Rapid sensory profiling methods for wine: Workflow optimisation for research and industry applications

by

Jeanne Brand

Dissertation presented for the degree of

Doctor of Philosophy (Wine Biotechnology)

at

Stellenbosch University

Institute for Wine Biotechnology, Department of Viticulture and Oenology

Faculty of AgriSciences

Supervisor: Dr H H Nieuwoudt

Co-supervisor: Prof M A Vivier

Co-supervisor: Prof D Valentin

Co-supervisor: Prof T Næs

April 2019

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 any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: April 2019

Summary

Descriptive sensory analysis techniques are widely used and trusted methodologies. Due to time and cost constraints, the demand for cost-effective methods for profiling is growing rapidly in food and beverage industries including the wine industry.

A number of rapid methods have been tested and validated for various food products. However, further work is needed to identify and address limitations of specific rapid methods, especially reference-based methods, when evaluating complex matrices such as wine. The majority of studies employed novice consumers or trained consumers as judges. The wine industry has an advantage over most food industries with: (1) product experts who can serve as judges and (2) having an extensive lexicon in the form of aroma wheels available that can be used as check-all-that-apply (CATA) questions.

The objective of this study was to identify cost-effective, rapid sensory methods that can be used for wine profiling by researchers and the wine industry alike. Furthermore, the study aimed to optimise the identified methods and to propose workflows that include sensory methods and statistical procedures suited for wine sensory analysis applications.

Four rapid methods were compared to descriptive analysis (DA). The methods tested were CATA, rate-all-that-apply (RATA), Napping, and sorting. Results obtained for the rapid sensory method and DA were similar. It can therefore be concluded that rapid methods are suitable for the sensory evaluation of wine. Industry professionals can therefore be used as sensory judges, and can use a pre-determined lists of attributes as verbalisation tools. CATA and sorting provided the highest quality profiles with the best discrimination between products. Sorting highlights similarities and differences whereas CATA provides more detailed descriptions. In addition, these two methods were found to be easier than rate-all-that-apply (RATA) and Napping to use.

Pivot profile (PP), a reference-based method, was validated against a CATA variant, namely frequency of attribute citation (FC). It was concluded that PP should be used with caution because the choice of pivot on the sensory space could have an influence. This method could, however, be useful when direct comparisons between samples are required, such as benchmarking.

In addition to sensory method development, a number of statistical procedures were also proposed to assist with the interpretation of rapid method data. A workflow to calculate drivers of quality and a strategy to calculate confidence ellipses for PP data were developed.

This study highlights the importance of selecting a fit-for-purpose method. The objective of the experiment being conducted, along with practical restrictions should be taken into account when deciding which method to use.

Opsomming

Beskrywende sensoriese evaluerings tegnieke word algemeen gebruik, en word as betroubare metodes erken. Weens tydsdruk en kostebeperkings is daar 'n toenemende aanvraag na vinniger sensoriese profileringsmetodes wat gebruik kan word in die voedsel- en drankbedrywe insluitende die wynbedryf.

Verskeie vinnige metodes ("rapid methods") is reeds getoets en gevalideer vir die sensoriese analise van 'n verskeidenheid voedselprodukte. Verdere navorsing is egter nodig om voordele en nadele van spesifieke metodes te identifiseer en aan te spreek. Dit is veral die geval wanneer verwysings-gebaseerde metodes gebruik word om komplekse matrikse soos wyn te evalueer. Die meeste studies wat reeds hieroor gedoen is, gebruik verbruikers as sensoriese beoordelaars. Die wynindustrie het 'n voordeel bo die meeste ander voedselindustrieë in die sin dat: (1) produkkenners as sensoriese beoordelaars kan dien en (2) omvattende beskrywings van die sensoriese eienskappe van wyn reeds in die vorm van, onder andere, geurwiele beskikbaar is. Laasgenoemde kan as merk-alles-wat-van-toepassing-is ("check-all-that-apply" of CATA) vraelyste gebruik word.

Die doel van hierdie studie was om koste-effektiewe, vinnige sensoriese metodes te identifiseer wat vir wynprofilering deur navorsers en die breë wynbedryf gebruik kan word. Dit het ook ten doel gehad om laasgeneomde metodes te optimaliseer en 'n getoetste werksvloei voor te stel, wat sensoriese metodes en statistiese prosedures insluit. Vier vinnige metodes is vergelyk met beskrywende sensoriese analise ("descriptive analysis" of DA). Die metodes is merk-alles-van-toepassing ("check-all-that-apply" of CATA), gradeer-alles-van-toepassing ("rate-all-that-apply" of RATA), Napping, 'n spesifieke variasie van projeksiekartering ("projective mapping" of PM), en sortering. Die vinnige sensoriese metodes en beskrywende analise het soortgelyke resultate opgelewer. Vinnige metodes blyk dus geskik te wees vir die sensoriese analise van wyn wanneer vooraf-opgestelde lyste met sensoriese eienskappe gebruik word as verbaliseringstap en produkkenners uit die industrie as beoordelaars optree.

Die hoogste kwaliteit profiele, wat die beste tussen produkte kon onderskei is met behulp van CATA en sortering verkry. Sortering het ooreenkomste en verskille tussen produkte uitgelig, terwyl CATA meer gedetailleerde beskrywings opgelewer het. Volgens die paneellede is beide hierdie metodes makliker om te gebruik as RATA of Napping.

Draaipuntprofilering ("pivot profile", PP), 'n verwysingsgebasseerde metode, is gevalideer en vergelyk met 'n variasie van die CATA metode, naamlik frekwensie van eienskap-aanhaling ("frequency of attribute citation", FC). Weens die invloed van die keuse van die draaipuntmonster op die sensorieseruimte met PP, moet hierdie metode met versigtigheid gebruik word. Hierdie metode kan wel waardevol wees wanneer 'n direkte vergelyking tussen produkte verlang word, byvoorbeeld wanneer een produk teen 'n ander een wat as maatstaf dien, vergelyk word.

Benewens die ontwikkeling van sensoriese metodes is statistiese prosedures voorgestel om die interpretasie van die data, wat met die vinnige metodes bekom is, te hanteer en beter te visualiseer. ‘n Werksvloei om kwaliteitsdrywers te identifiseer, sowel as ‘n strategie om vertrouensellipsoïede vir draaipuntprofilering te bereken, is ook ontwikkel. Die studie onderstreep hoe belangrikheid van die metodekeuse wanneer ‘n vinnige sensoriese metode vir wynprofilering gebruik word. Die metode wat gekies word moet die verwagte uitkomste van die studie ondersteun en die praktiese beperkings in ag neem.

This dissertation is dedicated to my husband, Jacques Brand,

my two beautiful children, Gretha and Marius,

the rest of my family

and

every reader who finds sensory and/or wine science fascinating.

Biographical sketch

Jeanne Brand (maiden name Treurnicht) was born in Pretoria, South Africa on 9 June 1980. She attended Stellenbosch Primary. At the age of ten, her parents moved to the Southern Free State where she attended several schools in the area and matriculated from Hopetown High in the Northern Cape in 1998. She obtained a bachelor's degree in chemistry in 2003 and an HonsBSc-degree in Wine Biotechnology in 2004 both at Stellenbosch University. She pursued a career in quality control in the wine industry and completed her MSc in Wine Biotechnology part-time in 2010. She is currently employed by the Department of Oenology and Viticulture as sensory laboratory manager and the coordinator of the Institute for Grape and Wine Sciences' Sensory Platform.

Acknowledgements

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

- **Dr Helene Nieuwoudt**, Institute for Wine Biotechnology (IWBT), Department of Oenology and Viticulture (DVO), Stellenbosch University, South Africa, for guidance, continuous support, inspiration and for helping me improve my ability to apply critical thinking.
- **Prof Dominique Valentin**, University of Burgundy, France, for teaching me about a completely new field in sensory methodology that I only started to explore during my PhD studies.
- **Prof Tormod Næs**, Nofima, Norway, for assistance with the formulation of new ideas, the initiative to conduct research concerning rapid sensory methods at Stellenbosch University, invaluable help with experimental design, statistical analysis approaches and results interpretation.
- **Prof Melané Vivier**, for insightful critical evaluation and support.
- **Prof Martin Kidd**, for helping with statistical analysis on a daily basis both from a practical and academic perspective.
- **Valeria Panzeri and Carla Weightman**, for technical and moral support as colleagues and friends.
- **Stellenbosch University sensory panels and industry professionals**, for their dedication, passion and hard work as sensory judges "…*without an instrument there will be no measurement…*" – *anonymous.*
- **Winetech** and **THRIP** for funding projects IWBT 13/02 and Tp14080888680
- **The Department of Science and Technology** for funding the "Cross-cultural wine quality perception profiling" project.
- **My colleagues and the postgraduate students at IWBT and DVO** for helping out, sometimes "last minute", when needed.
- **My family, especially my dear husband and two children** for support, understanding and walking this path with me every step of the way.

Preface

This dissertation is presented as a compilation of six chapters. Each chapter is introduced separately and written according to the style of the journal Food Quality and Preference. In order to keep to the style of the journal, the Latin abbreviation, “et al.” was not written in italics.

Chapter 1 **General Introduction and project aims**

Chapter 2 **Literature review**
The application of rapid methods to wine sensory evaluation: A Review

Chapter 3 **Research results**
In search of suitable rapid sensory methods for wine profiling using industry professionals: A comparison of Free Sorting, Napping, Check-All-That-Apply and Rate-All-That-Apply to Descriptive Analysis

Chapter 4 **Research results**
Validating Pivot© Profile by means of comparison to Frequency of attribute Citation: Analysing complex products with trained assessors

Chapter 5 **Research results**
Sorting in Combination with Quality Scoring: A Tool for Industry Professionals to Identify Drivers of Wine Quality Rapidly

Chapter 6 **General discussion and conclusions**

The chapters in this dissertation were written independently as scientific papers and submitted to scientific journals prior to the compilation of this document. Overlap in certain sections of the literature review chapter and the introductions of the research results chapters was unavoidable.

Some of the sensory methods, discussed and cited, are trademarked. Trademark signs were only included upon the first mention of the method in chapter 1 and not in the rest of the document.

List of Outputs

The work presented in this dissertation was submitted for publication to peer review scientific journals, presented at scientific conferences and communicated through publication of popular articles.

Scientific articles

Brand, J., Kidd, M., Van Antwerpen, L., Valentin, D., Næs, T., & Nieuwoudt, H.H. (2018). Sorting in Combination with Quality Scoring : A Tool for Industry Professionals to Identify Drivers of Wine Quality Rapidly, *South African Journal of Enology and Viticulture 39*, 163–175. (Chapter 5)

Brand, J., Valentin, D., Kidd, M., Vivier, M.A., Næs, T., & Nieuwoudt, H.H. Validating Pivot Profile by means of comparison to Frequency of attribute Citation: analysing complex products with trained assessors. (Chapter 4)

Submitted to *Food Quality and Preference* in August 2018, currently under review.

Brand, J., Næs, T., Kidd, M., Vivier, M.A., Valentin, D., & Nieuwoudt, H.H. In search of suitable rapid sensory methods for wine profiling using industry professionals: A comparison of Free Sorting, Napping, Check-All-That-Apply and Rate-All-That-Apply to descriptive analysis. (Chapter 3)

To be submitted to the *South African Journal of Enology and Viticulture* in December 2018.

Conference participation

Workshops

Brand, J., Louw, L., Van der Merwe, X., & Nieuwoudt, H.H. (2018). Rapid sensory profiling solutions for industry applications. *41st International SASEV / WINETECH conference*, 2–4 October 2018, Somerset West, South Africa

Oral presentations

Brand, J., Valentin, D., Vivier, M.A., Næs, T., & Nieuwoudt, H.H. (2015). Comparison of rapid sensory techniques for white wine profiling. *AfroSense Conference*, 23–26 November 2015, Stellenbosch, South Africa

Brand, J., Valentin, D., Næs, T., & Nieuwoudt, H.H., (2014). Sustainable sensory methods for profiling of wine: Pros and Cons of Rapid Methods. *36th International SASEV / WINETECH conference*, 12–14 November 2014, Somerset West, South Africa

Posters

Brand, J., Valentin, D., Vivier, M.A., Næs, T., Nieuwoudt, H.H. (2016). Comparing two frequency based sensory profiling methods using a trained panel: Pivot profile & frequency of attribute citation. *7th European Conference on Sensory and Consumer Research, EuroSense*, 10–14 September 2016, Dijon France

Brand, J., Van Antwerpen, L., & Nieuwoudt, H.H. (2016). Sorting in combination with quality scoring as a tool to identify drivers of wine quality. *38th International SASEV/WINETECH conference,* 23-25 August 2016, Somerset West, South Africa

Brand, J., Valetin, D., Vivier, M. A., Næs, T., & Nieuwoudt, H.H. (2015). Testing the robustness of Pivot profile when profiling wine with a trained panel. *AfroSense Conference*, 23–26 November 2015, Stellenbosch, South Africa

Brand, J, Van Antwerpen, L., & Nieuwoudt, H.H. (2015). Suitability of the sorting task in combination with quality scoring to identify drivers of quality using wine industry professionals as sensory judges. *AfroSense Conference*, 23–26 November 2015, Stellenbosch, South Africa

Popular articles

Brand, J., & Nieuwoudt, H.H. (2016). Sensory evaluation of wine (Part 2): Sorting – a fast and simple method to describe sensory differences and similarities between wines. Wineland (March) https://www.wineland.co.za/sensory-evaluation-of-wine-part-2/

Brand, J., & Nieuwoudt, H.H. (2016). Sensory evaluation of wine (Part 3): Projective mapping and Napping. Wineland (April) https://www.wineland.co.za/sensory-evaluation-of-wine-part-3/

Brand, J., & Nieuwoudt, H.H. (2016). Sensory evaluation of wine (Part 4): Check-all-that-apply (CATA) – profiling wine with multiple choice questions. Wineland (May) https://www.wineland.co.za/sensory-evaluation-of-wine-part-4/

Awards

Best poster presentation, *AfroSense 2015*, 23–26 November 2015, Stellenbosch, South Africa

Best poster, *38th International SASEV/WINETECH conference*, 23-25 August 2016, Somerset West, South Africa

Table of Contents

Chapter 1

General introduction and project aims

1. General introduction and project aims

1.1 Introduction

Sensory evaluation is used as a tool in the food and beverage industry throughout and after the production process to assist with the improvement and development of new products or brands, quality control and finally marketing and advertising (Lawless & Heymann, 2010).

