R,S
fresh fruit	R,S	sweet spice	R,S
fruitcake	R,S,C	tea	C
grape	R,S	tobacco	R,S,C
herbaceous	C	toffee	C
honey	R,S,C	vanilla	R,S,C
medicinal	R,C	viscous mouthfeel (viscous	R,C
muscat	S,C	watery mouthfeel (watery)	R,S,C
mushroom	R	woody	R
nutty	R,S,C		
a) R=Rectangle, S=Square, C=Circle			

6.3.1.2 Rectangular paper shape

The samples were positioned in a dispersed pattern on the consensus configuration (Figure 6.1). The two duplicated samples (OR1 and OR2) were positioned in the lower right quadrant (D_r = 0.27).Cluster analysis showed three clusters which are indicated by connective lines on Figure 6.1: the first cluster consisted of darker potstill brandies (OR1, OR2 and KG; yellowness index 61-83). This cluster was associated with roasted characters such as burnt, coffee, smoky, and also caramel, prunes, raisins and spices. It was also associated with oily mouthfeel. The second cluster consisted of CS, OV and KX which were all light in colour (yellowness index 14.5-48) and associated with apple, pear, soapy and vanilla flavours as well as watery mouthfeel, alcohol burn, sweet and sour taste. Finally, the third cluster consisted of CM and VR which were darker blended brandies (yellowness index 68-73). This cluster was associated with the positive end of F2. It is unclear which attributes separated this cluster from the others; not many characters were associated with this area, except for honey, mushroom and to some extent viscosity and apple. F2 was negatively associated with coffee and woody flavour, the position of CM and VR could possibly be explained by less citations for these attributes.

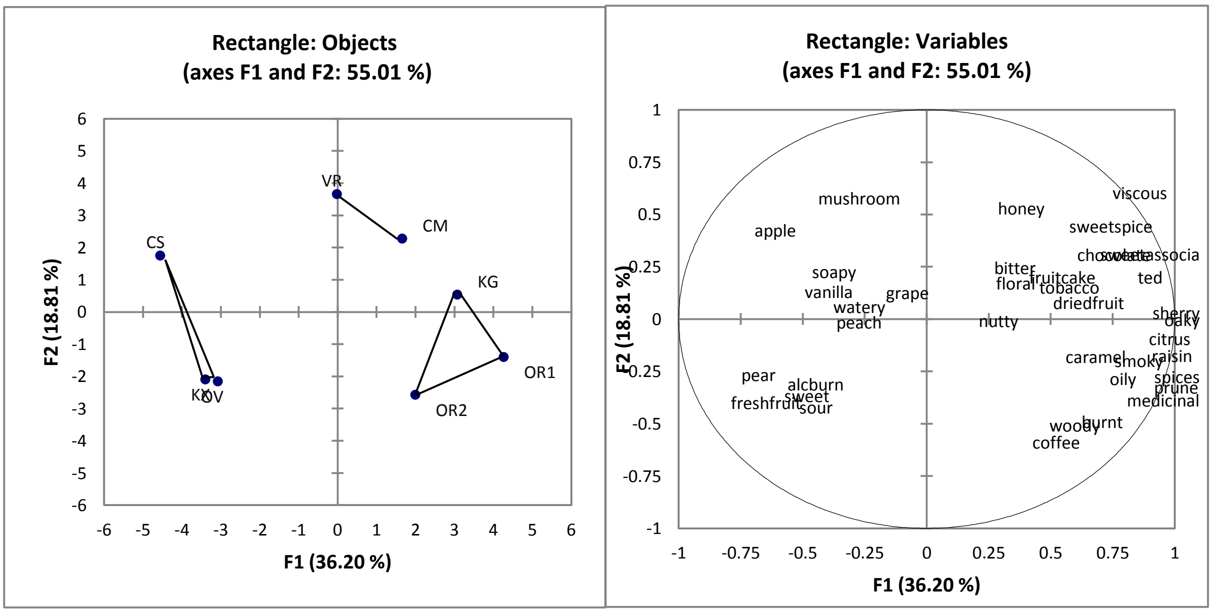

Figure 6.1 Consensus configuration showing the observations (left) and associated variables (right) from projective mapping of eight brandies using a rectangular tasting sheet. Connected samples belong to the same cluster as determined by agglomerative hierarchical cluster analysis.

6.3.1.3. Square paper shape

Figure 6.2 shows the square consensus configuration. The two duplicated samples were positioned close together (D_r= 0.19). Three clusters were observed with cluster analysis, shown by connective lines in Figure 6.2. The first cluster was located on the top left quadrant and consisted of OR1, OR2, KG and VR. VR was positioned a little further from OR1, OR2 and KG, which could be expected since it is a blended brandy, while the others are potstill brandies. It appears as if the cluster was associated with coffee, prune and sweet associated notes, as well as bitterness. VR seemed to be specifically associated with sherry and sweet associated notes, while OR1, OR2 and KG seemed to be associated with fruitcake, tobacco and oaky notes. The second cluster consisted of CM, KX and OV. These were positioned in the lower right quadrant, which was associated with floral, citrus and soapy notes and a drying mouthfeel. CS did not form part of a cluster and was associated with watery mouthfeel and fresh fruit. It must be noted that CS is considerably lighter in colour (yellowness index = 14.5), which may have contributed to a more watery perception as the product may have appeared diluted compared to the others. Although there was a slight overlap in yellowness index, the first cluster consisted of darker brandies than the second cluster (yellowness index 61-83 vs. 46.5-68).

6.3.1.4 Round paper shape

Figure 6.3 shows the round consensus configuration. The two duplicated samples (OR1 and OR2) were positioned very close together on the round consensus configuration (D_r= 0.04) (Figure 6.3). The same clusters, although tighter, were observed as in the rectangular configuration. OR1 and OR2, together with KG formed the first cluster (yellowness index 61-83) which was positioned on the far negative end of F1. It was associated with oaky, spice, smoky, fruitcake and caramel notes, as well as drying mouthfeel. CM and VR formed the second cluster (yellowness index 68-73), which is associated with toffee and dried fruit notes, more viscous mouthfeel and possibly more intense sweetness. On the positive end of F1, the third cluster (CS, KX and OV with yellowness index 14.5-48) were associated with vanilla, soapy, muscat, apple and pear notes.

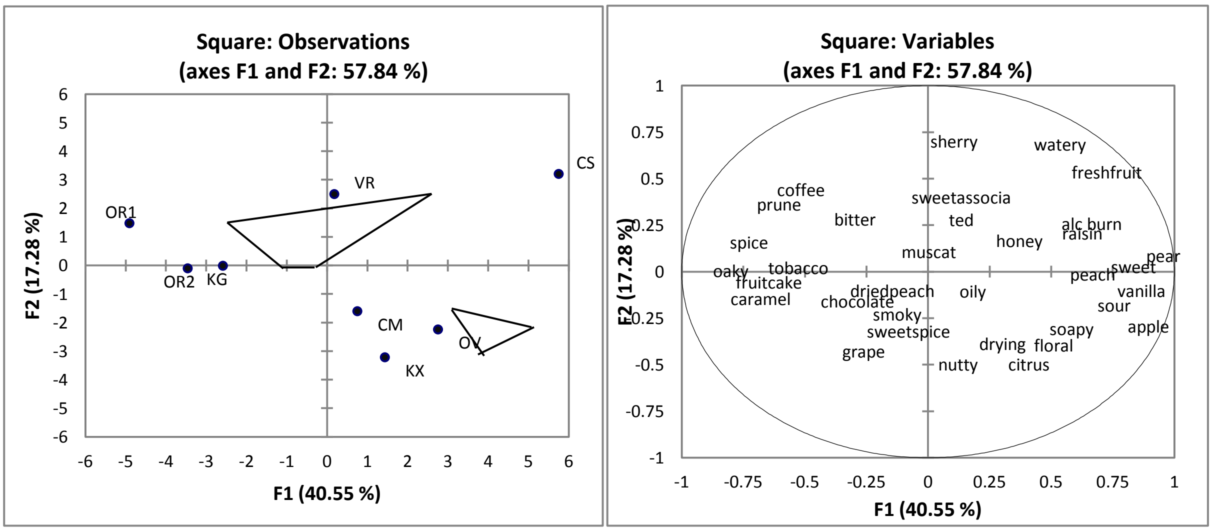

Figure 6.2 Consensus configuration showing the observations (left) and associated variables (right) from projective mapping of eight brandies using a square tasting sheet. Connected samples belong to the same cluster as determined by agglomerative hierarchical cluster analysis.

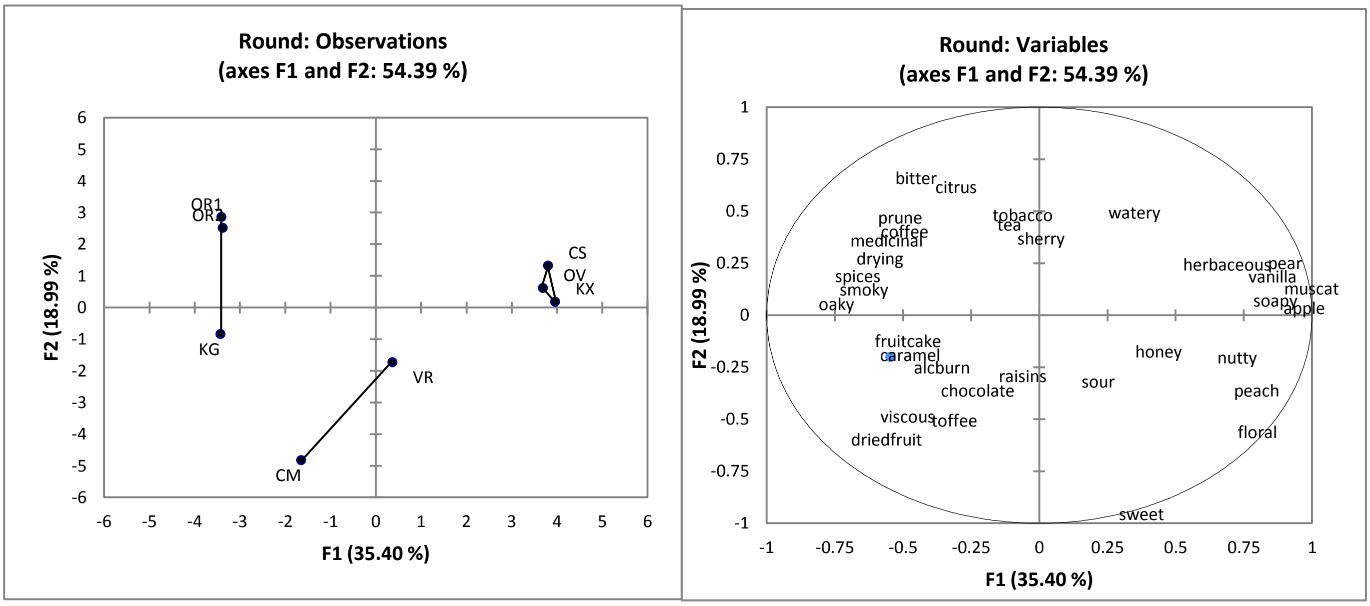

Figure 6.3 Consensus configuration showing the observations (left) and associated variables (right) from projective mapping of eight brandies using a round tasting sheet. Connected samples belong to the same cluster as determined by agglomerative hierarchical cluster analysis.

6.3.1.5 Comparison of shapes

Table 6.4 shows RV coefficients indicating the similarity between the tasting sheet shapes. Significant similarity was observed between all the tasting sheets. The rectangle and round shapes were most similar to each other. Both highlighted the differences between darker and lighter style brandies, but also between potstill and blended brandies. The square shape's configuration showed a progression on F1 from dark potstill brandies, to blended brandies (with no colour distinction) to a light potstill brandy. The square shape was more similar to the rectangle than to the round shape; both these shapes showed CS differentiating from KX and OV, while these three brandies were positioned very close to each other on the round shape's configuration. As seen in Figure 6.4, cluster analysis delivered similar sample clusters for the rectangular and round consensus configurations. However, notably different sample clusters were observed for the square tasting sheet.

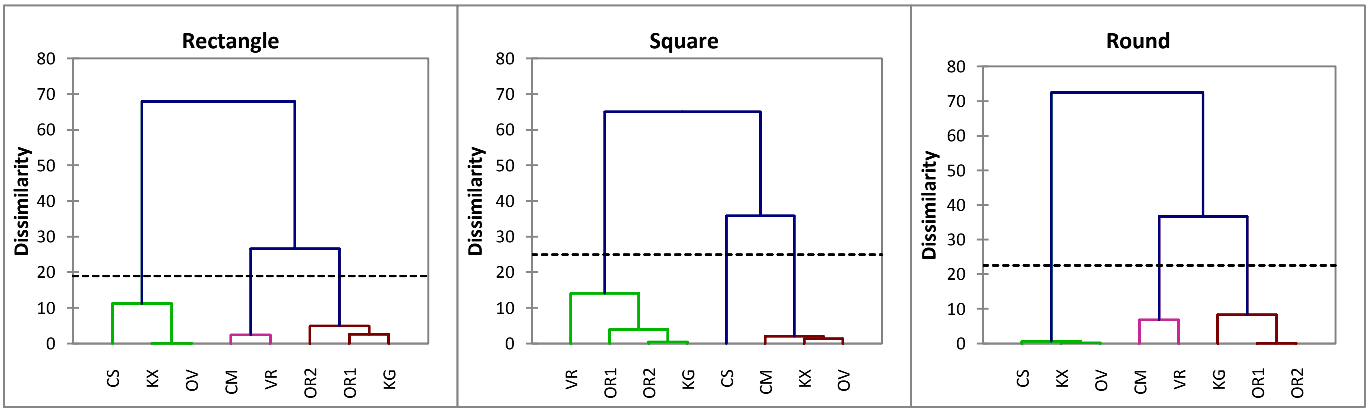

Figure 6.4 Dendograms showing the sample clusters obtained with agglomerative hierarchical cluster analysis of consensus configuration scores obtained from projective mapping of eight brandies on three different tasting sheet shapes.

Table 6.4 RV Coefficients (bottom, left) and p-values (top, right) showing significant similarity between projective mapping data captured on three tasting sheet shapes

Shape	Rectangle	Square	Round
Rectangle	1	0.003[b]	0.004[b]
Square	0.789[a]	1	0.006[b]
Round	0.822[a]	0.708[a]	1

[a]RV coefficient;[b]p-value

6.3.1.6 Influence of product colour as differentiating factor

All three sheet shapes highlighted lighter flavours (vanilla, fresh fruit) and watery mouthfeel, vs. darker flavours (oak, spice, caramel) and oily mouthfeel as the primary direction of differentiation. Previous research on wine evaluation has shown that panellists tend to associate the colour of the product with objects of a similar colour (e.g. a red wine smelling like red fruits) (Morrot et al., 2001; Parr et al., 2003).The terms used to describe the most important differences suggest that it has been influenced by sample colour intensity. Furthermore, the results from Chapter 5 have shown that colour has a strong, if subconscious, influence on panellists' projective mapping and sorting criteria of brandy. In this work, a strong correlation was observed between the X-axis en gold colour measurements for all three configurations (Table 6.5). A more intense golden colour correlated with a darker, oaky dimension, while lighter gold hues correlated with the watery, fresh fruit and vanilla dimension. The strongest correlation was observed for the square configuration while the weakest was observed for the round configuration.

Table 6.5 Pearson's correlation coefficients and p-values for correlation between sample scores on F1, F2 and gold colour hue.

Shape: Dimension	Pearson's correlation coefficient	p-value
Rectangle: F1	0.918	0.001
Rectangle: F2	-0.076	0.859
Square: F1	-0.925	0.001
Square: F2	0.019	0.964
Round: F1	-0.892	0.003
Round: F2	-0.023	0.957

6.3.2 Panellist performance

Previous studies reported that panellists are influenced by tasting sheet geometry in terms of the manner in which they use the horizontal and vertical axes (Dehlholm, 2013; Hopfer & Heymann, 2013). Therefor it was of interest to determine whether this influence affected their performance. Relative performance index (RPI) scores were calculated for each panellist for each tasting sheet shape (Figure 6.5). Panellist 10's scores were not included as she only completed one square evaluation. For the panellists as a group, no significant sheet shape effect was observed (p-value = 0.116). Since only one RPI value is recorded per panellist, it was not possible to determine the significance of a panellist*shape effect. Nevertheless, 60% of the panellists showed better performance using a round tasting sheet.

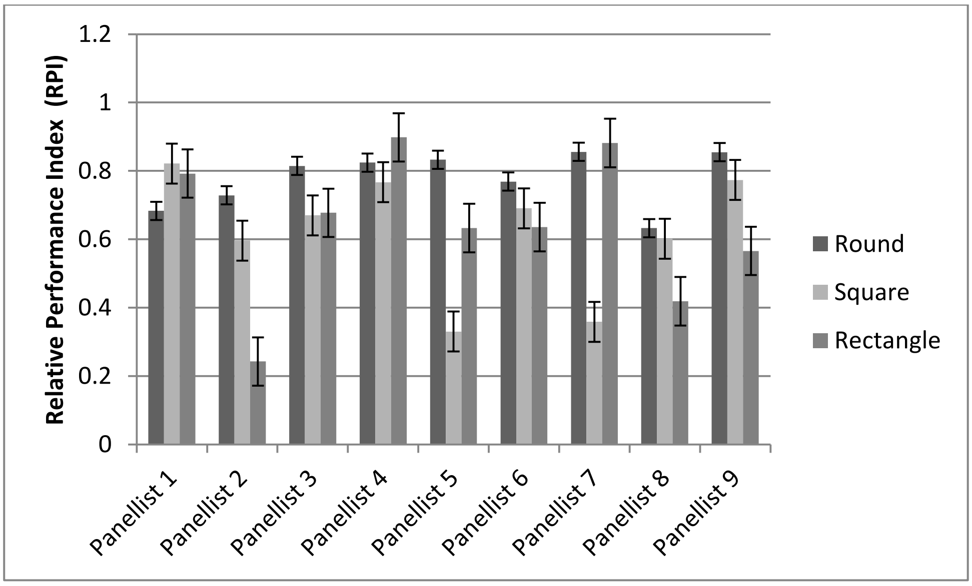

Figure 6.5 Relative performance index scores showing the nine trained panellists consistency in projective mapping of eight brandies using either a round, square or rectangular tasting sheet. Error bars denote standard errors.

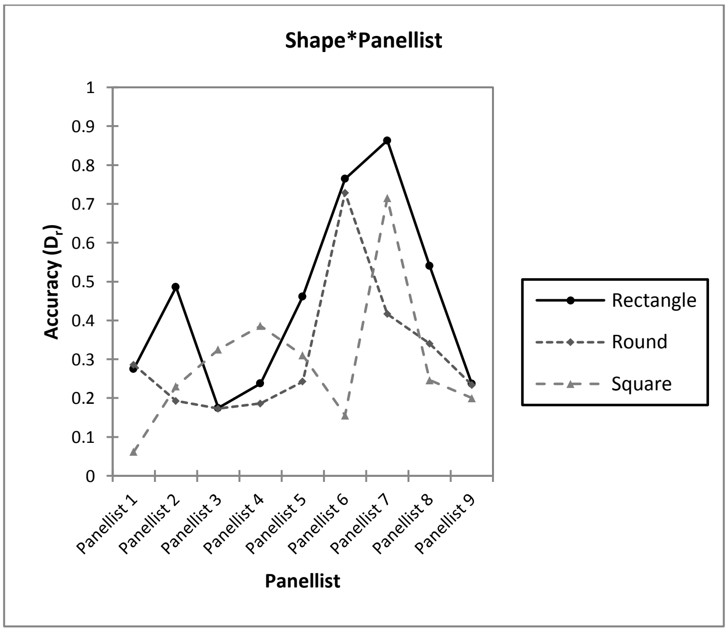

Figure 6.6 ANOVA interaction plot showing differences in nine panellists' accuracy in positioning duplicated samples close to each other during projective mapping of eight brandies using either a rectangular, round or square tasting sheet. Lower D_r scores indicate better accuracy. Panellists performed significantly worse on the rectangular sheet, while no significant panellists*shape effect was observed.

This observation is also reflected in the panellists' D_r scores, which measure the panellists' ability to position duplicated samples close to each other. There was a significant shape effect; overall the panellists D_r scores were worse for the rectangular tasting sheet. Although individual differences are visible on the panellist*shape interaction plot (Figure 6.6) the panellist*shape effect was not significant. No clear relationship could be established between the panellists' use of the X- and Y-coordinates and the sample colour.

6.4 Discussion and conclusion

This study confirmed previous reports that tasting sheet geometry impacts on projective mapping results. For the sample set tested, the round tasting sheet delivered by far more accurate results (i.e. proximity of the duplicated samples) than the square and rectangular tasting sheets. Overall, the rectangular tasting sheet delivered the least accurate results, although it should not be considered as inaccurate. The three tasting sheet shapes delivered very similar configurations. This is in agreement with Dehlholm's findings comparing rectangular and round tasting sheets, but not with Hopfer and Heymann's study on rectangular and square tasting sheets (Dehlholm, 2013; Hopfer & Heymann, 2013). Dehlholm observed that similarities between configurations of different shapes were larger where the explained variance is higher. In fact, the variance explained by the first two dimensions of the product configurations in this study is considerably higher than what was reported in Hopfer and Heymann's work (≈ 56% vs. ≈ 30%). The low reported explained variances in the latter study may have been a result of a very large and fairly homogenous sample set. Furthermore, Dehlholm's sample set consisted of six honeys while Hopfer and Heymann's work focussed on 18 wines. When a small number of samples are evaluated, it is more likely that similar sample configurations could be obtained even if different projective mapping strategies were used. Conversely, a large number of samples present more variations in which the samples can be projected onto the tasting sheet, increasing the chance that different configurations could be observed between sheet shapes. Our sample set of eight products were closer in size to Dehlholm's sample set than to Hopfer and Heymann's sample set, which may explain why the configuration differences in sheet shapes were more consistent with the former study.

Consistent with previous reports, the square tasting sheet resulted in different sample clusters than a rectangular sheet (Hopfer & Heymann, 2013). In our study, the square sheet delivered the most different results, while the round and rectangular sheets' results were more similar to each other. Considering that Hopfer and Heymann's work on square sheets reported a difference in overall product configuration while Dehlholm's work (2013) reported negligible differences between round and rectangular tasting sheets, our results suggest that changes in projective mapping sheet shapes are more relevant for square and rectangular shapes than for round shapes. Although it is possible that, as Dehlholm suggested, panellists' projective mapping strategies for square shape may be a compromise between their round shape strategy and rectangular shape strategy, it is not reflected in the overall product configurations.

There was a strong correlation between sample colour and the first dimension on all three consensus configurations; this was slightly lower for the round configuration compared to the square and rectangular configuration. This suggests that overall; colour was an important factor in the differentiation between products. However, no clear connection could be established between sample colour and the panellists' use of either the horizontal or vertical dimension of the tasting sheet.

The majority of the panellists provided more consistent results using a round tasting sheet. This is reflected in the tighter sample clusters observed in the round consensus configuration. However, this observation contradicts a previous statement that, based on the vertical aspect ratio of the human visual field, a rectangular tasting sheet could be a more natural tasting sheet geometry than a round one (Dehlholm, 2013). The small panel size is a limiting factor in this study; in order to fully understand the impact of tasting sheet shape on individuals it would be best to study the results from a very large pool of panellists.

It has been suggested that a rectangular tasting sheet with one axis being longer will guide panellists to use this dimension as a platform for indicating a few main differentiating factors between samples. Therefore, more symmetrical tasting sheets with equal perpendicular bisectors such as squares and circles could potentially highlight more subtle differences between samples, whereas elongated tasting sheets such as rectangles would highlight more obvious dimensions of differentiation (Dehlholm, 2013; Hopfer & Heymann, 2013). The results from the present study did not show that more subtle differences were highlighted by the round and square tasting sheets. This could have been an effect of the sample colour differences providing a very strong element of differentiation. It may be that any benefit of using an elongated tasting sheet to highlight gross differences vs. a more symmetrical tasting sheet to highlight subtle differences is more relevant in sample sets with a less obvious differentiating dimension.

Our study showed that the number of words generated by the panellists did not differ significantly between sheet shapes. This contradicts Dehlholm's study (2013) where more attributes were generated by the panel using a round sheet shape, but it must be taken into account that Dehlholm's study compared results from two different panels. From this study, it appears that using different sheet shapes would not affect descriptor yield.