A number of sensory methodologies are available in the literature to describe and profile the sensory characteristic of foodstuffs and beverages. Trademarked methods such as Flavour profile™ (Cairncross & Sjostrom, 1950), Quantitative Descriptive Analysis (QDA™, Stone et al., 1974) and the Spectrum method™ (Munoz & Civille, 1992), as well as non-trademarked methods, Texture profile (Brandt et al., 1963) and Quantitative Flavour Profiling (Stampanoni, 1993) are available. The most trusted, frequently used and adapted method from these descriptive profiling techniques is QDA™ (Valentin et al., 2012; Lestringant et al., 2019). Although QDA™ provides excellent results, the method is frequently adapted to allow more flexibility. These adapted methods are generally referred to as generic or conventional descriptive analysis (DA) in the literature. Conventional DA is conducted with eight to 15 judges rating the intensity of up to 20 sensory attributes on a line scale and can be described as a three-step process.

During the first step, familiarisation with the products and development of the lexicon used to accurately describe the products is achieved. In order to describe the product space thoroughly and discriminate efficiently between different samples, judges are exposed to a wide variety of different samples spanning the sensory space. In many cases, the samples to be analysed are used during the training process.

The second step involves standardisation and alignment of the panel using reference standards or definitions describing the sensory attributes in the lexicon generated during the first step. Lastly, as the third step, attributes are rated for intensity on a line scale and panel performance is tested. Testing the panel performance involves evaluating the repeatability, ability to discriminate between samples for each judge as well as the consensus amongst the different judges (Lawless & Heymann, 2010).

DA is a trusted methodology that produced accurate high-quality results for various foodstuffs including complex products such as wine (Campo et al., 2010; Heymann & Hopfer, 2013; Sokolowsky et al., 2015; Lestringant et al., 2019). Due to the difficulty of aligning a panel to rate intensity, DA has the drawback that it is time consuming, taking in some cases four to six weeks to profile products, and therefore, costly to perform. In addition, the vocabulary generated is specific for the sample set analysed and training has to be conducted for every different sample set to ensure that the entire sensory space is spanned and fully described. When complex products are evaluated, using a small number of attributes, a loss of information can occur explaining less variability within the data set than when a larger amount of attributes are used (Campo et al., 2010).

In order to address these issues, alternatives for DA were proposed. A class of methods, referred to as rapid methods, divided into three categories, namely verbal-based, similarity-based and reference-based are amongst these alternatives. Check-all-that-apply (CATA, Adams et al., 2007) is the most widely used verbal-based method. Since the introduction of CATA, variants such as pick-*K* attributes (Valentin et al., 2012), where the *K* most dominant attributes are selected to describe the sensory properties of the samples, emerged. Pick-*K* attributes is known as frequency of attribute citation (FC) when performed by a trained panel (Campo et al., 2008). When verbal methods are used the judges describe the products directly. The results obtained are thus dependent on the analytical abilities and verbal skills of judges. In addition, samples are described one at a time by means of a monadic presentation order. It is, therefore, not possible to take the properties of the rest of the samples in the set into consideration when profiling a specific sample.

Similarity-based methods follow an intuitive approach as a first step. Products are grouped or organised according to similarity first. Description of the sensory properties to explain the similarities between products follows as the second step. All of the samples are presented simultaneously, which gives the judge an idea of the entire sensory space while describing the samples. However, this also results in a limitation of the number of samples that can be evaluated and aggregation of the data is not possible. Sorting (Chollet et al., 2011) and projective mapping (PM, Risvik et al., 1994), including Napping® (Pagès, 2003), fall within this category.

Reference-based methods can provide solutions to some of the previously mentioned shortcomings of verbal-based and similarity-based methods, for example, the limited size of sample sets when performing similarity-based methods. When a large sample set has to be analysed, multiple sessions keeping reference standards constant can address this issue. Polarised sensory positioning (PSP, Teillet et al., 2010), pivot profile© (PP, Thuillier et al., 2015) and polarised projective mapping (PPM, Ares et al., 2013) fall within this category (Valentin et al., 2012; Varela & Ares, 2012). When reference-based methods are conducted a comparative approach is followed. Products are compared to one to three reference samples, also called pivots or poles, instead of to each product in the sample set.

The verbalisation step in these methods, the first step for verbal-based methods and the second step for similarity- and most reference-based methods can be conducted choosing one of two strategies. The first and most commonly used strategy is to generate sensory attributes through free description, where panellists rely on previous experience and memory. The second strategy, to use a predetermined list, is gaining popularity since the statistical analysis of the data is simpler and the task less tedious for sensory judges than free description.

All these rapid methods have the advantage over DA that training is not required and therefore, are faster to perform and more cost-effective. A drawback that most of these methods have is that frequency counts, nominal data, is obtained as opposed to intensity ratings, continuous data (Dooley et al., 2010). From a statistical viewpoint, a larger number of judges might be required for rapid methods than when performing DA. However, these methods have successfully been used to profile complex products such as wine using naïve consumers, formally trained judges and product experts / industry professionals. The majority of the studies involving rapid sensory methodologies employed consumers as judges (for a review consult Valentin et al., 2012; Varela & Ares, 2012).

It has been noted that industry professionals provide a similar but more accurate analytical description of products using rapid methods than consumers (Ballester et al., 2008; Botha, 2015). In addition, Louw et al. (2013) used trained panellist to describe brandies using Napping, Crous (2016) used PSP to discriminate between Chenin Blanc wines and Lelièvre-Desmas et al. (2017) to profile beers using PP.

Furthermore, it is important to note that the different methods have different disadvantages and restrictions. Therefore, methods should be selected to provide fit-for-purpose solutions aligned with the objective of the experiment. In the studies where rapid methods were compared for their suitability to profile wine, only two or three methods were compared to each other or to DA. Only a few comparative studies were conducted on wine using industry professionals as judges (Perrin et al., 2008; Johnson et al., 2013; Torri et al., 2013; Vidal et al., 2018).

It has, therefore, been shown that rapid methods have potential as sensory analysis techniques for complex product evaluation. However, there is no study to date that: (1) compared the most frequently used rapid methods to each other, (2) kept the matrix and panel constant, (3) critically evaluated the difference and similarities between methods keeping fit-for-purpose applications in mind and (4) investigated practical solutions for the sensory analysis of wine.

Furthermore, the need exists to identify and adapt rapid sensory methods using the reliable resources already available within the wine industry. Industry professionals can be used as sensory judges since they establish a common language through work experience, gained from tasting on various industry panels such as competition and certification panels. Lexicon developed over decades, e.g. aroma wheels, can serve as pre-determined lists for the descriptive steps while conducting rapid sensory analysis. The advantage of using industry professionals in combination with the existing lexicon is that a common language is used that is understood by most judges (Ballester et al., 2008; Torri et al., 2013) and giving them the option to add to the list when necessary will continuously update the lexicon in a formal scientific way. Additionally, the statistical handling of the data is easier and faster to conduct and less biased in the sense that the interpretation, coding and combination of terms by the sensory analyst are not as intensive as when free description is allowed as the descriptive step.

Testing these methods using trained panels is also a necessity since industry professionals might not always be available when analyses are required, e.g. during harvest time, it is, therefore, necessary to also test these methodologies using trained panels.

1.2 Project aims

This research project was conducted to provide the wine industry and research centres with information to develop and optimise existing sensory methodology.

The main aim was to evaluate and compare rapid sensory methods to test their suitability for wine profiling using industry professionals and trained panels to obtain analytical sensory profiling data.

Within the framework of this aim, the following specific research objectives were formulated:

1. To identify rapid methods suitable for profiling of wine, a complex product, where a list of terms can be used as descriptive step and industry professionals (experts) as judges.
2. To validate a reference-based rapid method, PP, for profiling of wine matrices with different within-set variability by:
    - Testing the stability of the sensory space when changing the reference sample, the pivot.
    - Comparing PP to a well-established and trusted verbal-based sensory method, frequency of attribute citation (FC), a CATA variant.
3. To propose a workflow with a rapid method as profiling tool to determine drivers of quality in a single sensory evaluation session by:
    - Testing the suitability of sorting in combination with quality scoring using industry professionals as sensory judges.

- Determining drivers of quality by means of statistical analysis and inspection of the multivariate sensory map on which the sensory attributes and quality scores were projected.

1.3 References

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. Poster presented at *The 7th Pangborn Sensory Science Symposium*. 16 August 2007, Minneapolis, USA. Delegate Manual, pp. 12–418.

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.

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

Botha, A. (2015). The use of different oak products during the fermentation and ageing of Chenin Blanc : sensory properties, perceived quality, and consumer preference. MSc thesis, Stellenbosch University. http://hdl.handle.net/10019.1/98006

Brandt, M.A., Skinner, E.A., & Coleman, J.A. (1963). Texture profile method. *Journal of Food Science, 28*, 404–409.

Cairncross, S., & Sjostrom, I. (1950). Flavor profile. A new approach to flavor problems. *Food Technology, 4*, 308–311.

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.

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.

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.

Crous, R. (2016). The sensory characterisation of old - vine Chenin blanc wine : an exploratory study of the dimensions of quality. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/100114

Dooley, L., Lee, Y. S., & 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.

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.

Lawless, H. T., & Heymann, H. (2010). *Sensory evaluation of food. Principles and practices*. (2nd ed.). New York, NY, USA: Springer.

Lelièvre-Desmas, M., Valentin, D., & Chollet, S. (2017). Pivot profile method : What is the influence of the pivot and product space? *Food Quality and Preference, 61*, 6–14.

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.

Munoz, A.M., & Civille, G.V. (1992). The spectrum descriptive analysis method. In: *Manual on Descriptive Analysis Testing for Sensory Evaluation* (edited by C. Hootman). West Conshohocken, PA: ASTM Manual Series MNL 13.

Pagès, J. (2003). Recueil direct de distances sensorielles: application à l'évaluation de dix vins blancs de Val de Loire. *Science des Aliments, 23*, 679–888.

Perrin, L., Symoneaux, R., Maître, I., Asselin, C., Jourjon, F., & Pagès, 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.

Risvik, E., 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.

Stampanoni, C.R. (1993). The quantitative flavor profiling technique. Perfumer & Flavorist, 18, 19–24.

Stone, H., & Sidel, J. L. (2004). *Sensory evaluation practices*. San Diego, USA: Elsevier, Academic, Elsevier.

Stone, H., Sidel, J., Oliviers, S., Woosley, A., & Singleton, R. C. (1974). Sensory evaluation by quantitative descriptive analysis. *Food Technology, 52,* 24–28.

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.

Thuillier, B., Valentin, D., Marchal, R., & Dacremont, C. (2015). Pivot© profile: A new descriptive method based on free description. *Food Quality and Preference, 42*, 66–77.

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.

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*, 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.

Vidal, L., Ares, G., Hedderley, D. I., Meyners, M., & Jaeger, S. R. (2018). Comparison of rate-all-that-apply (RATA) and check-all-that-apply (CATA) questions across seven consumer studies. *Food Quality and Preference, 67*, 49–58.

Chapter 2

Literature review

The application of rapid methods to wine sensory evaluation: A Review

2. Literature review

The application of rapid methods to wine sensory evaluation: A Review

Abstract

Sensory evaluation of alcoholic beverages, including wine, is essential during product development, production and marketing processes. A radical change in the focus of sensory method development research can be seen in the literature published in the past 20 years. Alternative, fast and cost-effective methods have been proposed, to complement conventional descriptive sensory methodologies and consumer liking tests. Panels with different levels of training and expertise including consumers, trained panels and industry professionals can perform these methods. Thus, consumer and expert sensory profiles, highlighting sensory properties and perceptions not captured with conventional descriptive analysis techniques, can be obtained. This category of methods is known as rapid sensory methods. In this review, the application and modification of these methods in the context of sensory evaluation of wine and alcoholic beverages are discussed. This review therefore complements previous reviews by Valentin et al. (2012) and Varela & Ares (2012) that focused on rapid sensory methods as applied in the food industry by: (1) incorporating the latest rapid method research specific to wine and alcoholic beverages and (2) discussing the application of rapid sensory methods within the alcoholic beverage industry.

Keywords: rapid sensory analysis, wine profiling, sorting, projective mapping, Napping, pivot profile, polarised sensory positioning, check-all-that-apply, rate-all-that-apply

2.1 Introduction

Measuring the perception of food through the senses specifically sight, smell, taste and touch, is crucial to understand the intrinsic and extrinsic properties of foodstuffs to produce products acceptable to consumers. The measuring and interpretation of human perception of food, in a systematic way, has emerged in the 1950s, due to industry demand (Pangborn, 1964). It was developed into a scientific field, namely sensory science, in the 1970s by researchers such as Pangborn (Lawless & Heymann, 2010). The development of sensory methods has been an ongoing process driven by industry demand ever since.

The first cycle of method development was driven by the notion of providing industry with formally validated, sensory methods where the quality of the data is monitored by means of testing various parameters for example panel performance and repeatability. These validated methods, including quantitative descriptive analysis (QDA™), where a trained panel is used and data analysed by means of formal statistical methods, are trusted and well-established (Lawless & Heymann, 2010). However, due to extensive panel training, these methods are expensive and can take up to six weeks to perform. Hence, one of the current focus areas of sensory method development is providing industry with cost-effective fast alternatives known as rapid sensory methods.

2.2 Rapid sensory evaluation methodologies

2.2.1 Categorisation of rapid sensory methods

Rapid sensory methods were categorised by Valentin et al. (2012) according to the psychological processes required from the sensory judges while evaluating products. Three categories were proposed namely verbal-based, similarity-based and reference-based. This classification system focuses on the cognitive process involved when sensory judges perform the main task. This task is in most cases the first step, responsible for the sample configuration of the sensory map.

A more detailed classification system where rapid methods are considered as combinations of different blocks or tasks was introduced by Bécue-Bertaut (2014). This classification system is based on the same principles and cognitive processes as those proposed by Valentin et al. (2012) with the difference that every task is categorised and not only the main task. Therefore, each task associated with a different cognitive process or generating a separate set of data is viewed as a separate block. This approach allows for a structured and detailed explanation of the statistical analysis techniques applied to rapid sensory method data. In addition, it highlights new possibilities for using different combinations of the subsequent building blocks to design fit-for-purpose methodologies (Table 1).

TABLE 1
Summary of the classification of rapid sensory methods.

Sensory method	Methodological building blocks					Reference
	Reference	Similarity	Verbal	Rating	Ranking	
DA[d]			✓[a]	✓		Lawless and Heymann, 2010
FP[e]			✓[a]		✓	Dairou and Sieffermann, 2002
CATA[f]			✓[a]			Adams et al., 2007
RATA[g]			✓[a]	✓		Ares et al., 2014; Reinbach et al., 2014
Sorting		✓[a]	Labeling[b]			Lawless et al., 1995; Chollet et al., 2011
Napping & PM[h]		✓[a]	UFP[b,c]			Pagès (2003, 2005); Risvik et al. (1994,1997)
PSP[i]	✓[a]	✓				Teillet et al., 2010
PPM[j]	✓[a]	✓	UFP[b,c]			Ares et al., 2013
PP[k]	✓[a]		✓		✓	Thuillier et al., 2015

[a]The task or block that determines the main classification of the method.
[b]Technique commonly used during a second supplementary step to obtain sensory descriptors.
Acronyms used for rapid sensory methods: [c]Ultra flash profile, [d]Descriptive analysis; [e]Flash profile ;[f]Check-all-that-apply ;[g]Rate-all-that-apply; [h]Projective mapping; [i]Polarised sensory positioning; [j]Polarised projective mapping; [k]Pivot profile.

2.2.2 Verbal-based methods

Verbal-based methods rely on the ability of sensory judges to express their perception using words, phrases (Valentin et al., 2012) or emoji's (Jaeger et al., 2018). A list of sensory attributes can be pre-determined and provided by the experimenter, or sensory judges can be asked to provide the terms themselves. Free comments (Lawrence et al., 2013), free listing (Hough & Ferraris, 2010), free choice profiling (FCP, Williams & Langron, 1984), repertory grid (RP, Veinand et al., 2011), flash profiling (FP, Dairou & Siefferman, 2002), check-all-that-apply (CATA, Adams et al., 2007; Lancaster & Foley, 2007) and open-ended questions are examples. These methods were used in sensory science to profile products since the late 1900s.

2.2.2.1 Free choice profiling (FCP) and repertory grid (RG)

FCP (Williams & Langron, 1984) and RP (Veinand et al., 2011) were among the first rapid methods proposed and tested, where the classical descriptive analysis' training step was bypassed. When performing FCP, sensory judges develop their own vocabulary to describe the samples and rate the intensities of the attributes on line scales. The data is then analysed by means of generalised procrustes analysis (GPA) since judges use different attributes. A similar approach is followed for RG with the difference that attributes are generated by providing triads of samples to judges (Kelly, 1955; Veinand et al., 2011) when the vocabulary is generated. When RG is performed judges are asked to explain in their own words how two of the three products in a triad differ or are alike when compared to the third. As the second step intensities for the vocabulary generated in step one are rated. The main difficulty with FCP and RP is the fact that sensory judges are asked to rate the intensity of attributes on a line scale without prior training.

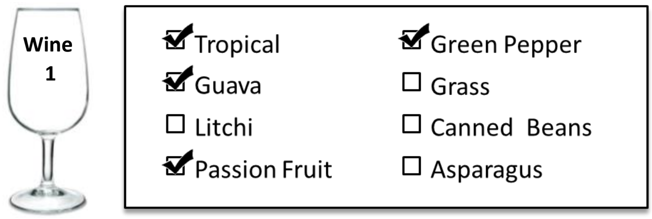
2.2.2.2 Free description, free listing, free comments and CATA

Free description, free comments or free listing of the sensory properties of products are frequently used to supplement liking data from consumers (Bécue-Bertaut, 2014). In addition it has been used in research for the description of products by industry professionals (Lawrence et al., 2013). Processing the data obtained from these methods is time consuming and prone to biases. These biases can occur during the semantic interpretation of the results by the experimenter since the experimenter together with colleagues and/or industry professionals reduce the number of attributes and not the sensory judges who evaluated the products. In order to simplify data processing and narrow down the variation in language used, CATA has become a popular technique for rapid profiling of food and lately wine (Etaio et al., 2010; Ares et al., 2015 ;Vidal et al., 2015, 2017, 2018; Lezaeta et al., 2017, 2018; Corsi et al., 2017; Alencar et al., 2018; Coste et al., 2018).

CATA questions consist of a list of words or phrases representing sensory attributes or terms related to emotions and product acceptability (Fig. 1). Sensory judges choose terms from the list to describe the products. Samples are presented according to a randomised monadic serving design where judges get samples in a different order one-at-a-time (Adams et al., 2007; Jaeger et al., 2018).

The most challenging aspect of CATA is to choose the specific terms and deciding on the number of terms when constructing the list. Lists can be constructed during focus group sessions or from literature and previous studies. CATA data are most frequently analysed by means of correspondence analysis (CA), multiple correspondence analysis (MCA) and multiple factor analysis (MFA, Ares et al., 2011a,b; Valentin et al., 2012).

An extension of CATA was proposed where the selected attributes' intensities are rated. Ares et al. (2014) called this method rate-all-that-apply, using a 3-point scale with "low", "medium" and "high" (RATA) and Reinbach et al. (2014) called it CATA with intensity rating using a 15-point scale ranging from "very weak" to "very strong". In addition, 5-point and 7-point scales have also been used (Ares et al., 2014; Franco-Luesma et al., 2016).



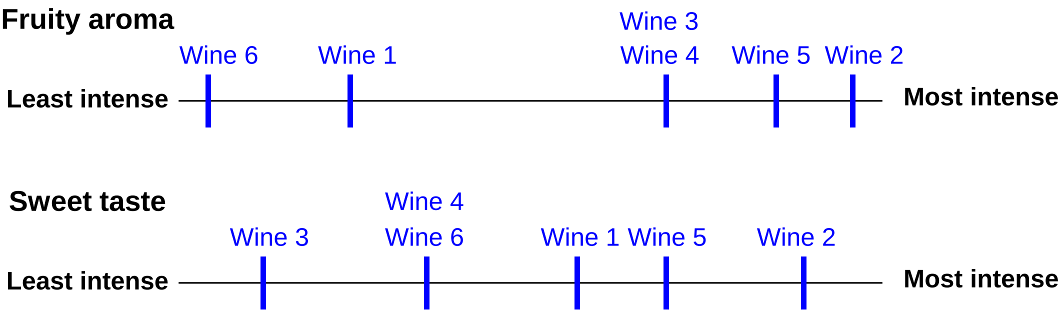
Judge 1	Tropical	Guava	Litchi	Green Pepper
Wine 1	**1**	**1**	**0**	**1**
Wine 2	1	0	1	0
Wine 3	1	0	0	1
Wine 4	0	0	0	1

Sum across Judges	Tropical	Guava	Litchi	Green Pepper
Wine 1	**21**	**15**	**0**	**18**
Wine 2	11	0	8	0
Wine 3	15	8	0	7
Wine 4	5	0	0	22

Fig. 1. Schematic presentation of a check-all-that-apply (CATA) question and the data capturing process where a contingency table is constructed with the sum of the attribute citations over all the judges for every sample for every attribute. Correspondence analysis (CA) can be performed on the contingency table (right-hand data table) or multi-block analysis, e.g. multiple factor analysis (MFA) can be performed on the individual data tables (left-hand data table).

2.2.2.3 Flash profile (FP)

FP is a combination of two methods; free choice profiling (FCP), also known as free description, and ranking (Dairou & Sieffermann, 2002). This method is used to rapidly profile products highlighting the most prominent attributes by ranking them (Fig. 2). When FCP is performed, the samples are presented simultaneously during a two-step process with a break between steps. Sensory judges are asked to evaluate the samples and generate a list of descriptors that will be sufficient to describe them, and discriminate between them. The descriptors from all the judges are pooled. Judges then select the most appropriate descriptors from the list and rank the samples from low to high for each attribute. The individual sensory judges' rank data are captured (Fig. 2). Due to the ranking step this method is not suitable for analysing large numbers of samples since short-term memory problems might compromise the quality of the data obtained (Valentin et al., 2012).



Judge 1				
Wines	Fruity (A1)	…	Sweet (A6)	…
Wine 1	2		4	
Wine 2	6		6	
Wine 3	3.5		1	
Wine 4	3.5		2.5	
Wine 5	5		5	
Wine 6	1		2.5	

	Judge 1			Judge 2				Judge 3			
Wines	A1	…	A6	A1	A2	A3	…	A1	A2	A3	…
Wine 1	2		4								
Wine 2	6		6								
Wine 3	3.5		1								
Wine 4	3.5		2.5								
Wine 5	5		5								
Wine 6	1		2.5								

Fig. 2. Schematic presentation of flash profile (FP), where FP is a combination of free choice profiling (FCP) and ranking. Rank data are collected after which generalised procrustus analysis (GPA) is performed with the data from the individual sensory judges as separate data tables.

2.2.3 Similarity-based methods

Similarity-based methods used for profiling of food products consist of a two-step process. The first step is to evaluate the entire samples set holistically to identify similarities and differences between samples. The second step is to describe the differences and similarities between the samples using sensory attributes either from a list or from memory. The second step is, therefore, a verbal-based method supplementing the similarity-based method. Sorting (Lawless, 1995; Chollet et al., 2011) and projective mapping (PM, Risvik et al., 1994) with its restricted version called Napping by Pagès (2003) fall into this category.

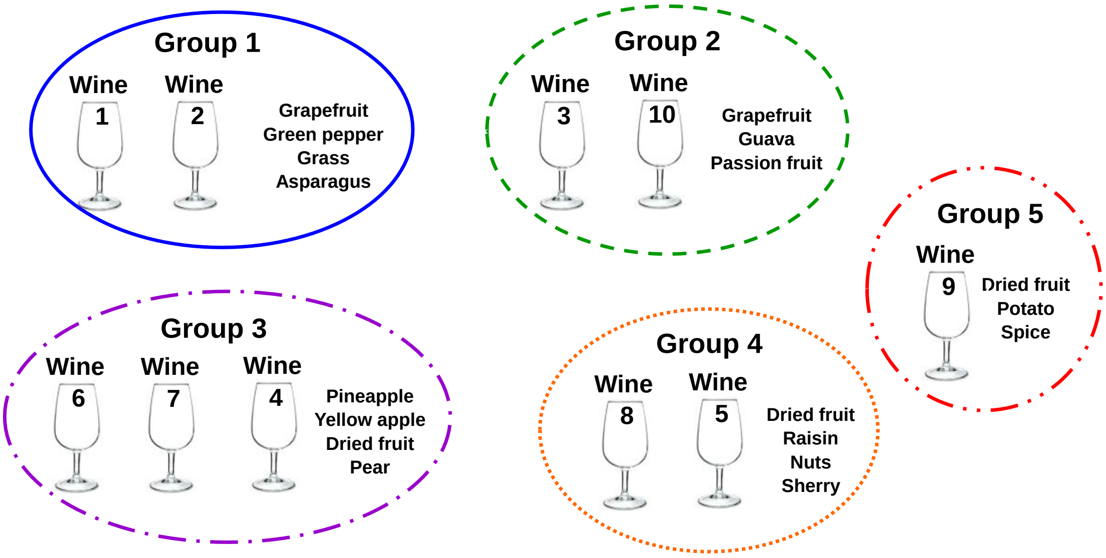
2.2.3.1 Sorting

Sorting is an intuitive task performed during everyday life by people while organising and interpreting their environment and is, therefore, regarded as an easy task (Qannari et al., 2010; Chollet et al., 2011). When sorting is conducted sensory judges receive all the samples simultaneously in a random order and are asked to group samples according to similarity (Fig. 3).

When judges are allowed to use their own criteria to sort the samples into as many groups as they deem necessary, it is known as the free sorting task (FST). Alternatively, directed sorting can be conducted where the experimenter provides sorting criteria by specifying: (1) the number of groups to be formed or (2) the nature of the properties of the groups. Examples of the latter scenario can be found in studies conducted where Sauvignon Blanc wines were sorted according to: (1) origin (Parr et al., 2010); (2) specific wine style groups called “green” and “not green; (3) ripeness levels with groups called “ripe” and “not ripe” and (4) typicality calling groups “good varietal definition” and “not good varietal definition” (Parr et al. 2007).

A variation of sorting, called hierarchical sorting, has been used and indicated by Courcoux et al. (2012) as more precise and stable than free sorting. During ascendant hierarchical sorting judges are asked to sort the samples according to similarities into groups and then merge groups until only one group exists (Coxon et al., 1999; Courcoux et al., 2012). Descendant sorting can also be conducted where groups are subdivided until no further groups can be formed (Santosa et al., 2010; Cadoret et al. 2011). Both ascendant and descendant can be used conjointly in the same experiment to obtain a full hierarchy of similarities and dissimilarities of the products (Honoré-Chedozeau et al., 2017).

A second step known as “verbalisation” or “labelling” (Bécue-Bertaut et al., 2011) is usually conducted after the judges sorted the wine samples. During this step words are provided to describe the groupings in such a way that the differences and similarities between the groups are highlighted. This step can be seen as a verbal step (Fig. 3) supplementary to the similarity-based main task where samples are grouped. The grouping data obtained during sorting is then captured by constructing similarity or distance matrices.



Judge 1		
Group	Samples	Descriptors
1	Wines 1, 2	Grapefruit, Green pepper, Grass, Asparagus
2	Wines 3, 10	Grapefruit, Guava, Passion fruit
3	Wines 6, 7, 4	Pineapple, Yellow apple, Dried fruit, Pear
4	Wines 5, 8	Dried fruit, Raisin, Nuts, Sherry
5	Wine 9	Dried fruit, Potato, Spice

	Wine 1	Wine 2	Wine 3	…	Wine 10
Wine 1	1	1	0	…	0
Wine 2	1	1	0	…	0
Wine 3	0	0	1	…	1
…	…	…	…	…	
Wine 10	0	0	1	…	1

	Passion fruit	Guava	Grapefruit	…	Grass
Wine 1	0	0	1	…	1
Wine 2	0	0	1	…	1
Wine 3	1	1	1	…	0
…	…	…	…	…	
Wine 10	1	1	1	…	0

	Wine 1	Wine 2	Wine 3	…	Wine 10
Wine 1	30	18	2	…	0
Wine 2	18	30	0	…	0
Wine 3	2	0	30	…	1
…	…	…	…	…	
Wine 10	0	0	1	…	30

	Passion fruit	Guava	Grapefruit	…	Grass
Wine 1	0	2	19	…	21
Wine 2	5	0	22	…	12
Wine 3	15	17	12	…	0
…	…	…	…	…	
Wine 10	7	9	10	…	0

Fig. 3. Schematic presentation of the sorting task. Products are grouped according to similarity and descriptors provided to describe the groups. The grouping data are captured as distance matrices and the descriptors compiled in contingency tables. DISTATIS can be applied to the individual distance matrices or multidimensional scaling (MDS) to a summed distance matrix with the attributes projected onto the multivariate map as supplementary variables.