The comparison of round tasting sheet shapes with the more readily available square and rectangular shapes is interesting from an explorative point of view, yet it must be kept in mind that from a practical point of view, the difference between square and rectangular tasting sheets are more important. It is recommended that a more visually homogenous sample set of a manageable size be used for further research in order to harness panellists' varying responses to different tasting sheet shapes. Black glasses could be used to reduce the role of colour. A future study testing the role of sample set homogeneity on the potential benefit of using elongated tasting sheets could also incorporate comparison of the two most common rectangular tasting sheets, 40 cm x 29.7 cm (A3) (Barcenas et al., 2004; Louw et al., 2013) and 60 x 40 cm (Pagès, 2005; Perrin et al., 2008), where the latter is double as long as the former.

6.5. References

Abdi, H.(2007). Rv and congruence coefficients. IN: *Encyclopedia of Measurement & Statistics.* (Ed (N.J. Salkind). SAGE publications Inc., 849-853.

Barcenas, P., Pérez Elortondo F., & Albisu, M. (2004). Projective mapping in sensory analysis of ewes milk cheeses: a study on consumer and trained panel performance. *Food Research International, 37,* 723-729.

Cartier, R., Rytz, A., Lecomte, A., Poblete, F., Krystlik, J., Belin, E., & Martin, N. (2006). Sorting procedure as an alternative to quantitative descriptive analysis to obtain a product sensory map. *Food Quality and Preference, 17,* 562-571.

Drake, S.L., Lopetcharat, K., & Drake, M.A. (2009). Comparison of two methods to explore consumer preferences for cottage cheese. *Journal of Dairy Science, 92,* 5883-5897.

Dehlholm, C. (2013) *Descriptive sensory evaluations: Comparison and applicability of novel rapid methodologies.* PhD Thesis, ISBN 978-87-7611-592-0.

Heymann, H., Machado, B., Torri, L., & Robinson, A.L. (2012). How many judges should one use for sensory descriptive analysis? *Journal of Sensory Studies, 27,* 111-112.

Hopfer, H. & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. *Food Quality and Preference, 28,* 164-181.

Kennedy J. & Heymann, H (2009). Projective mapping and descriptive analysis of milk and dark chocolates. *Journal of Sensory Studies, 24,* 220-233.

León, K., Mery, D., & León, J. (2006). Color measurement in L*a*b* units from RGB digital images. *Food Research International, 39 (10),* 1084-1091.

Louw, L., Malherbe, S., Næs, T., Lambrechts, M., Van Rensburg, P., & Nieuwoudt, H. (2013). Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products. *Food Quality and Preference, 30,* 192-201.

Morrot, G., Brochet, F., & Dubourdieu, D. (2001). *Brain and Language, 79,* 309-320.

Moussaoui, K.A., & Varela, P. (2010). Exploring consumer product profiling techniques and their linkage to a quantitative descriptive analysis. *Food Quality and Preference, 21,* 1088 – 1099.

Nestrud, M.A., & Lawless, H.T. (2008). Perceptual mapping of citrus juices using projective mapping and profiling data from culinary professionals and consumers. *Food Quality and Preference, 19,* 431-438.

Nestrud, M.A., & Lawless, H.T. (2010). Perceptual mapping of apples and cheeses using projective mapping and sorting. *Journal of Sensory Studies,* 25, 390-405.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire valley. *Food Quality and Preference, 16(7),* 642-649.

Parr, W.V., White, K.G., & Heatherbell, D.A. (2003). The nose knows: influence of colour on perception of wine aroma. *Journal of Wine Research, 14,* 79-101.

Perrin, L., Symoneaux, R., Maitre, I., Asselin, C., Jourjon, F., & Pages, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: case of wines from Loire Valley. *Food Quality and Preference, 19,* 1-11.

Risvik, E., McEwan, J.A., Colwill, J.S., Rogers, R., & Lyon, D.H. (1994). Projective mapping: a tool for sensory analysis and consumer research. *Food Quality and Preference, 5,* 263-269.

Risvik, E., McEwan, J.A., & Rødbotten, M. (1997). Evaluation of sensory profiling and projective mapping data. *Food Quality and Preference, 8,* 63-71.

Robert, P. & Escoufier, Y.(1976). A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Applied Statistics, 25,* 257-265.

Sharpe, F.R., Garvey, T.B., & Pyne, N.S. (1992). The measurement of beer and wort colour – a new approach. *Journal of the Institute of Brewers, 98,* 321-324.

South African Department of Agriculture (1989). Regulations 12 (1 d) and 13 (2) of the South African Liquor Products Act 60 of 1989, *Government Gazette, 33193.*

Torri, L., Dinella, C., Recchia, A., Næs, T., Tuorila, H., & Monteleone, E. (2013). Projective mapping for interpreting wine aroma differences as perceived by naïve and experienced assessors. *Food Quality and Preference, 29,* 6-15.

Veinand, B., Godefroy, C., Adam, C., & Delarue, J. Highlight of important product characteristics for consumers. Comparison of three sensory descriptive methods performed by consumers. *Food Quality and Preference, 22,* 474 – 485.

Chapter 7

Research results

Trained sensory panellists' response to product alcohol content in the projective mapping task: observations on alcohol content, product complexity and prior knowledge

This manuscript was published in
Food Quality and Preference

Chapter 7: Trained sensory panellists' response to product alcohol content in the projective mapping task: observations on alcohol content, product complexity and prior knowledge

Abstract

Projective mapping has been validated as a practical tool for the rapid sensory profiling of brandy products, although repeatability concerns necessitate repeated measurements in larger sample sets. The reason for poor repeatability could be linked to the complexity of the product type, as well as the physical and possibly psychological factors associated with its high alcohol content. To date no information has been published that tested the effect of these specific factors on panellist performance in projective mapping tasks. This study tested the effect of sample complexity and alcohol content on sensory panel repeatability and accuracy in projective mapping, using six types of commercial alcoholic beverages. In a second objective, the study also tested the effect of prior knowledge of alcohol content of a given product set on panellist performance in projective mapping. The results showed that complexity had the biggest impact on panel performance, while alcohol content had a secondary but decisive influence, largely due to its chemosensory fatiguing nature. Knowledge of the product alcohol content appeared to affect individuals differently, and also had an effect on the terminology used by the panellists to describe the products. The study also introduces the Relative Performance Index (RPI) as a new indicator for panel performance monitoring in projective mapping.

7.1 Introduction

Brandy is a complex grape-based distilled beverage with an alcohol content of at least 36% ABV (alcohol by volume), as specified by EU regulations (European Union, 2008). Many different styles and types of brandy are produced across the globe. Well-known and protected styles include French cognac, Spanish Brandy de Jerez, Portuguese Lourinhã brandy, Chilean Pisco and South African potstill brandies (Robinson, 1999). Sensory evaluation of these products is important to ensure quality products that meet consumer demands.

Projective mapping, also known as Napping®, (Risvik et al., 1994; Pagès, 2005) is a rapid sensory profiling method designed to obtain a holistic overview of the sensory differentiation between products in a given sample set, without the time- and cost impact of conventional sensory profiling methods such as quantitative descriptive analysis (QDA™) (Stone et al*.,* 1974). When it comes to alcoholic beverages, projective mapping has only been applied to wine (Pagès, 2005; Perrin et al*.,* 2008; Perrin & Pagès, 2009, Ross et al*.,* 2012, Hopfer & Heymann, 2013, Torri et al., 2013). The wines tested included white wines from the Loire valley as well as red wines from France, Italy and the USA. Although the alcohol contents were not specified, the expected range for these wine styles is 11-15% ABV. One of these studies reported a maximum alcohol content of 15.3% ABV (Hopfer & Heymann, 2013; chemical analyses reported in related study in Hopfer et al*.,* 2012). Spirit beverages, such as brandy, are typically diluted to 20-23% ABV before sensory evaluation (Louw & Lambrechts, 2012). Our earlier work was the first study on rapid sensory profiling of spirit beverages and projective mapping was validated as a suitable method for brandy evaluation (Louw et al., 2013). The results showed good accuracy and repeatability for a small set of six brandies per evaluation. However, in comparison, when a larger set of ten brandies per session was evaluated, the repeatability of the method decreased, and repeated measurements were recommended to improve the quality of the results (Louw et al., 2013).

Considering the nature of brandy, we speculated in our earlier work that the decrease in panel performance could be due to sensory fatigue caused by the samples. Different types of fatigue relevant to sensory evaluation have been identified (Sauvageot, 1990). Those relevant to brandy evaluation include sensory and mental fatigue that may be induced by the inherent product properties and possibly psychological fatigue that may be induced by panellists' expectations of the product type and what the evaluation thereof, would involve. For high alcohol beverages, panellists may for example expect the product to elicit a stronger burning sensation than a low alcohol beverage or that it may cause them to tire more easily.

Alcohol is a chemosensory irritant which may cause sensory fatigue through continuous stimulation of the trigeminal senses (Green, 1988; Prescott & Swain-Campbell, 2000). As projective mapping relies on holistic, comparative product evaluation, sensory analysts are more restricted in the measures that can be taken to compensate for fatigue induced by high alcohol content than in conventional profiling where samples are presented one at a time. However, the effect of alcohol content on panel performance, and subsequently data quality, has not specifically been explored in literature.

Product complexity also complicates sensory evaluation, by leading to mental fatigue amongst panellists and hence poor performance. It has been suggested that a less analytical sensory approach is more suitable to complex samples than intensity scaling, based on the argument that the overall odour perception of complex products cannot be accurately broken down into independent, measurable attributes (Lawless, 1999). This often results in a sensory lexicon that is limited to a few descriptors that can be accurately scaled, ignoring many other attributes that may be present but for which panel consensus regarding their definition and intensity could not be achieved (Lawless, 1999). An approach that could deal with this issue would be to allow panellists to indicate, instead of quantify, which terms are important to describe the product by providing them with an extensive list of descriptors relevant to the product category (Lelièvre et al., 2008; Campo et al., 2010), or allowing them to supply their own words to describe the product, as is done in the Napping® procedure (Perrin et al., 2008). Product complexity has been implicated to impact on the quality of projective mapping results (Nestrud & Lawless, 2010), although this observation was based on fruit and dairy studies. The complex volatile structure of brandy elicits a considerable number of sensory perceivable nuances (Jolly and Hattingh, 2001), which can complicate the projective mapping task by making it more difficult for the panellist to decide which attributes are the most important. To date, there is no information available on the effect of the complexity of alcoholic beverages on panel performance in projective mapping.

As mentioned previously, panellists' assessment of spirit beverages may be influenced by their expectations of the product and the task of evaluating it. Panellists' expectations from information received or inferred prior to product evaluation are some of the many cognitive factors that can influence the way that trained panellists perceive and evaluate products (Lawless and Heymann, 1999; Schifferstein, 1996). Panellists may expect to perceive certain attributes based on verbal cues given by the panel leader, or from non-verbal cues obtained from the product itself. Qualitative judgments made on product information such as nutritional information has shown to also affect quantitative product assessment (Schifferstein, 1996). Confidence in task competency has been linked to motivation and performance of trained sensory panellists (Lund et al., 2009). It is possible that panellists may form expectations around task difficulty based on product type and information; panellists may associate high alcohol beverages with sensory fatigue, mild intoxication and/or increased task difficulty. However, there is no information on whether sensory panellists' performance in the evaluation of spirit products is affected by their knowledge of the products' alcohol content.

Projective mapping studies tend to report on panel performance by comparing panellists with each other, but very few report on the individual panellists' internal consistency. Some researchers have used the panellists' physical projective maps to determine their task competency, i.e. whether samples were placed in straight lines, or scattered across the entire sheet (Pagès, 2005; Nestrud & Lawless, 2008). RV coefficients between data from repeated sessions have been used to determine the repeatability of individuals (Kennedy, 2010). Panellist performance has been evaluated by their ability to position two duplicated samples close to each other on the projective mapping sheet. This is expressed as a ratio of the Euclidean distance between the two duplicate samples and the maximum inter-sample Euclidean distance in the sample set. This ratio has been referred to as the People Performance Index (PPI) (Hopfer & Heymann, 2013) and also as a $D_{r\%}$ ratio (Torri et al., 2013). The drawback of this ratio is that it provides information on the panellists' consistency in positioning only one sample. Procrustes Analysis of Variance (PANOVA) has been used to determine panel consistency by evaluating total consensus variance for overall consistency and product residuals to determine whether there were any specific products that the panellists disagreed on (Nestrud & Lawless, 2008). Although this approach provides information the panel's consistency for all samples, it does not provide a single interpretable measure.