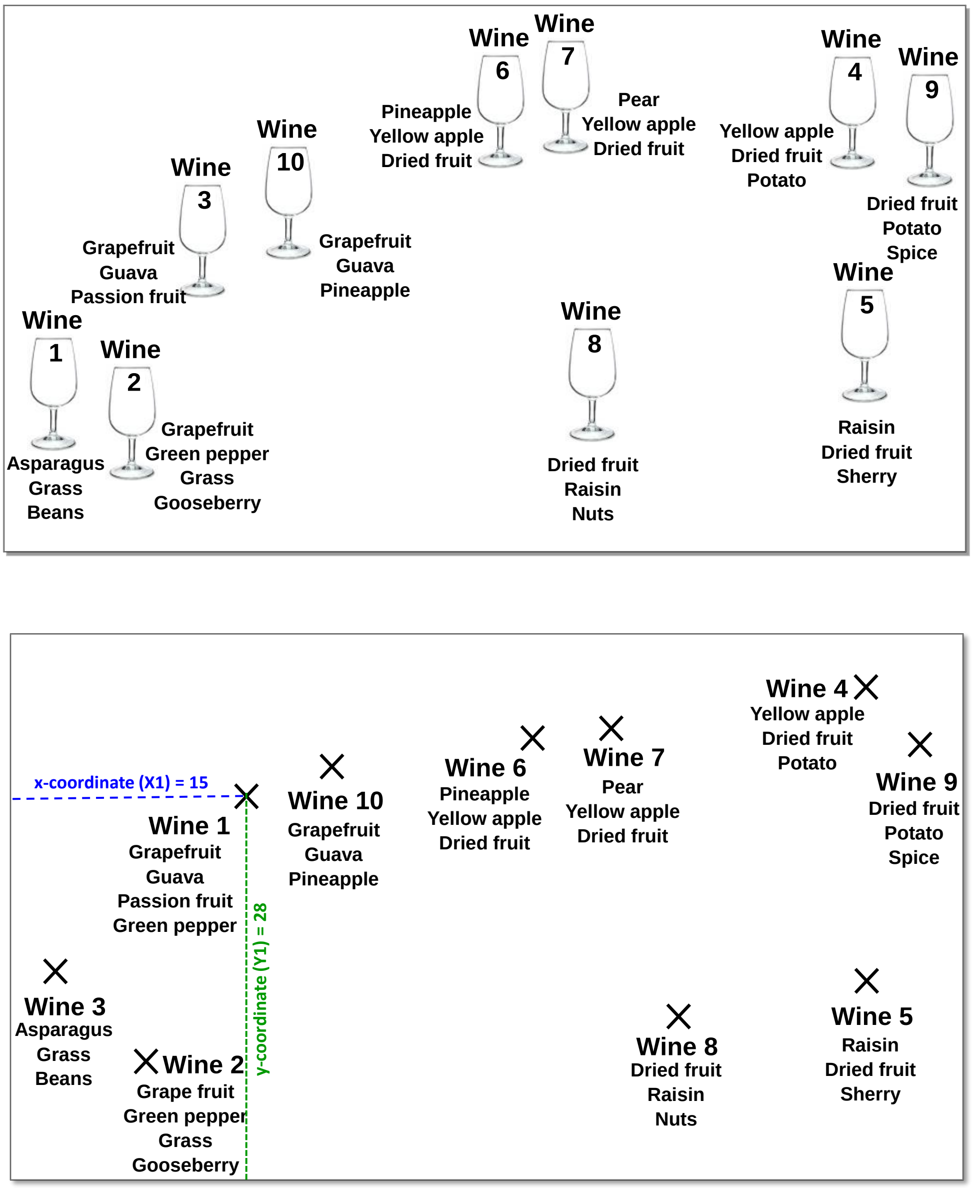
2.2.3.2 Projective mapping (PM)

Projective mapping (PM) depends on the ability of sensory judges to translate similarities and differences between products onto a two-dimensional space. Samples are presented simultaneously to sensory judges, and the judges are instructed to place samples close to each other if they are similar and far apart if they are different (Risvik et al., 1994, 1997). In addition to positioning the samples on a two-dimensional space, judges are asked to describe each sample using descriptors. Verbal-methods such as ultra flash profile (UFP, Perrin et al., 2008, 2009) or CATA can be used for that purpose (Fig. 4).

The *X* and *Y* coordinates for each product is measured and tabulated in a data matrix keeping the *X* and *Y* coordinates provided by each judge as a separate data table. The descriptor data are compiled in a contingency table (Fig. 4).

A restricted version of PM was introduced by Pagès (2003) where a 60 x 40 cm two-dimensional space is used for the organisation of the samples and MFA (Escofier & Pagès 1990, Pagès 2003, 2005) is used to analyse the *X* and *Y* coordinates of each sample as provided by the sensory judges.

Pagés et al. (2010) proposed to combine PM and sorting calling the new method “sorted Napping”. Sensory judges first organise the samples in terms of similarity by placing them close to or far from each other. As a second step, samples are grouped by drawing circles around them.



	PM Coordinates table					Supplementary data table			
	Judge 1		Judge 2		…	Sensory attributes– sum of all the tasters data			
Wines	X1	Y1	X2	Y2	…	Passion fruit	Guava	…	Grapefruit
Wine 1	15	28	12	32		23	16		24
…									
Wine 10									

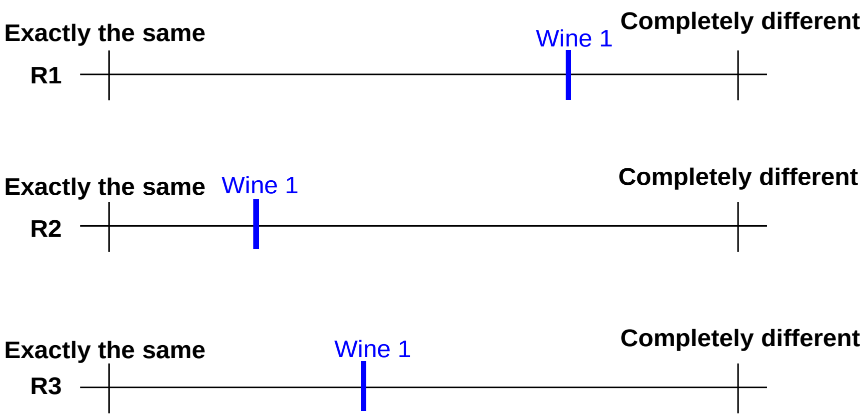
Fig. 4. Schematic presentation explaining how projective mapping (PM) is conducted, by arranging samples based on similarity and dissimilarity. The data are captured, measuring the distance from the left bottom corner to obtain *X* and *Y* coordinates for each product position. Multiple factor analysis (MFA) is most frequently used to analyse and visualise PM data.

2.2.4 Reference-based methods

The common factor between reference-based methods is the use of a reference sample(s), against which the other samples in the set are profiled. The main advantage of this category of methods is the possibility of aggregating data when large sample sets are evaluated, as separate subsets, during different sessions or when samples are analysed over a longer time period. However, a suitable, stable reference, is needed for these methods. Polarised sensory positioning (PSP, Teillet et al., 2010), polarised projective mapping (PPM, Ares et al., 2013) and pivot profile (PP, Thuillier et al., 2015) belong to this category.

2.2.4.1 Polarised sensory positioning (PSP)

When PSP is performed the similarity between evaluated and reference products, called poles, are rated. The products are served one at a time and judges have to indicate on a line scale anchored at "exactly the same" to "completely different" how similar each product is to each reference (Fig. 5). Teillet et al. (2010) proposed the use of three poles and in addition proposed Triad-PSP where sensory judges are asked to which one of the poles the evaluated product is respectively most similar and least similar.



	Judge 1			Judge 2				Judge 3			
Wines	A1	…	A6	A1	A2	A3	…	A1	A2	A3	…
Wine 1	2		4								
Wine 2	6		6								
Wine 3	3.5		1								
Wine 4	3.5		2.5								
Wine 5	5		5								
Wine 6	1		2.5								

Fig. 5. Schematic presentation of the line scale used during polarised sensory positioning (PSP). The mark on the line scale is measured for each comparison between the reference and the product evaluated. Averages are used if principal component analysis (PCA) or multidimensional scaling (MDS) unfolding is used for the data analysis and individual data if STATIS or multiple factor analysis (MFA) is used.

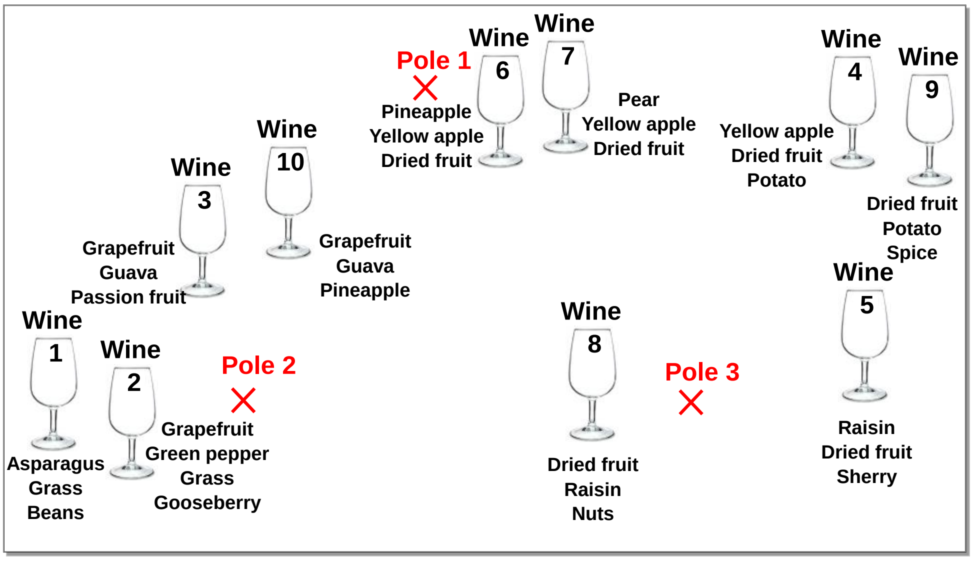
2.2.4.2 Polarised projective mapping (PPM)

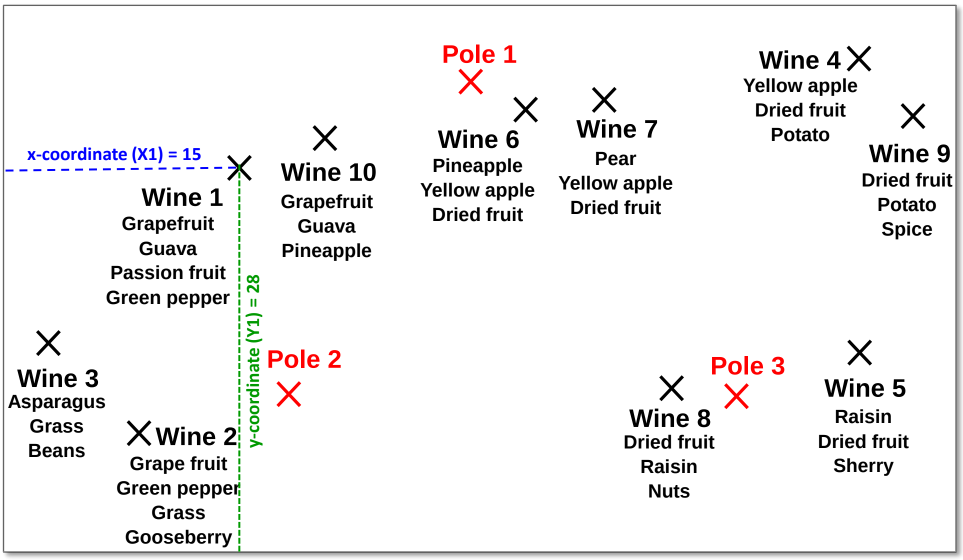
Ares et al. (2013) proposed to combine PSP and PM to address the limitations of the methods and combine their strengths. When PM is performed all samples are presented simultaneous, the number of samples that can be analysed is, therefore, limited. When PSP is performed samples are compared to the reference samples one at a time which does not facilitate direct comparison of the samples to each other. Furthermore, descriptive information is obtained only relative to the poles (Teillet et al., 2010).

The rationale behind the development of PPM was to develop a method where a direct description of each product could be obtained relative to reference samples. Data aggregation is possible, combining data from more than one sensory evaluation session, by keeping the poles constant for all the sessions. When PPM is performed, sensory judges are provided with a PM sheet where the poles are pre-located and their positions marked. All the samples to be evaluated are presented simultaneously and have to be located close to, if they are similar, or far away, if they are perceived differently, from the poles. A few words to describe each sample are usually provided after the positioning is finalised as for UFP (Perrin et al., 2008) performed during PM (Fig. 6). The data are analysed with MFA using the same protocol as for PM data (Pagès, 2005).

2.2.4.3 Pivot profile (PP)

Pivot profile (PP) was introduced by Thuillier et al. (2015) when profiling Champagne with wine industry professionals as sensory judges. When PP is conducted, sensory judges receive samples in pairs of two, the pivot and a sample to be evaluated. Attributes perceived as “more intense” and “less intense” in the sample than the pivot has to be provided (Fig. 7). Judges are restricted to only use words, no phrases and refrain from using the negative form of words. When PP data are collected -1 is recorded when an attribute is perceived less intense than the pivot and 1 if it is perceived more intense. The sum of all the negative and the positive frequencies over all the judges are computed for each attribute for each wine. The number of negative frequencies is subtracted from the number of positive frequencies. The data is translated by adding the absolute value of the minimum to all the values to obtain positive values (Fig. 7).





	PM Coordinates table					Supplementary data table			
	Judge1		Judge2		…	Sensory attributes– sum of all the tasters data			
Wines	X1	Y1	X2	Y2	…	Passion fruit	Guava	…	Grapefruit
Wine 1	15	28	12	32		23	16		24
…									
Wine 10									

Fig. 6. Schematic presentation of polarised projective mapping (PPM) showing how sensory judges place the samples relative to the poles on a 2-dimensional surface. The pre-located poles are indicated in red and the unknown products in black. Data capturing and analysis are conducted using the same approach used when projective mapping (PM) is conducted. Multiple factor analysis (MFA) with the descriptors added as supplementary variables is used frequently.

Wine 1	**LESS** intense than pivot sample	**MORE** intense than pivot sample
Appearance	Green	Brown, hazy
Aroma “on the nose”	Fruity, fresh, tropical, grass	Dried fruit, sherry
Taste and mouthfeel “on the palate”	Sour, astringent	Sweet, hotness
Aftertaste and length		Bitter

Judge 1						
Wines	Fruity	Tropical	Dried fruit	Sherry	Grass	…
Wine 1	-1	-1	1	1	-1	
Wine 2	1	1	-1	1	1	
Wine 3	1	0	1	1	1	
……						
Wine 6	1	0	-1	1	1	

Sum over all the judges						
Wines	Fruity	Tropical	Dried fruit	Sherry	Grass	…
Wine 1	-14	-6	15	8	-7	
Wine 2	12	13	-11	4	5	
Wine 3	7	6	-7	12	11	
……						
Wine 6	0	0	0	8	9	

Sum over all the judges - translated						
Wines	Fruity	Tropical	Dried fruit	Sherry	Grass	…
Wine 1	0	8	29	22	7	
Wine 2	26	27	3	18	19	
Wine 3	21	20	7	26	25	
……						
Wine 6	0	0	0	22	23	

Fig. 7. Schematic presentation of pivot profile (PP) showing how data from the tasting ballot are captured. The sum of the citation frequencies are recorded for the individual judges, summing over all the sensory judges and translation of the data is performed prior to statistical analysis. Correspondence analysis (CA) is used to visualise PP data.