In this study it was of interest to gain better understanding of the sensory, mental and psychological fatigue causing factors that influence panel performance in projective mapping of spirit beverages, and two separate research objectives were identified. The first was to investigate the effects of alcohol content and product complexity, using an experiment design to vary these two factors, on panellist performance in the projective mapping task. The aim of this experiment was to evaluate which of these product characteristics would be the most important risk factor in brandy evaluation. The second objective was to determine to what extent panellists' performance is affected by prior knowledge of the alcohol content of a given sample set. In other words, the objective was to gain insight into the cognitive impact of high alcohol content on panellist performance. With panellist performance being a key concern in this study, a new performance monitoring measure, will be introduced.

7.2 Methods and Materials

7.2.1 Panellists

The panels that participated in this study consisted of women between the ages of 23 and 60 that are employed as trained sensory panellists at Distell Ltd, South Africa. They were screened for sensory acuity according to the guidelines in Stone and Sidel (1992). The screening test included threshold testing for basic tastes, aroma identification, memory recall for aromas, discrimination ability, intensity ranking and participation in a mock panel situation. The panel was experienced in conventional sensory profiling as well as projective mapping of various types of alcoholic beverages, including brandy. Nine women participated in the study that investigated the effect of product alcohol content and complexity on panel performance, while ten women participated in the study investigating the effect of prior knowledge of alcohol content on panel performance.

7.2.2 Samples

7.2.2.1 Objective 1: The effect of product alcohol content and complexity on panel performance in projective mapping of alcoholic beverages

Six sets of commercial alcoholic beverages, ten products each, were presented to the panellists (Table 7.1). As the ultimate purpose of the study was to gain better understanding of the difficulties associated with brandy evaluation, it was decided to use ten samples per set, based on our previous work that showed that this number is challenging for the panel (Louw et al., 2013). The sets represented high alcohol products (20% ABV) and low alcohol products (≈ 7% ABV). Three levels of complexity were included in the design for each level of alcohol content. In this case, complexity was defined as the perceived complexity, i.e. the number of attributes used to describe the overall perception of the product, as well as the degree of homogeneity in the sample set, i.e. the ratio of the number of attributes per sample relative to the total number of attributes used to describe the whole product set. Data from previous sensory profiling studies on various types of alcoholic beverages, which were conducted at the sensory research facility at Distell Ltd, were evaluated to identify product types of varying complexity. These studies were conducted independently from the current study and from each other. The number of attributes for which a certain product scored higher than 0 on a 100 mm intensity scale was taken as an indication of perceived product complexity. This value was also divided by the total number of attributes measured in the study to give ratios of the number of attributes contributing to the overall perception of each product relative to the total number of attributes in the product set. This ratio can be regarded as an indication of sample set homogeneity. In each sample set, one product was presented twice as a blind duplicate to test for accuracy. The duplicated samples were chosen in such a way that they would not be obviously different from the others and therefore easily recognisable. Each sample set was presented twice to test repeatability. The panellists were not informed of the purpose of the study.

Table 7.1 Six sets (10 samples each) of commercial alcoholic beverage of different levels of alcohol content and perceived complexity.

	Low Complexity	Medium Complexity	High complexity
High alcohol (20% ABV)	HaLc[a] 20[b]; 0.48[c]	HaMc 22; 0.56	HaHc 27; 0.71
Low alcohol (7% ABV)	LaLc 12; 0.31	LaMc 15; 0.50	LaHc 20; 0.57

[a]Abbreviation used in text; [b]Sample complexity: Average number of attributes recorded per sample; [c]Sample set homogeneity: Ratio of number of attributes present per sample relative to total number of attributes used to describe sample set.

Standard serving practices were followed for each product type. The low alcohol products were refrigerated until 15 minutes prior to tasting. The high alcohol beverages were diluted with distilled, odourless water from their full alcohol strength (38% ABV to 43 % ABV) to 20% ABV one hour prior to evaluation. The high alcohol products were served at room temperature.

7.2.2.2 Objective 2: The effect of prior knowledge of product alcohol content on panel performance in projective mapping of spirit beverages

Five sample sets, consisting of eight products each were served in this study (Table 7.2). Each set consisted of three brandy brands (B1 – B3), diluted with three brands of non-alcoholic mixers of the same flavour (M1 - M3). The B3M3 combination was served twice to test the panellists' ability to identify that the two duplicate samples are the most similar. In three of the sample sets, the brandies were diluted from their original alcohol strength (38% ABV – 43% ABV) to 7% ABV (first three columns of Table 7.2). For these three, the panellists were either given no information (U), correct information (I) or incorrect information (M) regarding the alcohol strength. In the latter case (M) they were told that the samples were high in alcohol, when it was in fact low (7%). The remaining two sample sets were diluted to 20% ABV. For these, the panellists were given no information (U) or correct information about the alcohol content (I). It was not possible to create a credible misinformed scenario for the high alcohol content sample sets as the perceived alcohol burn would make it quite obvious that the samples are high in alcohol and not low alcohol as informed by the panel leader. Each sample set was presented twice to test repeatability.

Table 7.2 Abbreviations for samples tested at various alcohol strengths (L-low and H-high) and information regarding alcohol content given prior to evaluation (I-informed, U-uninformed, M-misinformed). In the sample abbreviations, the letters B1-B3 denote the brandy brands and the letters M1-M3 denote the mixer brands.

7% ABV (L) Uninformed (U)	7% ABV (L) Informed (I)	7% ABV (L) Misinformed (M)	20 % ABV (H) Uninformed (U)	20 % ABV (H) Informed (I)
B1M1_L_U	B1M1_L_I	B1M1_L_M	B1M1_H_U	B1M1_H_I
B1M2_L_U	B1M2_L_I	B1M2_L_M	B1M2_H_U	B1M2_H_I
B1M3_L_U	B1M3_L_I	B1M3_L_M	B1M3_H_U	B1M3_H_I
B2M1_L_U	B2M1_L_I	B2M1_L_M	B2M1_H_U	B2M1_H_I
B2M2_L_U	B2M2_L_I	B2M2_L_M	B2M2_H_U	B2M2_H_I
B2M3_L_U	B2M3_L_I	B2M3_L_M	B2M3_H_U	B2M3_H_I
B3M3_L_U	B3M3_L_I	B3M3_L_M	B3M3_H_U	B3M3_H_I
B3M3_L_U	B3M3_L_I	B3M3_L_M	B3M3_H_U	B3M3_H_I

ABV: percentage alcohol per volume; Uninformed: no information was given; Informed: Correct information was given; Misinformed: panellists were told the products were strong in alcohol when it was not.

7.2.3 The projective mapping task

The projective mapping task was performed in conjunction with ultra flash profiling as described by Perrin et al., (2008). Panellists were instructed to smell and taste all the samples in the order received and to position them on an A3 sheet of paper according to sensory similarity. Similar samples were to be positioned close together and very different samples, far apart. They were provided with scrap paper on which they could write down their perceptions. The panellists had to provide sensory descriptors for each sample.

7.2.4 Testing conditions

The tastings were conducted in white tasting booths under ambient lighting and controlled temperature. The samples were served in a random order with consideration for first order serving effects. The panellists received 30 ml of product in standard 250 ml tulip shaped tasting glasses. The panellists received all the samples at the same time.

7.2.5 Statistical analyses

All analyses were performed in XLStat version 2013.2.03. The projective mapping coordinates were measured relative to the centre of the tasting sheet. The coordinates were analysed with general procrustes analysis (GPA). RV coefficients were calculated as a measure of similarity between the repeated measurements of each panellist. During GPA, noise caused by rotation, translation and scaling are removed to generate an optimal consensus map. A Relative Performance Index (RPI), based on the variance explained by GPA after data transformation, was used as an indicator of the similarity between the product maps generated in the replicated sessions for each panellist. Both RPI and the RV coefficients test repeatability, but RV coefficients are more relevant to the data structure before statistical analysis, RPI values are more relevant to the resulting product maps. It can be said that the RV coefficient tests repeatability of the panellists' actual measurements, while the RPI tests the repeatability of the panellists' resulting product maps.

EQ 1: $$Relative\ performance\ index = \frac{\left(sum\ of\ variances - \frac{SSQ}{n_{samples} \times n_{configurations}}\right)}{sum\ of\ variances}$$

Where SSQ = Residual Sum of Squares from Procrustes ANOVA after compensating for rotation, translation and scaling during GPA and n = number of samples or configurations, as annotated.

Higher RV and RPI coefficients indicate better similarity. RV coefficients of 0.700 have been suggested as a cut-off point for good similarity (Cartier et al., 2006). An appropriate cut-off point for the RPI has not been confirmed, but will for the present be evaluated against the same standard. The panellists' accuracy was measured using the Peoples Performance Index (PPI) as suggested by Bertuccioli (2011) and applied by Hopfer and Heymann (2013). The index involves dividing the Euclidean distance between duplicated samples in a sample set by the maximum Euclidean distance in the sample set. Lower PPI values indicate better accuracy. The between treatment differences were tested for each performance indicator with ANOVA. Panellists were regarded as a fixed effect, since variance between panellists was a specific interest in this study. Significant differences were based on Type III Sum of Squares while individual differences were evaluated with the Fisher LSD post-hoc test.

7.3 Results

7.3.1 Objective 1: The effect of alcohol content and complexity on panel performance in projective mapping of alcoholic beverages

7.3.1.1 Repeatability of measurements

Figure 7.1 shows the effect of alcohol content and complexity on the RV coefficients between two repeated measurements by the same panellist. At low alcohol levels, sample set complexity did not appear to affect panellists' repeatability. However, at high alcohol levels, a significant step-wise decrease in repeatability was observed as the sample sets became more complex. In the high and medium complexity levels, alcohol content did not have a significant impact on repeatability. However, much higher RV coefficients was observed for HaLc than for LaLc.

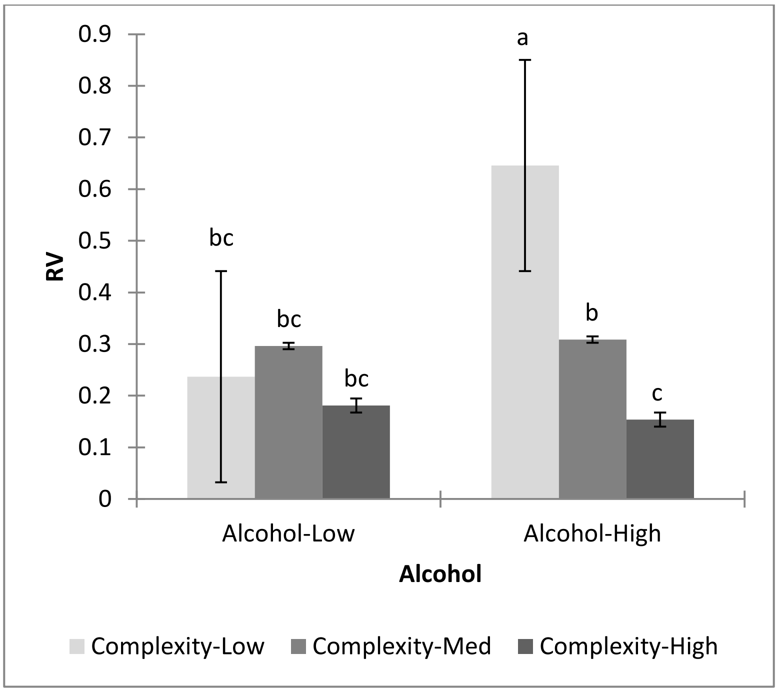

Figure 7.1 Average RV coefficients showing trained panellists' repeatability in projective mapping for beverage sample sets of varying degree of complexity and alcohol content. Higher values indicate better repeatability. Letter notations denote Fisher LSD values as an indication of the observed effects. Error bars denote standard error.

7.3.1.2 Repeatability of product maps

At 20% ABV, complexity had a significant effect on the panellists' RPI values (Figure 7.2). The panellists' produced the least repeatable product maps for the HaHc sample set. There were significant stepwise increases in the panellists' RPI values from HaHc to HaMc and finally to HaLc. At 7% ABV, an increase in complexity did not significantly affect the panellists' RPI values, although the panellists appeared to be somewhat less repeatable at high complexity than at medium and low complexity. Alcohol content did not appear to affect RPI values at the different levels of complexity. No difference was observed between HaHc and LaHc or between HaMc and LaMc. However, at low complexity, there was a significant difference between the alcohol levels, in favour of HaLc.

7.3.1.3 Accuracy in recognising duplicate samples

At both alcohol levels, the panellists' accuracy decreased as the sample sets became more complex (Figure 7.3). At 20% ABV HaLc had significantly lower PPI values than HaHc and HaMc. At 7% ABV the difference was less pronounced; with LaLc being significantly lower than LaHc but not than LaMc. At all three complexity levels, the panellists were more accurate in the low alcohol sample set than in the high alcohol sample set. This difference was statistically significant at medium complexity.

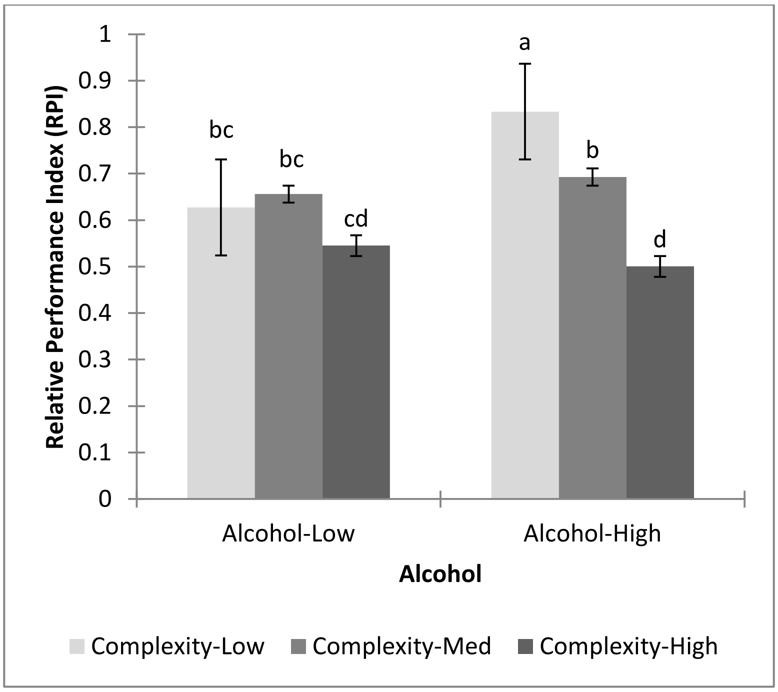

Figure 7.2 Relative performance index showing trained panellists' ability to generate reproducible product maps with projective mapping for beverage sample sets of varying degree of complexity and alcohol content. Higher values indicate better performance. Letter notations denote Fisher LSD values as an indication of the observed effects. Error bars denote standard error.

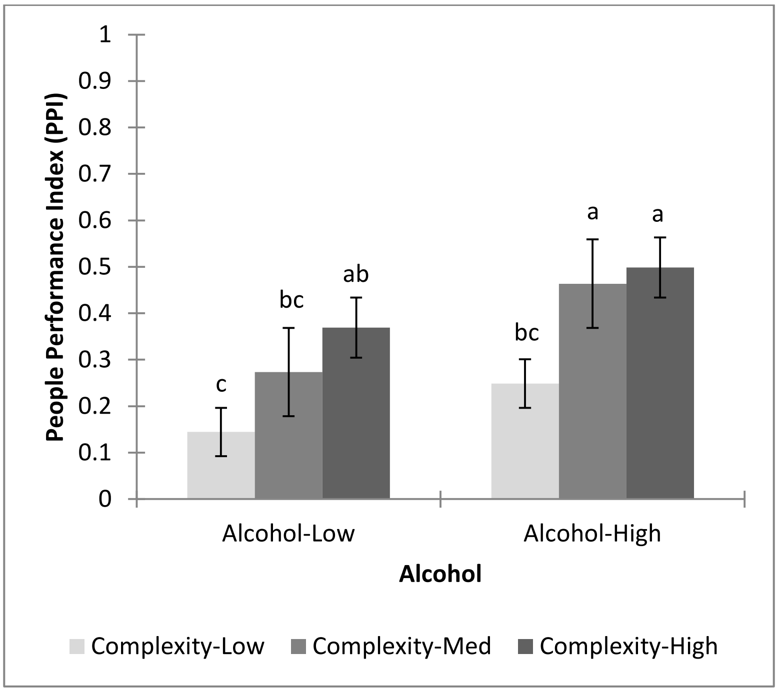

Figure 7.3 People performance index (PPI) showing trained panellists' accuracy in projective mapping for alcoholic beverage sample sets of varying degree of complexity and alcohol content. Lower values indicate more accurate responses. Letter notations denote Fisher LSD values as an indication of the observed effects. Error bars denote standard error.

7.3.2 Objective 2: The effect of prior knowledge of alcohol content on panellist performance in projective mapping of spirit beverages

7.3.2.1 Repeatability of the measurements and product maps

In this case, there was a very high correlation between the RV coefficients and RPI values (Pearson correlation coefficient = 0.945; p-value: <0.0001) and the same conclusions can be drawn from the two coefficients. For the sake of brevity, only the RPI values will be discussed as an indication of repeatability.

There was noticeably lower similarity between two replicate sessions when panellists were correctly informed that they were evaluating high alcohol products (H_I) compared to the other tasting conditions (Figure 7.4). The same effect was not observed in the case of L_M when panellists thought they tasted high alcohol beverages when they were in fact tasting low alcohol beverages. One could argue that the effect of the information served to strengthen the physiological impact of the high alcohol rather than reducing performance *per se*.

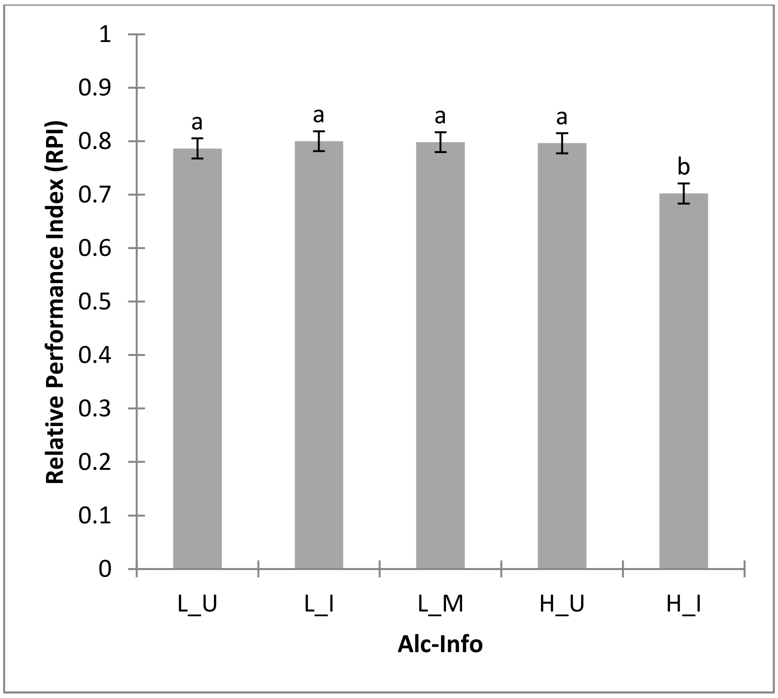

Figure 7.4 Panellist repeatability in the projective mapping task for sample sets with high alcohol (H) and low alcohol content (L). Panellists were either given no information (U), correct information (I) or incorrect information (M) regarding the alcohol strength. Error bars denote standard error while letter notations on bars denote Fisher LSD values.

7.3.3.2 Accuracy in recognising duplicate samples

Information regarding the alcohol content did not significantly affect accuracy in this experiment (Figure 7.5). However, a moderate panellist*treatment effect was observed (p =0.081). Figure 7.6 shows the average PPI score for three panellists for each treatment. Panellist 7 had very similar PPI scores over the different testing conditions, Panellist 6 performed better under low alcohol conditions, with her performance at L_M being more similar to H_I and H_U. In contrast, Panellist 2 performed better under high alcohol conditions, with her performance at L_M also being better than at L_I.

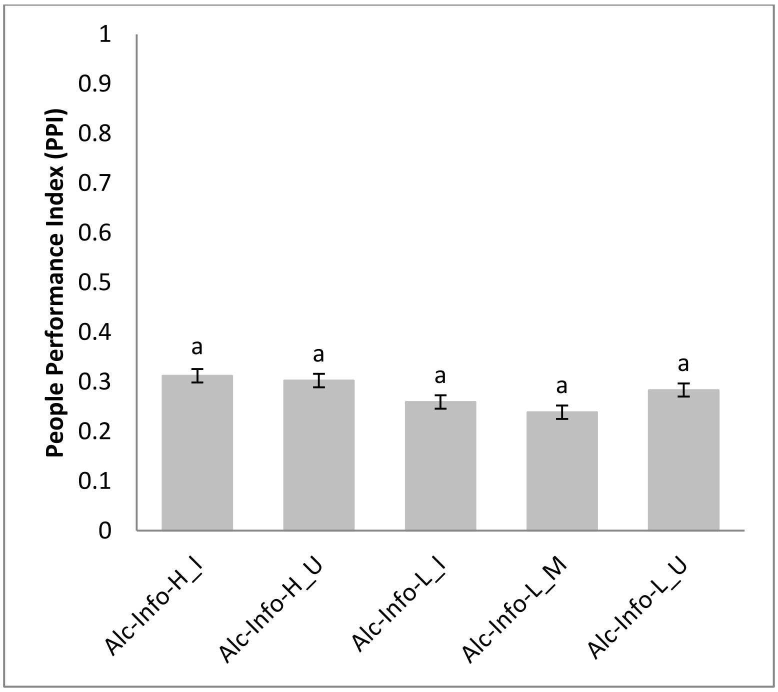

Figure 7.5 Panellist accuracy in the projective mapping task for sample sets with high alcohol (H) and low alcohol content (L). Panellists were either given no information (U), correct information (I) or incorrect information (M) regarding the alcohol strength. Error bars denote standard error while letter notations on bars denote Fisher LSD values.

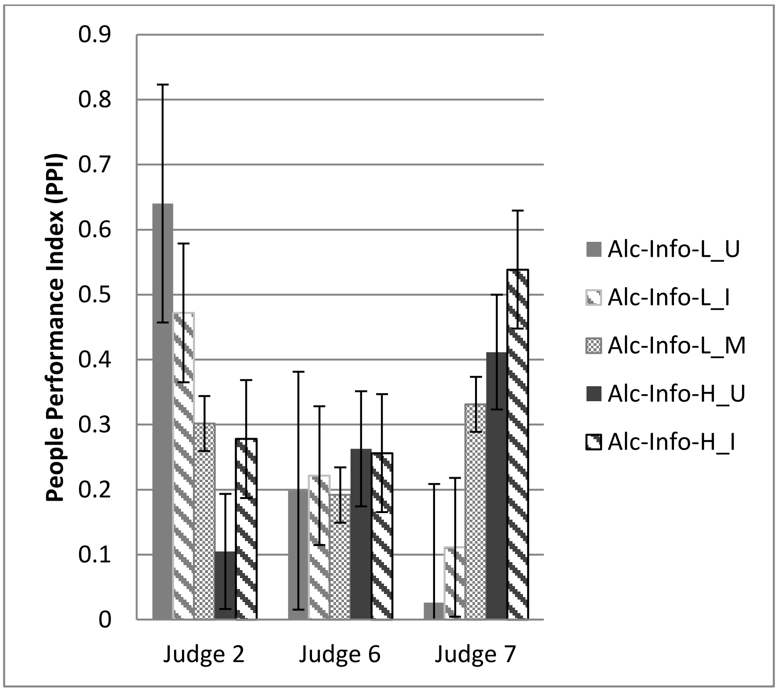

Figure 7.6 Individual differences in accuracy with which panellists are able to perform the projective mapping task under two alcohol content conditions (High alcohol (H) and low alcohol content (L)) and three information conditions where panellists were either given no information (U), correct information (I) or incorrect information (M) regarding the alcohol strength. Error bars denote standard error.