2.3 Sensory panels performing rapid sensory methods

In a review by Varela and Ares (2012) the authors noted that rapid sensory methods “cross the fine line between sensory testing and consumer acceptance testing”. Although consumers are most frequently used as sensory judges, product specialists referred to as industry professionals or experts are also used as sensory judges to perform rapid sensory analysis (Ballester et al., 2005, 2008, 2009; Parr et al., 2007; Perrin et al., 2008; Bester 2011, Johnson et al., 2013; Torri et al., 2013; Botha, 2015; Coulon-Leroy et al., 2017; Wilson et al., 2018). A number of studies reported the suitability of rapid methods for sensory testing using a trained panel (Delarue & Sieffermann, 2004; Louw et al., 2013; Thuillier et al., 2015; Liu et al., 2016; Moelich et al., 2017; Vidal et al., 2017, 2018).

These methods are, therefore, versatile and robust in the sense that training of the panel is not required to obtain good quality results. However, choosing the type of panel to perform the analysis depends largely on the expected outcome and detail needed.

2.3.1 Consumer panels

Consumers' data can contain a large number of descriptors with a low frequency of citation (Valentin et al., 2012). Therefore, larger numbers of sensory judges are recruited for consumer panels, consisting of up to a 100 participants, and expert panels, up to 30, than trained panels, where 8 to 15 judges are typically employed (Varela & Ares, 2012).

2.3.2 Trained panels

On the other hand, trained panelists might only use the few terms that they were trained for, resulting in a loss of information (Albert et al., 2011). It was, therefore, proposed to use product experts and rely on their work experience as industry professionals as sensory training. They are often referred to as expert panels in literature and are mainly used during wine sensory analysis with rapid methods (Perrin et al., 2008; Parr et al., 2010, 2015; Lawrens et al., 2013; Picard et al., 2015; Coulon-Leroy et al., 2017; Vidal et al., 2018).

2.3.3 Industry professionals / experts

Expert panels produce precise results with sufficient technical detail and describe products differently when compared to consumers since they know the production process. These differences were highlighted by several authors (Bester 2011, Botha, 2015; Torri et al., 2013; Ballester et al., 2005, 2008, 2009; Honoré-Chedozeau et al., 2017). It is, therefore, important to choose a rapid sensory method and panel that is fit-for-purpose. It was noted by Delarue and Sieffermann (2004) that when FP is used it is better to use "expert" judges, where "expert" judges were referred to as judges with experience in sensory evaluation.

2.4 Statistical analysis of rapid sensory method data

Rapid sensory methods are used to determine the sensory properties of products and how they are related in terms of similarities and dissimilarities. In order to achieve this aim multivariate sensory maps are constructed to visualise the data. The specific multivariate sensory technique applied, mainly depends on the type of data generated during the sensory analysis.

2.4.1 Single-block analysis techniques

Single-block statistical analysis techniques are performed on the averaged or summed data over the response of all the judges. This means that the difference between the individual judges' data is lost and not represented in the sensory map. If the experimenter wants to investigate differences between the individual judges' multi-block analysis should be conducted that will be discussed in section 2.4.2.

2.4.1.1 Principal component analysis (PCA)

PCA is one of the most commonly used multivariate statistical techniques, and forms the mathematical basis for various other techniques. PCA is used to transform a data matrix consisting of many variables, in sensory science attributes, into a space where those variables, their relationship to each other and the products they describe can be visualised by a reduced number of components (Esbensen, 2002). These components are plotted in a two or three-dimensional space to obtain a multivariate sensory map. PCA is most frequently performed on the correlations matrix which implies that the data was scaled prior to analysis and all the attributes will have the same weight. When analysing sensory data, where all the variables are within the same order of magnitude, PCA can be conducted on the covariance matrix. In this case, attributes with lower scores will be less important than attributes with higher scores.

PCA is used for the analysis and visualisation of DA data (Stone & Sidel, 1974; Lawless & Heymann, 2010; Tomic et al., 2010). Initially, a number of rapid method data sets were analysed with PCA prior to the development of customised methods tailor-made for that specific data type. A few of these cases were specified by Valentin et al. (2012). Recently PCA was used for the analysis of RATA (Reinbach et al., 2014) and PSP (Teillet et al., 2010) data (Table 2).

2.4.1.2 Correspondence analysis (CA)

CA (Takane, 1982) is a generalised PCA method adapted specifically for the analysis of ordinal data such as frequency data, where the number of times attributes were cited for different products are recorded. Therefore, CATA data sets are frequently analysed using CA (Ares et al., 2015). Chi-square distances are the most commonly used when CA is conducted, however, if Hellinger distances are used instead, attributes with low citation frequencies could be included in the analysis without taking the risk that those attributes will skew the multivariate map (Popper et al., 2011; Meyners et al., 2013). Ares et al., (2015) used CA based on Hellinger distances for the analysis of RATA data since RATA data can be seen as weighted CATA data. PP data is also analysed by CA (Thuillier et al., 2015)

2.4.1.3 Multidimensional scaling (MDS)

MDS can broadly be described as a method that rearrange products according to their similarities or dissimilarities to obtain the most efficient approximation of the distances between the products by minimising the stress which is a measure of the noise or error in the data set (Kurskal & Wish, 1978; King et al., 1998).

MDS is most commonly used for the analysis of sorting data by subjecting a similarity matrix consisting of the number of times each pair of products is grouped together to non-metric MDS (Lawless et al., 1995). However, similarity between two products can, in addition, be viewed as a distance and, therefore, an Euclidean metric which can be analysed by metric MDS (Abdi et al., 2007).

Teillet et al. (2010) used MDS unfolding for the analysis of PSP data. Originally, PM-type data was analysed with non-metric MDS (Valentin et al., 2012; Varela & Ares 2012) as well.

2.4.2 Multi-block analysis techniques

Multi-block analysis can address the limitation of the loss of individual data when single-block methods are used. Currently, in sensory science, multi-block analyses are mainly used to investigate the differences and similarities between the data obtained from (1) individual judges and (2) different sensory methodologies. In addition, it can be used to compare different panels (Bécue-Bertaut & Lê, 2011).

2.4.2.1 Multiple factor analysis (MFA) and multiple correspondence analysis (MCA)

MFA consists of multiple PCA or CA analyses depending on the data types of the different data blocks also called data tables (Pagès, 2005; Nestrud & Lawless, 2008; Le Dien & Pagès, 2003; Ares, et al., 2010a, 2010b).

MFA can be used to analyse PM data. When PM data are analysed the coordinates of the products can be subjected to MFA (Escofier & Pagès, 1990) keeping the data for each judge separate as a different data table in the MFA analysis. The Euclidean distance configuration of the products for each judge is calculated simultaneously and a biplot containing the data from all the sensory judges is obtained with this procedure. PCA is thus performed on the coordinate data from each judge. The descriptor data are added as a separate data table that is frequently added as supplementary data and projected onto the MFA compromise map. In this case both the product positioning and the attributes used, to describe the positioning, are represented on a single graph or sensory map (Perrin et al., 2008).

Another less frequently used application for MFA is the analysis of PSP data as proposed by Telliet et al. (2010). MFA can provide a measure for similarity between different data sets that can be visualised by inspection of the partial projections map. This map can be used to visualise differences between sensory judges. In addition, data from different sensory methods can be analysed as different data tables to be compared. Dehlholm et al. (2012a) used MFA to compare different sensory methods to each other.

MCA is a restricted version of MFA where multiple CA analyses are conducted. To be historically correct it should be noted that MCA is an older technique than MFA and from that perspective MFA is an enriched MCA that uses both CA and PCA to analyse the separate data blocks. MCA was used in sensory research for the analysis of CATA (Varela & Ares, 2012), sorting (Cadoret et al., 2009) and PSP data (Ares et al., 2013).

2.4.2.2 Generalised procrustes analysis (GPA)

Until recently GPA (Gower, 1975) was a popular statistical analysis method used for the analysis of PM data (Risvik et al., 1994). When GPA is conducted the data is transformed by: (1) translation, where all the individual PM configurations, obtained from the different sensory judges, are moved to the middle of the PM sheet; (2) rotation and reflection to align the individual PM data sets and (3) isotopic scaling, where the individual data is stretched or shrank to obtain the best fit and reduce the individual differences.

MDS was compared to GPA for the analysis of PM data by King et al. (1998), even though it was concluded that higher dimensions could be investigated using MDS than GPA, MDS is not commonly used for analysing PM data

Tomic et al. (2015) found that MFA and GPA produced similar results for simulated data, but different results for “real data”, in a study where these two methods were compared when analysing PM. MFA can provide data with a higher dimensionality than GPA which is an advantage. Thus, MFA is the most popular technique for analysing PM data currently.

Kennedy et al. (2009) proposed procrustes multiple factor analysis (PMFA) a method where procrustes rotation is incorporated into the MFA analysis. This method is, however not commonly used. GPA can also be used to analyse FP data by performing PCA for each individual judges’ data which is then subjected to translation, rotation and isotopic scaling while integrating the different data sets to obtain a single multivariate map (Gower, 1975 ; Moussaoui & Varela, 2010).

2.4.2.3 INDSCAL

INDSCAL is a multi-block generalisation of MDS applied to the individual sensory judges' distance matrices. The weighted Euclidean model is used to transform the product coordinates into distances (Bárcenas et al., 2004; Nestrud & Lawless, 2011).

In a recent study Næs et al. (2017) compared INDSCAL and MFA when analysing PM data. It was found that, even though MFA is based on coordinate data and INDSCAL on distance data similar results were obtained. MFA, however, performed slightly better as a consensus indicator, explaining how well judges agreed in terms of sensory perception of the products.

2.4.2.4 DISTATIS

DISTATIS (Abdi et al., 2007) was proposed as a generalised MDS-based method to address the fact that individual differences between judges are not taken into account when MDS is performed. When DISTATIS is performed the individual distance matrix of each judge is transformed into a cross-product matrix which is normalised. The individual matrices are combined prior to eigenvalue decomposition producing a DISTATIS compromise cross-product matrix. The DISTATIS compromise map is used to visualise the similarities between the products. The attributes used to describe the groups made by the sensory judges are projected onto the DISTATIS compromise map as supplementary variable not playing a role in the product configuration. DISTATIS is currently the most popular statistical analysis technique used for analysing sorting data (Abdi et al., 2007).

2.4.2.5 Less frequently used methods

The FAST method was proposed by Cadoret et al. (2009) to optimally represent all sensory judges, using MFA, and samples, using MCA, when the multivariate sensory map is constructed (Cadoret et al. 2009). SORT CC was proposed by Qannari et al. 2009 as another multi-block technique for the analysis of sorting data where individual data are taken into account. Hierarchical multiple factor analysis (HMFA) was used by Bécue-Bertaut and Lê (2011) to analyse and compare sorting data generated by more than one panel. FAST, SORT CC and HMFA is not currently frequently used although the ideas behind the development of these techniques are scientifically justified. These methods might be used more frequently in future.

2.4.3 Visualisation of multivariate maps

The multivariate sensory maps produced during statistical analysis are used to visualise the relationships between: (1) the different products, e.g. score plots, (2) the attributes, e.g. loadings plots and (3) the products and the attributes, e.g. biplots (Fig. 1). The magnitude of these relationships, mainly similarities and dissimilarities, are described by identifying positive and negative correlations when PCA-based methods or distances when MDS-based methods are performed. Interpreting these graphs by means of inspection is the most common practise.

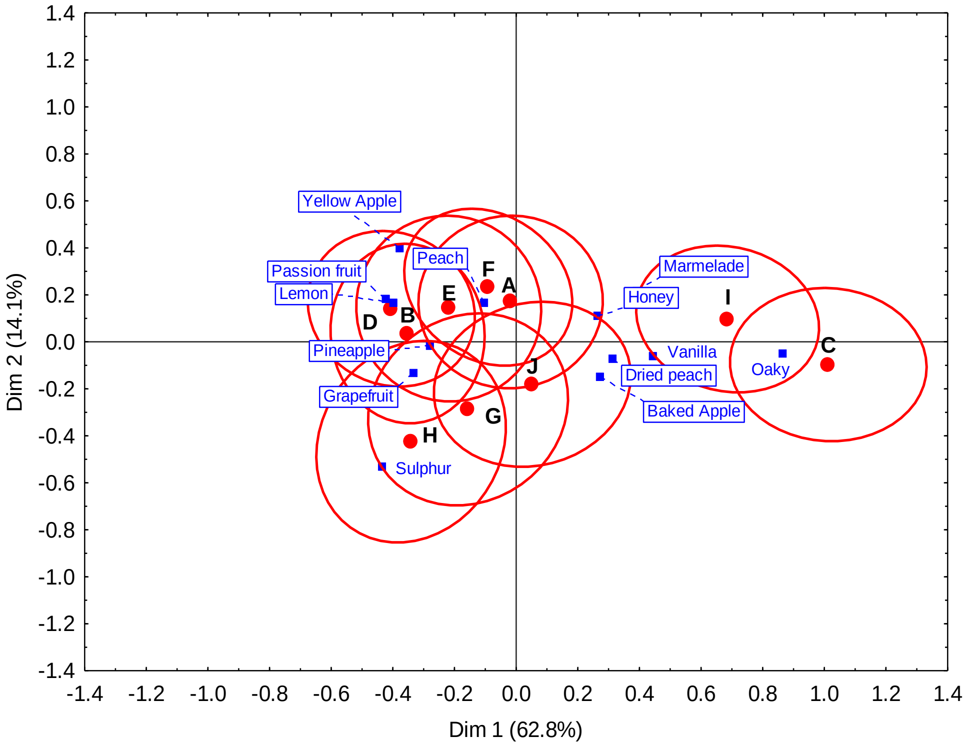


Fig. 1. An example of a PCA biplot with the samples represesnted as scores and variables, in this case attributes, as loadings (own data).

A second approach is to perform statistical analysis on the multivariate configurations which are often represented by the coordinates of the sensory map. Hierarchical cluster analysis (HCA) with Ward's linkages as aggregation criteria (Fig. 2) are commonly used and were for example used to investigate the clustering of products (Chollet et al., 2011; Veinand et al., 2011), attributes or sensory judges (Ferrage et al., 2010) to explain product similarities and evaluate panel performance.