7.3.3.3 Use of alcohol related descriptors

Generally, the panellists used only one or two attributes relating to alcohol to describe the samples tested in this experiment. However, as a group, their use of alcohol related words was significantly more prevalent in the high alcohol conditions than in the low alcohol conditions (Figure 7.7). Their word use in the L_M condition did not differ significantly from the high alcohol conditions, even though the samples were low in alcohol.

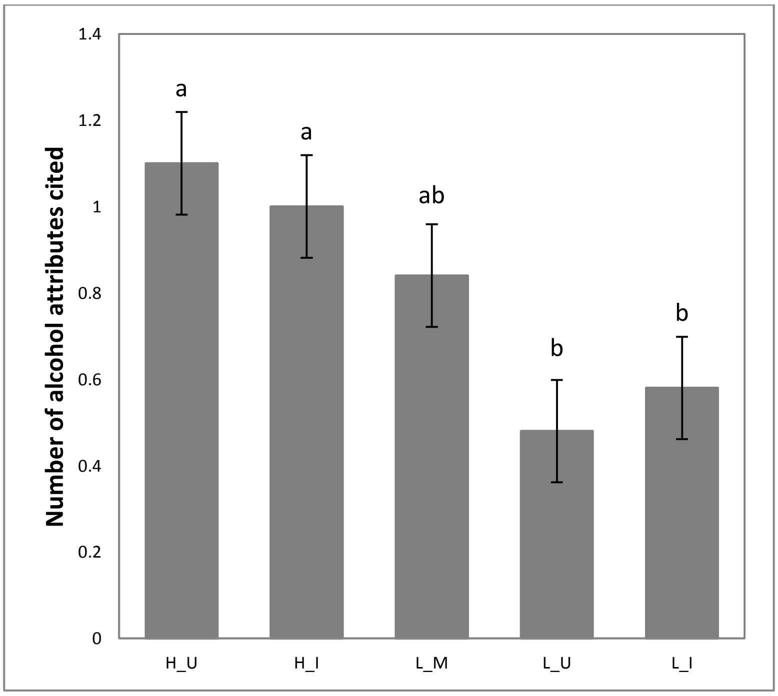

Figure 7.7 Differences in the overall number of alcohol related attributes used in the projective mapping task under two alcohol content conditions (High alcohol (H) and low alcohol content (L)) and three information conditions where panellists were either given no information (U), correct information (I) or incorrect information (M) regarding the alcohol strength. Error bars denote standard error while letter notations on bars denote Fisher LSD values.

7.4 Discussion

7.4.1 The effect of alcohol content and complexity on panel performance in projective mapping of alcoholic beverage

The purpose of this part of the study was to gain a better understanding of the influence of high alcohol content and sample complexity on panel performance in projective mapping of brandy. Towards this end panellists' performance in projective mapping of six types of commercial beverages, including brandy, was compared. One must first consider the implications of the choice for evaluating commercial beverages over model beverages. Previous studies on wine suggest that alcohol content has a significant impact on the odour activity and perception of volatile compounds, but that the effect is not to the same for all compounds. Alcohol does not only affect the solubility of individual compounds, but also the perceptual synergy between compounds (Goldner et al., 2009; Le Berre et al., 2007; Villamor et al., 2012). Furthermore, it was also suggested that alcohol content can influence the extent to which other components such as glycerol and polysaccharides, influence aroma perception in wine (Jones et al., 2008). This implies that even if all other components remain constant, panellists' sensory perceptions of a set of products will invariably differ as soon as the alcohol content is adjusted. For this reason, it was decided not to use model beverages, but rather commercial beverages. However, this also means that conclusions drawn from the results of this study must take the variation of the product type into account.

Keeping this information in mind, two key observations can be made regarding the effect of alcohol content and complexity on panel performance. The first is that the complexity of the sample set had a larger impact on panellist repeatability than alcohol content, in the range tested. In fact, despite some differences observed between LaLc and HaLc, alcohol content did not appear to have a significant effect on panellist repeatability at all. Considering the minimal effect at medium and high complexity, the difference observed at low complexity was likely due to the fact that the HaLc samples were more one-dimensional than the LaLc samples rather than the difference in alcohol content. It must be noted that the repeatability of LaMc was slightly lower than LaLc, although not significantly so. It is possible that the panellists changed their criteria on which they based their product positioning from one session to another, which could have a big effect if a sample set is very heterogeneous like LaLc. Shifts in projective mapping and sorting criteria have been reported in previous studies on sorting and projective mapping (Chollet & Valentin, 2001; Kennedy, 2010). Repeated measurements are recommended to compensate for criteria changes in projective mapping. Previous studies have also highlighted the value of repeated measurements towards ensuring valid results (Hopfer & Heymann, 2013; Louw et al*.,* 2013).

The finding that sample set complexity proved to have a significant effect on panel performance, supports circumstantial evidence from other studies on the impact of product complexity on stability of results in rapid sensory profiling (Nestrud & Lawless 2010; Delarue & Siefferman, 2004). The complexity of the sample set proves to be an important factor to take into account for the effective execution of projective mapping. Evaluating sample sets with a relatively low degree of differentiation in the set is a more complicated task with a higher risk of affecting panel performance than a sample set with a high degree of differentiation e.g. brandies compared to flavoured vodkas. If the samples are also perceptually complex, as is the case with brandy and wine, this can further complicate the task and increase the risk of poor panel performance. This supports Hopfer and Heymann's speculation that sample set homogeneity may be a contributing factor to high variability in panellist performance in projective mapping (Hopfer & Heymann, 2013).

A second important observation is that the effect of complexity on panellist repeatability was enhanced at higher alcohol content. Alcohol in itself can add to product complexity in the sense that flavoured vodkas can be perceived as more complex than flavoured waters. Among high alcohol beverages, different spirit product types can vary vastly in terms of volatile complexity. Vodka is by definition a neutral spirit with very few volatile components. In contrast, brandies have a very complex volatile structure, and even within the brandy category, product types can differ in complexity, depending on the percentage copper pot distilled content. Similar variations in compositional complexity can be found among low alcohol products such as spirit coolers, ciders and beers. Having a high alcohol content does not make a product set difficult to evaluate, as is indicated by HaLc. This study shows that the effect of complexity on task difficulty, as inversely expressed by panellist performance, is larger for spirit beverages than for low alcohol beverages. The enhancing effect of alcohol content could possibly be attributed to underlying sensory fatigue caused by the chemosensory irritation from the ethanol. Another hypothesis is that the perceived complexity of a compositionally complex spirit product is enhanced by its alcohol content and that this adds to task difficulty. The enhancing effect of alcohol content is likely due to a combination of these factors, although other explanations are not excluded.

7.4.2 The effect of prior knowledge of alcohol content on panellist performance in projective mapping of brandy

The second objective of this study was to determine whether sensory panellists' performance in brandy evaluation is influenced by their knowledge of the alcohol content. Although the panel leader can attempt to minimise the amount of information conveyed to the panel, the panellists are still able to infer information regarding the product. This is especially relevant in alcoholic beverage studies since the type of product can be a cue to the possible alcohol content, and also because alcohol itself is detectable with the senses. This part of the study investigated whether trained panellists' knowledge of the alcohol content of the products they evaluated affected their performance in projective mapping. It is uncertain exactly what information panellists would associate with high alcohol products. Panellists may expect that alcohol related attributes such as alcohol burn may be more prevalent in high alcohol beverages. This may influence their perception of the products and possibly the way in which they convey their perceptions in projective mapping. They may also expect, or at least be aware of the risk of the mild intoxication that may accompany the evaluation of high alcohol beverages. Such awareness may influence their expectation of the task difficulty and subsequently their performance.

Brandies mixed with different non-alcoholic mixers were chosen as stimuli as it would not be immediately obvious what the product type is and that the panellist could therefore not make any inferences regarding the alcohol content before tasting the product. Our results provide indicative information on the possible effect of panellists' expectation of alcohol content on their performance and approach to the projective mapping task.

Two important observations are highlighted in this study. Firstly, the effect of prior knowledge of a product's alcohol content on accuracy in the projective mapping task can differ from panellist to panellist as was illustrated in Figure 6. The differences in their responses to the information given about the sample conditions could mean that the panellists formed different expectations from the information that they received. For instance, Panellist 6 had very similar PPI scores over the different testing conditions; she may not have had any expectations from her knowledge of the alcohol content of the beverages she had to evaluate. Panellist 7 performed much better under L_U and L_I conditions than under the H_U, H_I and L_M conditions. She may have expected that a low alcohol task will be easier than a high alcohol task based on previous experiences. The information provided about the alcohol content, whether true or false, may have influenced her confidence and motivation and subsequently her performance in the task. On the other hand, Panellist 2 had the exact opposite results, which is performing better under high alcohol conditions than low alcohol conditions. It may be that she also expected the high alcohol task to be more difficult, but instead of responding with demotivation, she responded by concentrating harder on the task. Collecting data from more than two sessions, as was done in this case, and perhaps also in different product types, would provide more insight into whether these tendencies persist.

A second important finding relates to the use of alcohol related characteristics to describe the differences they perceived between the products. In the low alcohol conditions, the panel used significantly less alcohol related words than in the high conditions. However, for the L_M sample set, their use of alcohol related words did not differ significantly from high alcohol sample sets. A possible reason for this is that they may have allocated more value to the alcohol related attributes they perceived, thinking that the samples are high in alcohol than they may have felt necessary under the low alcohol conditions. It must be considered that the results of projective mapping rely on the relative value that panellists allocate to the different sensory attributes present in the products being evaluated. The results of this study provides reason for caution that alcohol content expectancies may influence which product attributes the panellists may consider as important or not. As with other stimulus errors, providing panellists with product information should be avoided.

7.4.3 The relative performance index (RPI) as a measure of panellist performance

A final observation must be made regarding the use of RPI versus RV coefficients to measure panellist repeatability. The conclusions drawn from RV coefficients and RPI values were largely similar. However, some important differences could be observed. Firstly, the panellists RPI values comparing sessions were consistently higher than their RV coefficients comparing sessions. Previous studies on projective mapping have also reported that, based on RV coefficients, that panellist repeatability can be poor despite stable overall configurations (Risvik et al., 2007; Kennedy, 2010). In our study, the average RV coefficient between sessions was 0.402, also indicating quite poor repeatability. In comparison, the average RPI were much higher, around 0.702. In a study on granola bars, Kennedy (2010) speculated that poor panellist repeatability may be related to the possibility that panellists change their criteria by which they differentiate between samples. RPI is based on the optimal consensus between the replicate sessions after procrustes data transformation and therefore corrects for variation in the panellist's assessments, whether due to environmental factors or a conscious decision on the panellist's part. RPI may be a more realistic assessment of the panellist's perception of the products rather than the way that they approached the products and could be a very relevant panel management tool. Also, RPI values can be used as a single measure to determine consistency over several data configurations, whereas the RV coefficient only compares two data configurations. Therefore, if multiple repeated sessions are conducted, RPI values would provide a more comprehensive measure for panellist repeatability. In the case of brandy evaluation, where the risk of the product complexity to panel performance is enhanced by its high alcohol content, three replicate sessions may be more practical than two sessions, in which case RPI would be an especially useful measure of panel performance.

7.5 Conclusion

The high alcohol content and complexity of brandy has been identified as risk factors that can potentially affect panellist performance in the sensory evaluation thereof. Complexity appears to be the most important contributing factor, although alcohol content plays an important secondary role. From the preliminary results from this study, the role of alcohol content appears to be largely physiological through its chemosensory fatiguing effect. However, the performance of some individuals may also be influenced on a cognitive level by their knowledge of the products' alcohol content. As projective mapping is increasingly being used as an alternative to conventional profiling, it would be prudent to take all possible risk factors into account when applying the method to brandy. By using repeated measurements, restricting the amount of information given to panellists and using enough panellists to compensate for individual differences, many of the identified risks could be effectively addressed in future studies.

7.6 References

Bertuccioli, M. (2011). *The role of sensory analysis to identify factors contributing to wine performance.* Graz, Austria: Keynote lecture at the In Vino Analytica Scientia.

Campo, E., Ballester, J., Langlois, J., Dacremont, C., & Valentin, D. (2010). Comparison of conventional descriptive analysis and a citation frequency-based descriptive method for odor profiling: An application to Burgundy Pinot noir wines. *Food Quality and Preference*, *21,* 44-55.

Cartier, R., Rytz, A., Lecomte, A., Poblete, F., Krystlik, J., Belin, E., & Martin, N. (2006). Sorting procedure as an alternative to quantitative descriptive analysis to obtain a product sensory map. *Food Quality and Preference, 17,* 662-571.

Chollet, S., & Valentin, D. (2001). Impact of training on beer flavour perception and description: are trained and untrained subjects really different? *Journal of Sensory Studies*, *16,* 601-618.

Delarue, J., & Siefferman, J.M. (2004). Sensory mapping using Flash Profile. Comparison with a conventional descriptive method for the evaluation of the flavour of fruit dairy products. *Food Quality and Preference, 15,* 383-391.

European Union. (2008). Regulation (EC) No 110/2008 of the European Parliament and of the Council of 15 January 2008. *Official Journal of the European Union*, L 39/30-31.

Goldner, M. C., Zamora, M.C., Di Leo Lira, P., Gianninoto, H. & Bandoni, A. (2009). Effect of ethanol level in the perception of aroma attributes and the detection of volatile compounds in red wine. *Journal of Sensory Studies*, *24,* 243-257.

Green, B.G. (1988). Spatial and temporal factors in the perception of ethanol irritation on the tongue. *Perception and Psychophysics*, *44,* 108-116.

Hopfer, H., Ebeler, S.E., & Heymann, H. (2012). How blending affects the sensory and chemical properties of red wine. *American Journal of Enology and Viticulture*, *63,* 313-324.

Hopfer, H. & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. *Food Quality and Preference*, *28,* 164-181.

Jolly, N.P., & Hattingh, S. (2001). A brandy aroma wheel for South African brandy. *South African Journal of Enology and Viticulture*, *22,* 16-21.

Jones, P.R., Gawel, R., Francis, I.L., & Waters, E.J. (2008). The influence of interactions between major white wine components on the aroma, flavour and texture of model white wine. *Food Quality and Preference, 19,* 596-607.

Kennedy, J. (2010). Evaluation of replicated projective mapping on granola bars. *Journal of Sensory Studies*, *25,* 672-684.

Lawless, H.T. (1999). Descriptive analysis of complex odors: Reality, model or illusion? *Food Quality and Preference, 10 (4-5),* 325-332.

Lawless, H.T., & Heymann, H. (1999). *Sensory Evaluation of Food: Principles and Practices* (pp 330) Gaithersburg, Maryland: Aspen Publishers.

Le Berre, E., Atanasova, B., Langlois, D., Etiévant, P. & Thomas-Danguin, T. (2007). Impact of ethanol on the perception of wine odorant mixtures. *Food Quality and Preference*, *18,* 901-908.

Lelièvre, M., Chollet, S., Abdi, H., & Valentin, D. (2008). What is the validity of the sorting task for describing beers? A study using trained and untrained assessors. *Food Quality and Preference*, *19,* 697-703.

Louw, L., Malherbe, S., Næs, T., Lambrechts, M., Van Rensburg, P., & Nieuwoudt, H. (2013). Validation of two Napping® techniques as rapid sensory screening tools for high alcohol products. *Food Quality and Preference, 30,* 192-201.

Louw, L., & Lambrechts, M.G. (2012). Grape-based brandies: production, sensory properties and sensory evaluation. IN: *Alcoholic beverages: sensory evaluation and consumer research* (Ed J Piggot). Cambridge: Woodhead Publishing Limited. p292-294.

Lund, C.M., Jones, V.S., & Spanitz, S. (2009). Effects and influences of motivation on trained panellists. *Food Quality and Preference*, *20,* 259-303.

Nestrud, M.A., & Lawless, H.T. (2008). Perceptual mapping of citrus juices using projective mapping and profiling data from culinary professionals and consumers. *Food Quality and Preference, 19,* 431-438.

Nestrud, M.A., & Lawless, H.T. (2010). Perceptual mapping of apples and cheeses using projective mapping and sorting. *Journal of Sensory Studies, 25,* 390-405.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire valley. *Food Quality and Preference. 16(7),* 642-649.

Perrin, L., & Pagès, J. (2009). A methodology for the analysis of sensory typicality judgments. *Journal of Sensory Studies, 24,* 749-773.

Perrin, L., Symoneaux, R., Maitre, I., Asselin, C., Jourjon, F., & Pages, J. (2008). Comparison of three sensory methods for use with the Napping® procedure: case of wines from Loire Valley. *Food Quality and Preference, 19,* 1-11.

Prescott, J., & Swain-Campbell, N. (2000). Responses to repeated oral irritation by capsaicin, cinnamaldehyde and ethanol in PROP tasters and non-tasters. *Chemical Senses, 25,* 239-246.

Risvik, E., McEwan, J.A., Colwill, J.S., Rogers, R., & Lyon, D.H. (1994). Projective mapping: a tool for sensory analysis and consumer research. *Food Quality and Preference, 5,* 263-269.

Risvik, E., McEwan, J.A., & Rødbotten, M. (1997). Evaluation of sensory profiling and projective mapping data. *Food Quality and Preference, 8,* 63-71.

Robinson, J. (1999). *Oxford Companion to Wine,* (2nd Ed.). Oxford, Oxford University Press.

Ross, C.F., Weller, K.M., & Alldredge, J.R. (2012). Impact of serving temperature on sensory properties of red wine as evaluated using projective mapping by a trained panel. *Journal of Sensory Studies, 27,* 463-470.

Sauvageot, F.(1990). Fatigue in sensory evaluation: myth or reality? In: *Flavours and Off-flavours.* Ed G Charalambous. Elsevier, London, pp 577-593.

Schifferstein, H.N.J., 1996.Cognitive factors affecting taste intensity judgments. *Food Quality and*

Preference, 7, 167-175.

Stone, H., Sidel, J., Oliver, S., Woolsey, A., & Singleton, R.C. (1974). Sensory evaluation by quantitative descriptive analysis. *Food Technology, 28:11,* 24-34

Stone, H., & Sidel, J.L. (1992). *Sensory Evaluation Practices.* (2nd Ed.). Orlando, FL: Academic Press.

Torri, L., Dinnella, C., Recchia, A., Næs, T., Tuorila, H., & Monteleone, E. (2013). Projective Mapping for interpreting wine aroma differences as perceived by naïve and experienced assessors. *Food Quality and Preference*, *29,* 6-15.

Villamor, R.R., Evans, M.A., Mattinson, D.S., & Ross, C.F. (2013). Effects of ethanol, tannin and fructose on the headspace concentration and potential sensory significance of odorants in a model wine. *Food Research International*, *50,* 38-45.

Chapter 8

General discussion and conclusions

Chapter 8: General discussion and conclusions

8.1 General discussion

Brandy has been defined as a compositionally complex grape-based spirit beverage containing at least 36% ABV (alcohol by volume) and often more than 43% ABV. It is consumed across the world and is especially important in the South African spirit beverage market. The combination of its high alcohol content and complex volatile chemical composition makes sensory evaluation of brandy very difficult. Rapid sensory profiling methods are increasingly being used to understand the sensory characteristics of food and beverage products and its importance to consumers. Despite these successful applications to other product types, rapid sensory profiling methods have not been validated for spirit beverages and prior to this study it was not clear whether applying these methods to brandy would be feasible. Rapid sensory profiling methods require comparative sample presentation, thereby limiting the systems that can be put in place to control sensory fatigue during brandy evaluation, especially for larger sample sets. Moreover, previous reports suggested that product complexity can impact on the stability of the results obtained with rapid sensory profiling methods.

The primary objective of this work was to determine whether rapid sensory profiling methods are feasible for brandy and other spirit products such as whisky. Although there are various rapid sensory profiling methods, each with their own advantages and disadvantages, projective mapping was chosen as a starting point. Projective mapping, or Napping®, is a flexible and intuitive method that is well adapted to product experts and trained panellists. It can be used in its basic format to obtain a snapshot of overall differentiation in a sample set. Using the partial napping approach it can also be used to elicit information focussed on specific sensory modalities. It not only provides information on the sensory differentiation between samples, but also provides sample specific descriptors. This intuitive approach encourages the elicitation of complex attributes that are not easily measured.

The study was conducted in two, inter-related, tiers. The first tier, indicated with dashed arrows in Figure 8.1, addressed research questions identified from the literature review: 1) the need to validate projective mapping for brandy evaluation, 2) the lack of measurable panel performance monitoring indicators for projective mapping, 3) the limited knowledge on factors influencing panel performance in projective mapping of alcoholic beverages and 4) the practical significance of using different projective mapping tasting sheet shapes. The second tier, indicated with solid arrows in Figure 8.1, addressed further research opportunities identified from the outcomes of the first tier. These were closely related to the first tier objectives.

The first step in this study was to validate projective mapping for brandy evaluation (Figure 8.1, step 1). Two versions of projective mapping were tested: global napping and partial napping. Partial napping is a restrictive form of the basic method where the task is focussed on a specific sensory modality. This study confirmed that partial napping of brandy delivered results more similar to conventional profiling than what was be obtained with global napping of brandy. This is in agreement with observations on other product types (Dehlholm et al., 2012; Pfeiffer & Gilbert, 2008). This study managed to build further on this statement by observing that, in the context of brandy evaluation, this advantage of partial napping over global napping is only evident in a larger sample set of 10 products and not in a small sample set. As the sample set gets larger, there are more factors to take into account in having to position the samples on the projective mapping space; providing more focus through the partial napping approach makes the task easier for the panellists and hence more reliable results can be obtained. One of the drawbacks of both partial napping and global napping was that mouthfeel attributes that were highlighted in conventional profiling did not appear as important in the projective mapping results. The reason for this was that though the panellists perceived the differences in mouthfeel, they disregarded the difference during sample placement. It was also found that less experienced panels relied more on visual cues than panellists that had a little more exposure to the method and products. In fact, the results suggested that prior training on smaller sample sets may prepare a trained panel better for a larger sample set, resulting in more reliable results. In general partial napping was found to be more suitable for larger sample sets, although reliable results could be obtained with global napping, provided that repeated sessions were captured.

Considering that partial napping was found to be the more effective of the two methods, a further investigation was initiated to optimise the method for brandy evaluation by scrutinising the results from the individual modalities evaluated (Figure 8.1, step 2). The panellists' performance was compared based on three modalities: visual, smell and in-mouth sensations (defined as a combination of retronasal odours, taste and mouthfeel). The results showed that the construct of in-mouth sensations were not very reliable and did not contribute to the overall product configuration; this is in line with suggestions in the spirit industry that nosing may be a sufficient proxy for retronasal flavour (Jack, 2003; Peña y Lilo et al., 2005). In fact, in-mouth sensations seemed to be predominantly driven by retronasal odour descriptions. In view of the potential importance of mouthfeel to consumers, it was necessary to investigate ways to harness panellists' perceptions of mouthfeel during projective mapping more effectively. It was already observed in the difference between global napping and partial napping that a semantic restriction delivered richer vocabulary regarding the modality of interest. Explicitly restricting the task to taste and mouthfeel (i.e. the deliberate exclusion of retronasal odour) provided more useful information. Since panellists tended to place a larger emphasis on colour and smell, it was important to determine whether the presence of these modalities would affect the panellists in terms of their ability to evaluate taste and mouthfeel accurately as a separate modality. It was shown that as individuals, the panellists performed equally well, regardless of whether they could perceive the colour or the retronasal odour of the product. As a group, being able to see and smell the product provided the most accurate and clearly interpretable results and for this reason it is suggested that brandy be served in clear glasses during projective mapping unless the exclusion of appearance is relevant to the specific research question.

Guidelines have been obtained towards the successful implementation of partial napping for brandy from Chapters 3 and 4. However, partial napping is more labour intensive than global napping as more data are generated, even if only one session is captured. With the aim of further time and cost-reduction, the question was raised whether an alternative holistic method could be used to capture an overview of a larger set of brandies in one session (Figure 8.1, step 3). Sorting was chosen as the alternative as it has already been shown to be compatible with larger data sets. By means of a hypothetical sample screening study in Chapter 5, it was found that representative sample selections can be made using a single session of sorting. However, the results also indicated that sorting is less descriptive and less accurate than global napping. Therefore, global napping is recommended for studies with no follow-up, where only rough descriptions of the samples and overall differentiation in the product set are required. However, sorting appeared to be a more suitable method for screening samples down into a subset where further action will be taken that would provide more detailed sample descriptions.