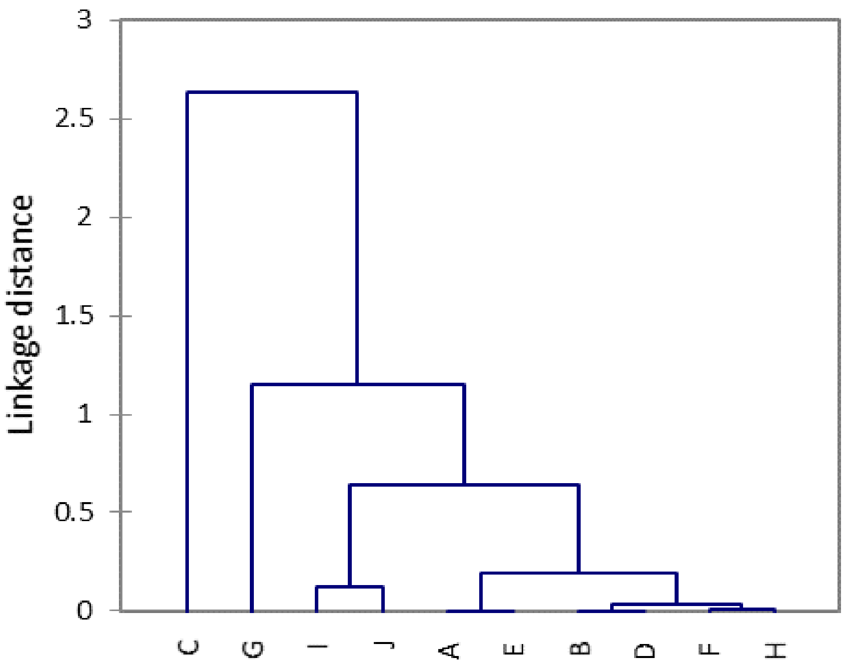


Fig. 2. An example of a hierarchical cluster analysis (HCA) dendrogram where Euclidean distances and Ward's aggregation criteria was used during statistical analysis to identify groupings on a multivariate sensory map (own data).

Confidence ellipses (Fig. 3) provide an estimation of the fluctuation of a product on the multivariate sensory map. In the ideal scenario these fluctuations should be small. Strategies to calculate confidence ellipses were developed for many of the statistical analysis methods used to analyse rapid sensory data, but not all. Amongst these are the strategies proposed by Cadoret and Husson (2013), Dehlholm et al. (2012b) for MFA, Abdi et al. (2009) for DISTATIS and Courcoux et al. (2012) for MDS. It is important to note that the statistical calculations and outputs differ to some extent and a fit for purpose solution should be chosen. Further work in this field could make a valuable contribution to the statistical analysis portfolio of methods currently available.

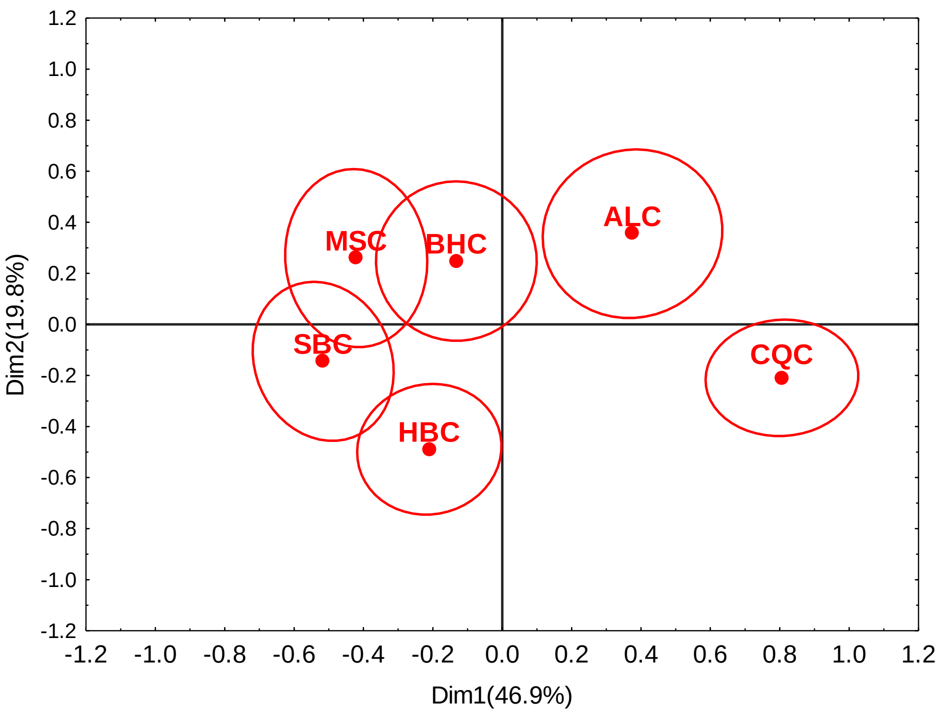


Fig. 3. An example of correspondence analysis (CA) performed on CATA data where confidence ellipses were calculated using bootstrapping (own data).

When different multivariate maps are compared RV coefficients are commonly used (Perrin et al., 2008; Reinbach et al., 2013: Vidal et al., 2018; Dehlholm et al., 2012a). RV coefficients are measures of the similarity between two data sets or in this case spaces (Robert & Escofier, 1976).

2.4.4 Pre-treatment of descriptor data

In some cases, pre-treatment of the data is conducted prior to statistical analysis. CA is commonly applied to PP data, however, the data is recorded as -1 if the attribute was perceived as “less intense” in the sample than the pivot and 1 if it was perceived as “more intense”. CA cannot be conducted on negative values, therefore a translation step is incorporated to create a data matrix consisting of positive values. After all the citations of a specific attribute are summed over all the judges translation is conducted. During this step, the absolute value of the largest negative value is added to all the values in the data matrix to obtain only positive values with the lowest scoring attribute being zero (Thuillier et al., 2015).

When verbal-based methods are used sensory judges use language to express perception. Different individuals often use different words to describe the same concept or stimuli. Therefore, textual data analysis including techniques such as coding is required to convert text to data typically captured in contingency tables (Fonseca et al., 2016). This is the case when methods such open-ended questions, are used.

When data, obtained from CATA, are captured contingency tables are constructed directly from the sensory judges' responses. However, different sensory judges can still choose different but similar terms when the same or similar sensory attribute is perceived, therefore, these attributes can be reduced by means of: (1) lemmatisation, combining linguistic and semantic synonyms or (2) statistical analysis prior to constructing the multivariate sensory map (Campo et al., 2008, 2010; Thuillier et al., 2015; Symoneaux et al., 2012; Fonseca et al., 2016; Wilson et al., 2018).

Attributes can be reduced even further, by setting a cut-off point taking into account: (1) the quotation frequency percentage or (2) the number of sensory judges that used an attribute. Cartier et al. (2006) took into account only attributes with a quotation frequency of 3% or higher. Campo et al. (2008, 2010) combined terms used by less than 15% of the panel with an appropriate synonym. If no synonym could be found the word was not used during the statistical analysis process. Wilson et al. (2018) followed a similar procedure combining or discarding terms used by less than 20% of the panel. Symoneaux et al. (2012) and Fonseca et al. (2016) only kept attributes used by at least 5% of the panel for multivariate analysis. After reducing the attributes, CA can be performed on the sum of the citations over all the judges for each attribute for each product compiled in a contingency table.

Semantic combination of attributes is not standard practice, many authors choose to select and refine attributes by means of statistical analysis prior to constructing the multivariate sensory map. Cochran's Q test is applied to determine which attributes are perceived significantly different for the different products (Parente et al., 2011).

In the majority of studies where rapid sensory methods are used the authors do not specify if all the attributes were included in the multivariate analysis. It is, therefore, assumed that all attributes were included. In some studies, it was motivated why all the attributes were included, e.g. Santos et al. (2013) argued that since an attribute was cited it was important to that particular sensory judges and should be included in the statistical analysis.

TABLE 2
Most frequently used statistical methods for the analysis of rapid sensory data.

Statistical method		Sensory method									Reference
	Multi-block	DA[a]	FP[b]	CATA[c]	RATA[d]	Sorting[e]	PM[f]	PSP[g]	PPM[h]	PP[i]	
PCA[j]		✓			✓			✓			Tomic et al., 2010[a]; Teillet et al., 2010[g]; Reinbach et al., 2014
CA[k]				✓	✓					✓	Ares et al., 2015[c,q], Ares et al., 2014[d,r], Thuillier et al., 2015[i], Picard et al., 2003; Soufflet et al., 2004
MCA[l]	✓			✓		✓		✓			Takane, 1982; Popper et al., 2011[rc], Cadoret et al., 2009[e]; Ares et al., 2013[g,h]
MFA[m]	✓			✓			✓	✓	✓		Escofier and Pages, 1990; Abdi and Valentin 2007; Ares et al., 2010b[c]; Pagès, 2003; Perrin et al., 2008, 2009; Teillet et al., 2010[g]; Ares et al., 2013[g,h]
HMFA[n]	✓						✓				Le Dien and Pagès, 2003[f]; Perrin et al., 2008, 2009; Bécue-Bertaut and Lê, 2011[e]
GPA[o]	✓		✓				✓				Gower, 1971; Risvik et al., 1994[a,f]; Moussaoui and Varela 2010[b]
MDS[p]						✓		✓			Kurskal and Wish, 1978; Lawless et al., 1995[e]; Teillet et al., 2010[g] (MDS unfolding)
DISTATIS	✓					✓					Abdi et al., 2007[e]
FAST	✓					✓					Cadoret et al., 2009[e]
SORT CC	✓					✓					Qannari et al. 2009[e]
INDSCAL	✓					✓					Bárcenas et al., 2004[e]; Nestrud and Lawless 2011[e]

Acronyms used for rapid sensory methods: [a]Descriptive analysis; [b]Flash profile; [c]Check-all-that-apply; [d]Rate-all-that-apply; [e]Sorting task, [f]Projective mapping which includes Napping; [g]Polarised sensory positioning; [h]Polarised projective mapping; [i]Pivot profile.
Acronyms used for statistical methods: [j]Principal component analysis; [k]Correspondence analysis; [l]Multiple correspondence analysis; [m]Multiple factor analysis; [n]Hierarchical multiple factor analysis; [o]Generalised procrustus analysis; [p]Multidimensional scaling.
[q]Chi square distances or [r]Hellinger distances were calculated when CA was performed

2.5 Rapid sensory analysis applied to alcoholic beverages

The popularity of rapid sensory methods to evaluate the sensory properties of alcoholic beverages such as wine, beer and brandy has increased tremendously during the past 20 years (Valentin et al, 2012; Varela & Ares, 2012; Louw et al., 2013, 2014, 2015; Navajas et al. 2014; Lezaeta et al, 2017, 2018; Vidal et al., 2015, 2017; 2018).

Rapid sensory methodologies are used for various purposes including sensory profiling and as pre-screening tools prior to conducting detailed sensory analysis, e.g. Piombino et al. (2004) used sorting conducted with consumers as pre-selection tool prior to DA conducted by a trained panel. The two most popular rapid methods used to date for wine sensory analysis are CATA and sorting. CATA performed by consumer panels to profile wine became increasingly popular lately with a number of articles published in the last 3 years.

2.5.1 Check-all-that-apply (CATA)

2.5.1.1. Lexical differences obtained from consumers of different cultural groups

Weightman (2018) studied consumer perception of different white wine cultivars using CATA as profiling technique to investigate differences between cultural groups within South Africa. No significant difference between different cultural groups were found, however, differences between male and female consumers were found.

Corsi et al. (2017) investigated the lexical differences between Chinese and Western wine consumers. It was shown that no significant difference were found when generic terms were used. Generic terms were used three times more than culture-specific terms, for which a significant difference was observed.

2.5.1.2 Consumer perception of different wood treatments

Alencar et al. (2018) investigated the sensory perception of Syrah subjected to different wood treatments. The methods was applied successfully and it was shown that consumers can distinguish between different oak treatments and a large segment, more or less 50%, of consumers disliked wine when oak chips was used. Botha (2015), who compared consumer acceptance of unwooded and wooded Chenin Blanc, found that consumers preferred either unwooded or barrel fermented wine. Wine subjected to alternative wood treatments were less liked or slightly disliked.

2.5.1.3 Astringency perception

Vidal et al. (2015) investigated consumers' perception of astringency in red wine and found that wine involvement did not contribute to segmentation of consumers. Furthermore, only 17 of the 31 descriptors on the mouthfeel wheel were used by more than 10% of the participants. The authors concluded that the mouthfeel wheel might not be an efficient tool to communicate astringency to consumers.

The astringency of commercial Tannat wine was characterised (Vidal et al., 2017) and the relationship between astringency and phenolic composition studied (Vidal et al., 2018) using a trained panel. CATA was applied to assess astringency sub-categories where main categories were analysed by means of time-intensity analysis (Vidal et al., 2018). It was shown that the astringency profiles of those specific Tannat wines were not correlated to the vintage, price segment or aging in oak barrels.

2.5.1.4 Other studies investigating the effect of oenology and viticulture practices on wine profiles

Coste et al. (2018) used a CATA list representing emotional terms in combination with optimised descriptive profile (ODP), to distinguish between red wines originating from cool and warm regions. The two wine styles could be distinguished, analysing the data obtained from CATA performed by consumers. The cool climate wines were less liked and described as “most surprising”, “red brick colour”, “complex smell” and “aggressive mouthfeel”.

Lazeata et al. (2017, 2018) used consumers conducting CATA to profile enhanced Sauvignon Blanc wines prepared from enriched grape must and compared enhanced wines to the “ideal product”. Ares et al. (2015) investigated the differences in sensory profiles obtained when CATA was performed by consumers and trained panels, one of the matrices investigated was Sauvignon Blanc wine.

RATA, a variant of CATA, was used by Franco-Luesma et al. (2016) to study the effect of reductive volatile sulphur compounds on the sensory profile of young red wine using semi-trained expert judges, mostly university students specialising in wine sciences.

In all of these cases differences between the products’ sensory properties could be described showing that CATA is an efficient rapid method for analysing the wine matrices.

2.5.2 Sorting

Sorting is a popular method for the analysis of both wine and beer using product experts as sensory judges (Chollet & Valentin, 2001; Abdi et al., 2007; Lelièvre et al., 2008, 2009). Sorting is the most popular rapid sensory method used for wine sensory analysis (Valentin et al., 2012). Campo et al. (2010) and Bester (2011) obtained similar results when sorting was compared to descriptive techniques, where a trained panel was used, while evaluating white wines.

2.5.2.1. Studies on cultivar concepts and wine style

In addition, Bester (2011) performed both directed and undirected sorting to investigate the ability of wine industry professionals and consumers to identify “wooded”, “rich and ripe” and "fresh and fruity" Chenin Blanc styles. In similar studies, using free sorting, Hanekom (2012) and Van Antwerpen (2012) also investigated Chenin Blanc wine style perception, where Hanekom analysed Chenin Blanc wines made from grapes grown on bush vines. It was found that both consumers and experts could distinguish between wooded and unwooded wines.

Ballester et al. (2005) studied the "Chardonnay concept" by conducting sorting using experts and consumers. In another study, Ballester et al. (2008) compared the ability of consumers and experts to discriminate between Melon de Bourgogne and Chardonnay from Burgundy.

The concept of "typicality" was studied conducting sorting experiments with industry professionals as sensory judges by Parr et al. (2007). "Minerality" was studied similarly by Parr et al. (2015) and Ballester et al. (2013) using the sorting task.