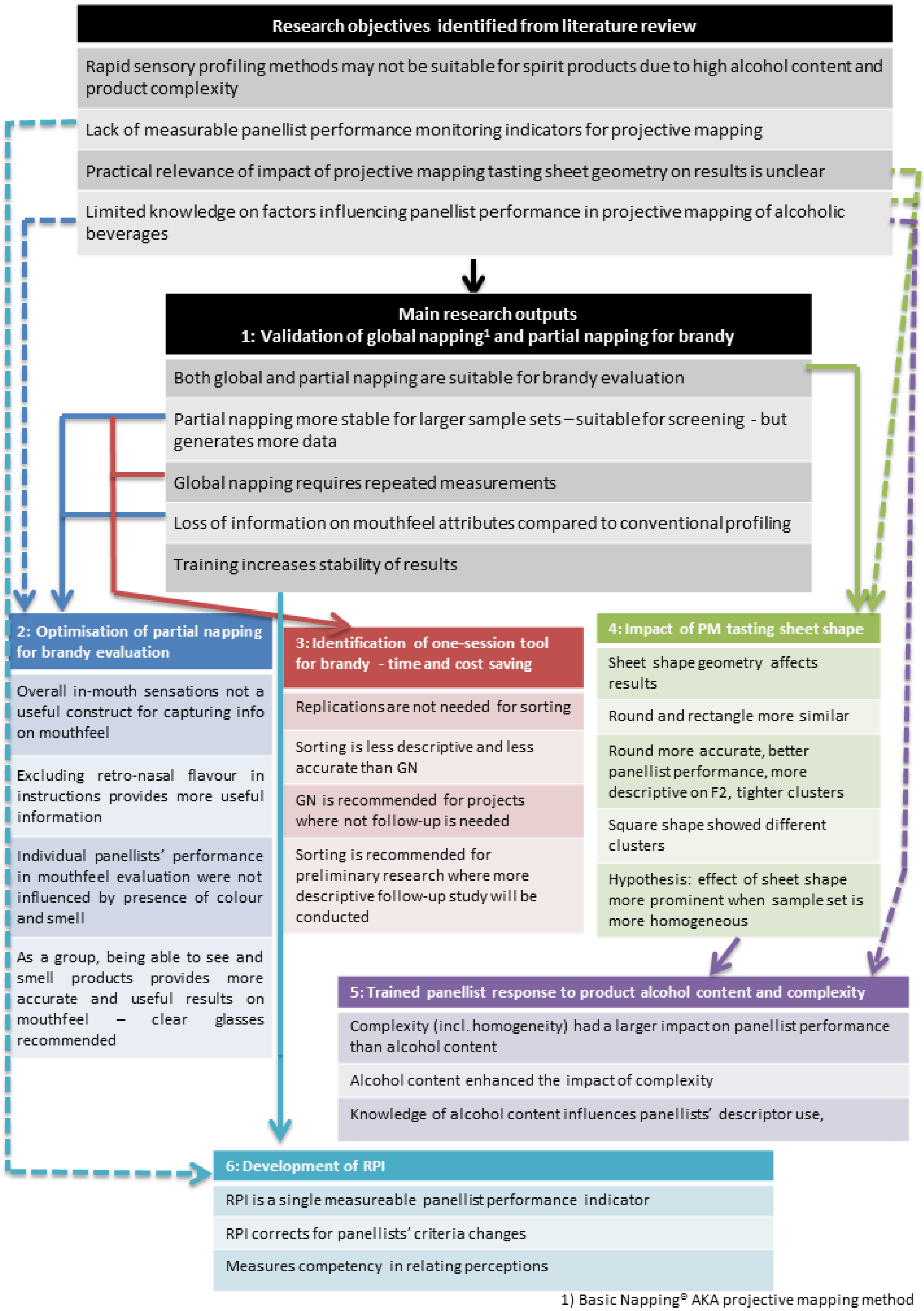

Figure 8.1 Feasibility assessment and optimisation of rapid sensory profiling methods for brandy evaluation: process flow and key findings.

The feasibility of projective mapping as a rapid sensory profiling method for brandy was evident at this point. Yet, there were still some uncertainties regarding the practical application of the method in general as well as to high alcohol beverages specifically. A general practical aspect was the influence of tasting sheet shape on projective mapping results. As of yet, this aspect has got very little attention, and tasting sheet shapes have not been standardised. From the two previous reports on this topic it was still unclear whether any changes to the overall product configuration due to variance in tasting sheet geometry had any practical implications (Dehlholm, 2013; Hopfer & Heymann, 2013). Although there was evidence that panellists treated the horizontal and vertical axes differently depending on tasting sheet shapes, there is limited evidence on whether their performance were affected. By comparing results from rectangular, square and round tasting sheets, this study confirmed in Chapter 6 that tasting sheet geometry affects projective mapping results (Figure 8.1, step 4) to some extent. Between round and rectangular sheets, the results were very similar, although tighter clusters, better accuracy and clearer descriptions of the second dimension of the product configuration. The first dimension of the square sheets overall product configuration had the strongest correlation with sample colour intensity and also delivered different sample clustering compared to the round and rectangular sheet shape. In terms of panellist performance, the majority of the panellists performed better using the round tasting sheet. It has previously been hypothesised that an elongated tasting sheet would guide panellists to bring out major product differences, while more subtle product differences could be obtained using a more symmetrical sheet shape with equal perpendicular bisectors where they are not given dimensional guidance. Besides the fact that the round tasting sheet explained the second factor better in terms of descriptor use, this effect was not obvious from the product configurations in this study. It further appears as if the benefit of using a more symmetrical tasting sheet to highlight subtle differences may be related to the homogeneity in the sample set.

Sample set homogeneity was addressed as part of an experiment investigating the role of complexity in panel performance in projective mapping in Chapter 7. Considering the complex volatile composition of brandy in the light of previous research showing that complexity can affect projective mapping results, the question was raised as to whether the difficulty in brandy evaluation was related to sample complexity, alcohol content of both. It was attempted to answer this question by comparing panellist performance in the projective mapping of sets of low and high alcohol commercial beverages of varying complexity (Figure 8.1, step 5). In this case complexity was defined as the homogeneity of the sample set as well as the perceived complexity of the product i.e. the number of nuances perceived. The results showed that complexity had a larger impact on panellist performance than alcohol content, but that high alcohol content enhanced the effect of complexity. A second question regarding the difficulty experienced by panellist in the sensory evaluation of brandy is whether knowing that a product is high in alcohol affects panellist performance. Knowing that a product is high in alcohol may create expectations of task difficulty, which have been reported to affect trained panellists’ motivation and task performance. The results from this study showed that any effect of the knowledge of alcohol content of a product on panellist performance varied across panellists. More importantly, knowledge of the high alcohol content of the samples changed the panel’s descriptor usage to include more alcohol-related words. It is hypothesised that the panellists may place more value on alcohol related words, knowing that they are evaluating high alcohol products.

Several studies have pointed out that projective mapping is better suited to trained panels and expert panels than consumer panels. The use of a trained panel warrants a suitable performance monitoring tool, especially considering that panellists task competency may increase with practice of the projective mapping method. Panellist performance monitoring in the projective mapping task has got very little attention in projective mapping studies and there was a lack of a single interpretable measure for the stability of panellist performance. A new panellist performance monitoring tool was developed, namely the Relative Performance Index (RPI). The RPI measure provides a single measure of the consistency of panellists' results over several sessions. Unlike the People Performance Index (PPI) that focusses on the positioning of two replicated samples, the RPI takes the positioning of all the samples into consideration. The measure is based on the optimal consensus between replicated sessions after procrustes data transformation. By correcting for variation in the panellists' assessments, a useful measure of the panellists' competency in relating their perceptions of the products is provided. This is very valuable considering that panellists have been reported to change their criteria over projective mapping sessions. This benefit was confirmed by the fact that higher RPI values can be obtained compared to RV coefficients when comparing individual panellists' results over several sessions. In contrast to the RV coefficient, the RPI measure can be applied to more than two sessions; considering the need for replicated measures for brandy evaluation, this can be especially useful. A suitable cut-off point for the interpretation of the RPI measure must still be established.

8.2 Practical applications and recommendations for future research

In summary, three useful rapid sensory profiling methods have been validated for brandy, each with their own recommended applications. Partial napping is best used where details are required of product dimensions that are not the most prominent differentiating factors; of the three methods tested it is the best approach for obtaining detailed mouthfeel evaluation. To obtain useful information on mouthfeel attributes, panellists must be instructed to focus strictly on this modality, but it is not necessary to use dark glasses or nose clips to eliminate other sensations. Partial napping can reliably be used to evaluate larger sample sets.

Global napping is useful for obtaining a quick overview of the most important differentiating factors in the product set, with relevant product descriptions. For larger sample sets, repeated measurements are highly recommended. On the other hand, sorting is recommended for studies where a larger sample set needs to be screened down to a representative subset for a follow-up study, provided that sample descriptions are not particularly important. Considering the negligible contribution of in-mouth sensations, it could be considered to perform these methods by nosing only, which may increase the potential upper limit for sample set size as the samples are not ingested. However, this upper limit must still be confirmed.

In terms of tasting sheet shape, there is not enough evidence that the effect of changing the sheet shape on overall product configurations has a significant *practical* impact. Panellists performed better on the round tasting sheet, and the round tasting sheet was more similar to the rectangular tasting sheet. Since round tasting sheets are less practical than rectangular tasting sheets, the latter shape is recommended for the moment. Regarding the theory that more symmetrical tasting sheet shapes could facilitate the elicitation of more subtle sample differentiations, it appears that, if this effect in fact exist, its practical application would be more useful in large and fairly homogeneous sample sets. The sample set in this project was perhaps too small to illustrate the potential use of a specific sheet shape to create a better fit with project objectives. Further studies are recommended and should include applications in sample sets varying in size and homogeneity. The length of the rectangular sheet shape should also be investigated in order to ease the choice between the two most commonly used rectangular sheet shapes namely 600 x 400 mm and 297 x 420 mm (A3).

The significant impact of both product complexity and alcohol content on panellist performance provides compelling evidence for the careful monitoring of projective mapping of brandy. It is highly recommended to include at least two duplicated samples in a product set to monitor the reliability of the results. For both global napping and sorting, there were sessions where the duplicated samples positioned very far apart on the overall consensus configuration. Including duplicated samples can be regarded as a safety net for ensuring valid and reliable results. Furthermore, training panellists on smaller sample sets can be useful to improve their internal consistency. If projective mapping is to be applied on a regular basis, it is recommended that the performance of the panellists are monitored using the RPI measure.

8.3 References

Dehlholm, C., Brockhoff, P.B., Meinert, L., Aaslung, M.D., & Bredie, W.L.P. (2012). Rapid descriptive sensory methods – comparison of Free Multiple Sorting, Partial Napping, Napping, Flash Profiling and conventional profiling. *Food Quality and Preference 26,* 267-277.

Dehlholm, C. (2013) *Descriptive sensory evaluations: Comparison and applicability of novel rapid methodologies.* PhD Thesis, ISBN 978-87-7611-592-0.

Hopfer, H., & Heymann, H. (2013). A summary of projective mapping observations – The effect of replicates and shape, and individual performance measurements. *Food Quality and Preference*, *28,* 164-181.

Jack, J. (2003). Development of guidelines for the preparation and handling of sensory samples in the Scotch whisky industry. *Journal of the Institute of Brewing, 109,* 114-119.

Pena Y Lilo M., Latrille E., Casaubon, G., Agosin, E., Bordeu E., & Martin, N. (2005). Comparison between odour and aroma profiles of Chilean Pisco spirit. *Food Quality and Preference, 16,* 59-70.

Pfeiffer, J.C. & Gilbert, C.C. (2008). Napping by Modality: a happy medium between analytical and holistic approaches. *Proceedings of Sensometrics Conference 2008,* Canada.

List of abbreviations used

ABV: Alcohol by volume
ANOVA: Analysis of variance
CA: Correspondence Analysis
CP: Conventional descriptive profiling
D_r: Distance ratio
GN: Global napping, also Napping®
GPA: General Procrustes Analysis
HMFA: Hierarchical Multiple Factor Analysis
MCA: Multiple Correspondence Analysis
MFA: Multiple Factor Analysis
PANOVA: Procrustes analysis of variance
PCA: Principal Component Analysis
PM: Projective mapping
PMFA: Procrustes Multiple Factor Analysis
PN: Partial napping
PPI: People Performance Index
RPI: Relative Performance Index
RSP: Rapid sensory profiling