2.5.2.2. Wine origin and cross-cultural experts' perception

Parr et al. (2010) applied directed sorting according to origin when evaluating French and New Zealand Sauvignon Blanc wines. French and New Zealand winemakers could successfully sort wine according to the country of origin. The French wines could also be divided into sub-regional groups. This was however, not the case for New Zealand wines. Bécue-Bertaut and Lê (2011) used sorting during a cross-cultural study, evaluating the perception of French and Catalan wines.

Johnson et al. (2013) investigated the sensory attributes perceived for Australian Shiraz from different regions conducting a free sorting exercise with wine industry professionals. Clear differences between regions were difficult to identify and even more so when the region had diverse geography and climatic conditions.

2.5.2.3 Measuring quality perception

Navajas et al. (2014) used a directed sorting task specifying four quality groups, "very high", "high", "low" and "very low" into which consumers had to sort wines assessing extrinsic cues. A difference in quality perception could be seen between judges with high and low involvement with wine. Judges with low involvement used the origin as a decisive factor.

2.5.3 Projective mapping (PM)

In addition to the studies conducted by means of sorting, Heymann et al. (2013) studied the concept of "minerality" using a trained panel to conduct DA and an expert panel performing PM.

Pages (2005) performed PM on wine from Touraine using wine experts and supplemented the similarity-based data with DA data obtained from a trained panel. Perrin et al. (2008, 2009) proposed UFP as an alternative to DA for verbalisation to supplement PM when profiling Loire wine using experts. Ross et al. (2012) studied the effect of serving temperature on red wine using PM.

Torri et al. (2013) conducted PM while studying the relationship between the sensory profiles, consumer liking and the expert quality perception of Italian red wine. It is interesting to note that judges did not have to provide descriptors after PM was performed in this study even though this study was performed the study performed by Perrin et al. (2009).

2.5.3.1. Validation of projective mapping and Napping for alcoholic beverage description

A number of studies were conducted to validate and test PM for the evaluation of alcoholic beverages. Savidan et al. (2015) compared PM results obtained when paper versus computer screens were used for data capturing while evaluating beer with consumers. Similar results were obtained with paper and computer screens as capturing systems.

Hopfer and Heymann (2013) tested the effect of the: (1) paper shape; (2) the number of replicated tastings done by sensory judges and (3) proposed the people performance index (PPI), as a measure to evaluate individual sensory judges. This study was conducted on red wine blends. Louw et al. (2013, 2014, 2015a, 2015b) validated the restricted version of PM, called Napping for the sensory analysis of brandy using trained panels. Hopfer and Heymann (2013) found that the product representation was dependant on the provided space. Louw et al. (2015a) found that, when profiling brandy with Napping, similar results were obtained when rectangular, square and round paper sheets were used.

Vidal et al. (2014) investigated the number of consumers needed, for a PM experiment, to produce a stable sensory space, analysing many different matrices including wine and found that 50 consumers are sufficient.

Liu et al. (2016) used Napping and FP to study small differences in model wine solutions using FP and Napping. It was found that training with regards to the method or the product space improved the quality of data obtained. It was also noted that Napping highlighted qualitative differences between samples where FP provided more detail about quantitative sample differences.

2.5.4 Polarised sensory positioning (PSP), polarised projective mapping (PPM) and pivot profile (PP)

Only a few examples of reference-based rapid sensory methods used for wine evaluation could be found in literature by the authors. When PP was introduced by Thuillier et al. (2015) a case study on Champagne was conducted.

Crous (2016) compared PSP to DA using a trained panel to gain insights in the sensory properties of Chenin Blanc wines made from old vine grapes. The sample configurations obtained with PSP and DA were similar. Crous note that PSP is useful for a broad description of the sample set relative to the poles, if a detailed profile of each wine is required then DA is more suitable.

Wilson et al. (2018) used PPM to explore the possibility of data aggregation when profiling South African Chenin Blanc wines. Good results were obtained with PPM and it was shown to be suitable for this purpose. However, further investigation is required to investigate other procedures to determine the positioning of the poles on the sheet, prior to evaluation, to avoid distortion of the sensory space that is at risk when poles are simply placed in a triangle.

In all of these studies it was shown that reference-based rapid sensory method make a unique contribution to the rapid method category specifically for the evaluation of wine but require further study to identify and address methodological limitations.

2.5.5 Comparative rapid sensory method studies

Since a number of rapid sensory methodologies has been proposed as alternatives for DA the question how well these methods compare to DA and each other had to be answered. Several studies were conducted to answer this question by testing if: (1) sufficient product discrimination is achieved for the matrix analysed and (2) the rapid method is suitable for the type of panel used by comparing rapid methods to DA and each other.

The modern view that consumers could provide more detailed information than only hedonic information related to preference and liking has been tested. Rapid sensory methods, where product characteristics have to be recognised and verbalised, has been performed using consumers as sensory judges on products with varying matrix complexities (Valentin et al., 2012; Varela & Ares 2012).

Similar results to DA were reported in literature using (1) consumer panels to evaluate food products such as chocolate using PM (Kennedy & Heymann, 2009) and oil emulsions by means of RATA (Oppermann et al., 2017); (2) trained panels evaluating chocolate using PM (Risvik et al., 1994); fruit dairy products using FP (Delarue & Sieffermann, 2004); breakfast cereals by means of sorting (Cartier et al., 2006) and honeybush tea using PM (Moelich et al., 2018) and (3) panels with different degrees of training when hot served food was evaluated by means of FP and PM (Albert et al., 2011).

In addition a few comparative studies (Table 2) performed on alcoholic beverages also concluded that DA results compared well to rapid method results when rapid methods were performed by: (1) consumers performing the sorting task (Bester, 2011), CATA (Ares et al., 2015; Weightman, 2017; Lezaeta et al., 2018) and free listing (Mapheleba, 2018); (2) trained panels performing Napping (Louw et al., 2013), PSP (Crous, 2016) and sorting (Bester, 2011); (3) panels with different levels of training performing the sorting task (Bester et al., 2011) and Napping (Torri et al., 2013) and (4) industry professionals or experts performing Napping with

UFP (Perrin et al., 2008), directed sorting (Johnson et al., 2013), free comments (Lawrens et al., 2013) and FCP (Coulon-Leroy et al., 2017).

Although rapid sensory methods provide useful alternatives to DA, it will never replace DA. Quantitative intensity score data is produced by DA and can provide a more detailed description of products partially due the training and alignment of the sensory judges and is, therefore, statistically more robust (Albert et al. 2011). This statement, even though it is valid for food products still has to be verified for alcoholic beverage analysis, e.g. Louw et al. (2013, 2015a, 2015b) showed that Napping is better adapted for brandy description than DA. Campo et al. (2008) made a similar observation showing that “frequency of attribute citation” also known as “pick-*k* attributes” (Valentin et al., 2012), which is an adapted version of CATA, might be more suitable for wine sensory evaluation.

Most novel rapid methodologies provide citation frequencies, and the assumption is made that frequently cited attributes are more intense than attributes cited rarely. However, when DA is performed the sensory judges are restricted to use a relatively small number of attributes. When complex matrices such as wine are evaluated intensity scores might be less important than simply noting whether an attribute is present or not (Campo et al., 2008). In this case, rapid sensory methods might provide a richer vocabulary than DA. This however lead to the difficult and time consuming task associated with rapid methods when the verbalisation or labelling attributes have to be processed (Veinand et al., 2011; Valentin et al. 2012, Varela & Ares., 2012;).

In addition, although most studies where rapid methods were compared to each other noted that similar results were obtained (Teillet et al., 2010; Veinand et al., 2011; Ares et al., 2013; Cadena et al., 2014; Reinbach et al., 2014; Fleming et al., 2015; Fonseca et al., 2015; Liu et al., 2016; Esmerino et al., 2017; Vidal et al., 2017; Deneulin et al., 2018; Lezaeta et al., 2018; Liu et al., 2018), the cognitive task and logistical possibilities differ slightly for the different rapid methods (Valentin et al., 2012, Varela & Ares, 2012).

It is, therefore, important to carefully consider the pros and cons of the different descriptive sensory methods when the objectives and aims of a sensory study are formulated in order to choose a fit-for-purpose method (Fig. 8).

2.6 Conclusions

Rapid sensory analysis methods play an increasingly important role in the field of sensory and consumer science in both the broader food and beverage industry as well as the wine and alcoholic beverage industry. Continuously adaptation and validation are performed for different product matrices. However, further studies on validation and optimisation are still needed for many product matrices, e.g. complex matrices such as wine, specifically methods falling into the reference-based category.

Reference-based methods can be tested for their suitability as benchmarking tools and evaluation measures for product consistency between batches since a direct comparison between the reference and evaluated sample is made by the sensory judge. Sample sets consisting, of many different products can be evaluated, using reference-based methods over multiple sessions, aggregating data by using the same product as reference for all sessions. However, studies to: (1) identify the limitation of the reference-based methods and (2) compare the different reference-based methods to identify the best method and sensory panel for specific wine applications are needed.

In all the studies where rapid methods were compared to DA it was shown that similar results were obtained when a simple broad description of the sensory properties of a product set was required. These methods are, therefore, ideal screening tools and even profiling tools if detailed information is not required. In addition, it was reported that the results obtained from different rapid method were similar.

However, in most scenarios, a specific rapid method might be more suitable due to the objectives of the study. Another consideration or restriction might be logistical and practical implications. When PM and sorting are conducted it was noted that between 8 and 20 products should be evaluated to get reliable results that make sense. It is, therefore, not suitable for small or large samples sets. Reference-based methods like PSP, PP and PPM can be used to analyse large sample set over multiple sessions. CATA and RATA can be used for small sample sets, but can only be used if enough information of the product attributes is known to establish a predetermined list of attributes. Even though many studies have been conducted where rapids methods were compared using different types of panels, further studies could highlight matrix and panel type specific pros and cons if the same panel evaluates the same product using different methods.

2.7 References

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. *Poster presented at the 7th Pangborn Sensory Science Symposium.* 16 August 2007, Minneapolis, USA. Delegate Manual, pp. 12–418.

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.

Alencar, N. M. M., Ribeiro, T. G., Barone, B., Barros, A. P. A., Marques, A.T.B., & Behrens, H. J. (2018). Sensory profile and check-all-that-apply (CATA) as tools for evaluating and characterizing Syrah wines aged with oak chips. *Food Research International*, (March), 1–9.

Ares, G., Antúnez, L., Bruzzone, F., Vidal, L., Giménez, A., Pineau, B., Beresford, M. K., Jin, D., Paisley, A. G., Chheang, S. L., Roigard, C. M., & Jaeger, S. R. (2015). Comparison of sensory product profiles generated by trained assessors and consumers using CATA questions : Four case studies with complex and / or similar samples, *Food Quality and Preference 45*, 75–86.

Ares, G., Bruzzone, F., Vidal, L., Cadena, R. S., Giménez, A., Pineau, B., Hunter, D. C., Paisley, A.G., & Jaeger, S. R. (2014). Evaluation of a rating-based variant of check-all-that-apply questions: Rate-all-that-apply (RATA). *Food Quality and Preference*, *36*, 87–95.

Ares, G., Deliza, R., Barreiro, C., Giménez, A., & Gámbaro, A. (2010a). Comparison of two sensory profiling techniques based on consumer perception. *Food Quality and Preference*, *21*, 417–426.

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.

Ares, G., Giménez, A., Barreiro, C. & Gámbaro, A., (2010b). Use of an open-ended question to identify drivers of liking of milk desserts: comparison with preference mapping techniques. *Food Quality and Preference, 21*, 286–294.

Ares, G., Varela, P., Rado, G. & Giménez, A., (2011a). Identifying ideal products using three different consumer profiling methodologies. Comparison with external preference mapping. *Food Quality and Preference, 22*, 581–591.

Ares, G., Varela, P., Rado, G. & Giménez, A., (2011b). Are consumer profiling techniques equivalent for some product categories? The case of orange-flavoured powdered drinks. *International Journal of Food Science & Technology, 46*, 1600–1608.

Ballester, J., Abdi, H., Langlois, J., Peyron, D., & Valentin, D. (2009). The odor of colors: Can wine experts and novices distinguish the odors of white, red, and rosé wines? *Chemosensory Perception, 2*, 203–213.

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., 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.

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

Bárcenas, 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.

Bécue-Bertaut, M. (2014). Tracking verbal-based methods beyond conventional descriptive analysis in food science bibliography. A statistical approach. *Food Quality and Preference*, *32*, 2–15.

Bester, I. (2011). Classifying South African Chenin blanc wine styles. MSc thesis, Stellenbosch University. http://hdl.handle.net/10019.1/18048

Botha, A. (2015). The use of different oak products during the fermentation and ageing of Chenin Blanc : sensory properties, perceived quality, and consumer preference. MSc thesis, Stellenbosch University. http://hdl.handle.net/10019.1/98006

Cadena, R. S., Caimi, D., Jaunarena, I., Lorenzo, I., Vidal, L., Ares, G., Deliza, R., & Giménez, A. (2014). Comparison of rapid sensory characterization methodologies for the development of functional yogurts. *Food Research International*, *64*, 446–455.

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, S., & Pagès, J. (2009). A factorial approach for sorting task data (FAST). *Food Quality and Preference*, 20, 410–417.

Cadoret, M., Lê, S., & Pagès, J. (2011). Statistical analysis of hierarchical sorting data, *Journal of Sensory Studies, 26*, 96–105.

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.

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.

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.

Coulon-Leroy, C., Symoneaux, R., Lawrence, G., Mehinagic, E., & Maitre, I. (2017). Mixed Profiling : A new tool of sensory analysis in a professional context . Application to wines. *Food Quality and Preference*, *57*, 8–16.

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

Corsi, A. M., Cohen, J., Lockshin, L., & Williamson, P. (2017). Testing lexical equivalences for wine flavours in emerging markets : Do hawthorns taste like blackberries ? *Food Quality and Preference*, *62*, 296–306.

Coste, A., Sousa, P., & Malfeito-ferreira, M. (2018). Wine tasting based on emotional responses : An expedite approach to distinguish between warm and cool climate dry red wine styles. *Food Research International*, *106*, 11-22.

Coxon, A.P.M. (1999). *Sorting Data: Collection and Analysis.* Thousand Oaks, CA: Sage.

Crous, R. (2016). The sensory characterisation of old - vine Chenin blanc wine : an exploratory study of the dimensions of quality. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/100114

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*, 826-834.

Delarue, J., & Sieffermann, 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*, 383–392.

Dehlholm, C., Brockhoff, P.B., Meinert, L., Aaslung, M.D., & Bredie, W.L.P. (2012a). 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. (2012b). 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.

Esbensen, K.H. (2002). Multivariate Data Analysis in practice – an introduction to multivariate data analysis and experimental design (5th ed.). Oslo: CAMO Process Ås

Escofier, B., & Pagès, J. (1990). Multiple factor analysis. *Computational Statistics & Data Analysis, 18*, 121–140.

Etaio, I., Albisu, M., Ojeda, M., Gil, P. F., Salmerón, J., & Elortondo, F. J. P. (2010). Sensory quality control for food certification: A case study on wine. Method development. *Food Control*, *21*, 533–541.

Fleming, E. E., Ziegler, G. R., & Hayes, J. E. (2015). Check-all-that-apply (CATA), sorting, and polarized sensory positioning (PSP) with astringent stimuli. *Food Quality and Preference*, *45*, 41–49.

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.*

Fonseca, F. G. A., Esmerino, E. A., Filho, E. R. T., Ferraz, J. P., Cruz, A. G., & Bolini, H. M. A. (2016). Novel and successful free comments method for sensory characterization of chocolate ice cream : A comparative study between pivot profile and comment analysis. *Journal of Dairy Science, 99*, 3408–3420.

Franco-Luesma, E., Sáenz-navajas, M., Valentin, D., & Ballester, J. (2016). Study of the effect of H_2S , MeSH and DMS on the sensory profile of wine model solutions by Rate-All-That-Apply (RATA). *FRIN*, *87*, 152–160.

Gower, J. C. (1975). Generalized procrustes analysis. Psychometrika, 40, 33–51. Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification, 2*, 193–218.

Hanekom, E. (2012). Chemical, sensory and consumer profiling of a selection of South African Chenin blanc wines produced from bush vines. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/71812

Heymann, H., Hopfer, H., & Bershaw, D. (2013). An exploration of the perception of minerality in white wine by projective mapping and descriptive analysis, *Journal of Sensory Studies, 29,* 1–13

Honoré-chedozeau, C., Lelièvre-desmas, M., Ballester, J., Chollet, S., & Valentin, D. (2017). Knowledge representation among assessors through free hierarchical sorting and a semi-directed interview : Exploring Beaujolais wines, *Food Quality and Preference*, *57*, 17–31.

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.

Hough, G., & Ferraris, D. (2010). Free listing: A method to gain initial insight of a food category. *Food Quality and Preference*, *21*, 295–301.

Jaeger, S. R., Lee, P., & Ares, G. (2018). Product involvement and consumer food-elicited emotional associations : Insights from emoji questionnaires. *Food Research International*, *106*, 999–1011.

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.

Kelly, G. A. (1955). *The psychology of personal constructs*. New York: Norton.

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.

King, M.C., Cliff, M.A., & Hall, J.W., (1998). Comparison 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.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional Scaling*. Thousand Oaks: SAGE publications, Inc.

Lancaster, B., & Foley, M., (2007). Determining statistical significance for choose-all- that-apply question responses. *Seventh pangborn sensory science symposium*, Minneapolis, USA.

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.

Lawless, H. T., & Heymann, H. (2010). *Sensory evaluation of food. Principles and practices (2nd ed.)*. New York: Springer.

Lawrence, G., Symoneaux, R., Maitre, I., Brossaud, F., Maestrojuan, M., & Mehinagic, E. (2013). Using the free comments method for sensory characterisation of Cabernet Franc wines: Comparison with classical profiling in a professional context. *Food Quality and Preference*, *30*, 145–155.

Lezaeta, A., Bordeu, E., Næs, T., & Varela, P. (2017). Exploration of consumer perception of Sauvignon Blanc wines with enhanced aroma properties using two different descriptive methods. *Food Research International*, *99*, 186–197.

Lezaeta, A., Bordeu, E., Agosin, E., Pérez-correa, J. R., & Varela, P. (2018). White wines aroma recovery and enrichment : Sensory-led aroma selection and consumer perception. *Food Research International*, *108*, 595–603.

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.

Liu et al., 2016 Liu, J., Schou, M., Di, R., Giacalone, D., & Bredie, W. L. P. (2016). Performance of Flash Profile and Napping with and without training for describing small sensory differences in a model wine. Food Quality and Preference, 48, 41–49.

Louw, L., Oelofse, S., Næs, T., Lambrechts, M., Van Rensburg, P., & Nieuwoudt, H. (2015a). Optimisation of the partial napping approach for the successful capturing of mouthfeel differentiation between brandy products. *Food Quality and Preference*, *41*, 245–253.

Louw, L., Oelofse, S., Næs, T., Lambrechts, M., Van Rensburg, P., & Nieuwoudt, H. (2015b). The effect of tasting sheet shape on product configurations and panellists ' performance in sensory projective mapping of brandy products. *Food Quality and Preference*, *40*, 132–136.

Louw, L., Oelofse, S., Næs, T., Lambrechts, M., Van Rensburg, P., & Nieuwoudt, H.H. (2014). Trained sensory panellists' response to product alcohol content in the projective mapping task: Observations on alcohol content, product complexity and prior knowledge. *Food Quality and Preference, 34*, 37-44.

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

Mapheleba, A. (2018). Perceptions of South African Chenin blanc wine among Consumers and Industry professionals. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/103719

Meyners, M., Castura, J. C., & Carr, B. T. (2013). Existing and new approaches for the analysis of CATA data. *Food Quality and Preference*, *30*, 309–319.

Moelich, E. I. (2018). Development and validation of prediction models and rapid sensory methodologies to understand intrinsic bitterness of Cyclopia genistoides, PhD dissertation, Stellenbosch University, http://hdl.handle.net/10019.1/103543

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.

Næs, T., Berget, I., Hovde, K., Ares, G., & Varela, P. (2017). Estimating and interpreting more than two consensus components in projective mapping : INDSCAL vs multiple factor analysis (MFA). *Food Quality and Preference*, *58*, 45–60.

Navajas et al., 2014 Sáenz-navajas, M., Ballester, J., Peyron, D., & Valentin, D. (2014). Extrinsic attributes responsible for red wine quality perception : A cross-cultural study between France and Spain, *35*, 70–85.

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. (2011). Recovery of subsampled dimensions and configurations derived from napping data by MFA and MDS. Attention, *Perception, & Psychophysics, 73*, 1266–1278.

Oppermann, A. K. L., Graaf, C. De, Scholten, E., Stieger, M., & Piqueras-fiszman, B. (2017). Comparison of Rate-All-That-Apply (RATA) and Descriptive sensory Analysis (DA) of model double emulsions with subtle perceptual differences. *Food Quality and Preference*, *56*, 55–68.

Pagès, J. (2003). Recueil direct de distances sensorielles: application a` l'é´valuation de dix vins blancs de Val de Loire. *Science des Aliments, 23*, 679–888.

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.

Parente, M.E., Manzoni, A.V., & G. Ares. (2011). External preference mapping of commercial antiaging creams based on consumers' responses to a check-all-that-apply question. *Journal of Sensory Studies, 26*, 158–166.

Parr, W. V., Green, J. A., White, K. G., & Sherlock, R. R. (2007). The distinctive flavour of New Zealand Sauvignon blanc: Sensory characterisation by wine professionals. *Food Quality and Preference, 18*, 849–861.

Parr, W. V., Valentin, D., Green, J. A., & Dacremont, C. (2010). Evaluation of French and New Zealand Sauvignon wines by experienced French wine assessors. *Food Quality and Preference*, *21*(1), 56–64.

Parr, W. V., Ballester, J., Peyron, D., Grose, C., & Valentin, D. (2015). Perceived minerality in Sauvignon wines: Influence of culture and perception mode. *Food Quality and Preference*, *41*, 121–132.

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.

Picard, D., Dacremont, C., Valentin, D., & Giboreau, A. (2003). Perceptual dimensions of tactile textures. *Acta Psychologica*, 114, 165–184.

Picard, M., Tempere, S., de Revel, G., & Marchand, S. (2015). A sensory study of the ageing bouquet of red Bordeaux wines: A three-step approach for exploring a complex olfactory concept. *Food Quality and Preference*, *42*, 110–122.

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*, 27-34.

Popper, R., Abdi, H., Williams, A., & Kroll, B.J. (2011). Multiblock Hellinger analysis for creating perceptual maps from check-all-that-apply questions. In: *9th Pangborn Sensory Science Symposium.* September 7th, 2011, Toronto, CA.

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

Reinbach, H.C., Giacalone, D., Ribeiro, L.M., Bredie, W.L.P., & Frøst, M.B. (2014). Comparison of three sensory profiling methods based on consumer perception: CATA, CATA with intensity and Napping®. *Food Quality and Preference, 32*, 160-166.

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. & Escouffier, 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 PM by a trained panel. *Journal of Sensory Studies, 27*, 463-470.

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 Preference, 21*, 881-892.

Savidan, C. H., & Morris, C. (2015). Panelists' Performances and Strategies in Paper-Based and Computer-Based Projective Mapping. *Journal of Sensory Studies*, *30*(2), 145–155.

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

Soufflet, I., Calonnier, M., & Dacremont, C. (2004). A comparison between industrial experts' and novices' haptic perception organization: a tool to identify descriptors of handle of fabrics. *Food Quality and Preference, 15*, 689–699.

Symoneaux, R., Galmarini, M.V., & Mehinagic. (2012). Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples. *Food Quality and Preference, 24,* 59–66.

Takane, Y. (1982). IDSORT: an individual differences multidimensional scaling program for sorting data. *Behavior Research Methods & Instrumentation, 14*, 546.

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.

Thuillier, B., Valentin, D., Marchal, R., & Dacremont, C. (2015). Pivot© profile: A new descriptive method based on free description. *Food Quality and Preference*, *42*, 66–77.

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.

Tomic, O., Berget, I., & Næs, T. (2015). A comparison of generalised procrustes analysis and multiple factor analysis for projective mapping data. *Food Quality and Preference*, *43*, 34–46.

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*, 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 Antwerpen, L. (2012). Chemical and sensory profiling of dry and semi-dry South African Chenin blanc wines. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/71853

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.

Vidal, L., Cadena, R. S., Antúnez, L., Giménez, A., Varela, P., & Ares, G. (2014). Stability of sample configurations from projective mapping: How many consumers are necessary? *Food Quality and Preference*, *34*, 79–87.

Vidal, L., Antúnez, L., Giménez, A., Medina, K., Boido, E., & Ares, G. (2017). Sensory characterization of the astringency of commercial Uruguayan Tannat wines. *Food Research International*, *102*, 425–434.

Vidal, L., Antúnez, L., Rodríguez-haralambides, A., Giménez, A., Medina, K., Boido, E., & Ares, G. (2018). Relationship between astringency and phenolic composition of commercial Uruguayan Tannat wines : Application of boosted regression trees. *Food Research International*, *112*, 25–37.

Vidal, L., Giménez, A., Medina, K., Boido, E., & Ares, G. (2015). How do consumers describe wine astringency ? *FRIN*, *78*, 321–326.

Weightman, C. J. (2018). Consumer attitudes and sensory perceptions of wine : A South African cross-cultural study. PhD dissertation, Stellenbosch University, http://hdl.handle.net/10019.1/103475

Williams, A.A,. & Langron, S.P. (1984). The use of free-choice profiling for the evaluation of commercial ports. *Journal of the Science of Food and Agriculture, 35*, 558–568.

Wilson, C., Brand, J., Toit, W., & Buica, A. (2018). LWT - Food Science and Technology Polarized projective mapping as a rapid sensory analysis method applied to South African Chenin Blanc wines. *LWT - Food Science and Technology*, *92*, 140–146.

Chapter 3

Research results

In search of suitable rapid sensory methods for wine profiling using industry professionals:

A comparison of Free Sorting, Projective mapping, Check-All-That-Apply and Rate-All-That-Apply to descriptive analysis

This manuscript is in preparation for publication in **The South African Journal of Enology and Viticulture**

3. Research results

In search of suitable rapid sensory methods for wine profiling using industry professionals: A comparison of Free Sorting, Napping, Check-All-That-Apply and Rate-All-That-Apply to Descriptive Analysis

J. Brand[a,b*], T. Næs[c], M. Kidd[d], M. A. Vivier[b], D. Valentin[e], H.H. Nieuwoudt[b]

[a]*Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[b]*Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[c]*Nofima AS, Raw material and Process Optimization, Oslovegen 1, P.O. Box 210, N-1431, Ås, Norway*

[d]*Centre for Statistical Consultation, Department of Statistics and Actuarial Sciences, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[e]*Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, CNRS, INRA, Univ. Bourgogne, Franche-Comté, F-2100, Dijon, France*

Keywords: Descriptive analysis, sorting, Napping, projective mapping, check-all-that-apply, rate-all-that-apply, wine, industry professionals

Abstract

Rapid sensory analysis techniques are gaining popularity as alternatives for traditional descriptive analysis (DA) to evaluate the sensory properties of wine. The suitability of these methods for white wine profiling has not been studied in depth. In this study, four rapid sensory methods were tested and compared to DA. Wine industry professionals profiled 10 Chenin Blanc wines by means of free sorting, Napping, check-all-that-apply (CATA) and rate-all-that-apply (RATA). The same wines were analysed by a trained panel conducting DA. The sample configurations of the multivariate sensory maps obtained from the different methods were compared by inspection, HCA and RV coefficients. In addition, the attributes obtained from the different methods to describe the wines were compared by means of Multiple Factor Analysis (MFA) and inspection. The sample configurations of the multivariate sensory maps obtained with the different rapid methods were similar to the map constructed with DA data with RV coefficients ranging from 0.69 to 0.83. CATA and sorting provided the best separation between the different Chenin Blanc wine styles assessing the overlap of the confidence ellipses on the multivariate sensory maps. Napping and RATA were perceived as the most difficult methods and sorting and CATA the easiest to perform. Therefore, sorting and CATA were identified as

the most suitable rapid methods to use as alternatives for DA for rapid profiling of white wine using industry professionals.

3.1 Introduction

Rapid sensory methods that are less time consuming and more cost-effective than classical descriptive analysis (DA) have received ample attention in recent research. These methods are attractive to the food and beverage industry and are becoming increasingly popular during wine sensory evaluation. The main reasons are: (1) training is not required and (2) they are suitable for profiling products using different types of panels including consumers, trained panellists or product experts and industry professionals. Check-all-that-apply (CATA, Adams et al., 2007) and it's variants, rate-all-that-apply (RATA) also known as CATA with intensity rating (RATA, Reinbach et al., 2014; Ares et al., 2014), Sorting (Lawless et al., 1995), and projective mapping (PM, Risvik et al., 1994) techniques including Napping (Pagès 2003, 2005) form part of the rapid method category.

3.1.1 Check-all-that-apply (CATA)

CATA is based on a multiple choice approach where participants select the appropriate choices from a list to best answer the question asked. It was first used in marketing research to study different brands as perceived by consumers (Coomb, 1964). When CATA is used as a rapid sensory method, the list consists of words, phrases or emoji's (Jaeger et al., 2018). These words can be sensory attributes, hedonic terms or emotional phrases. Sensory judges receive products according to a monadic serving order, where every judge receives one sample at a time and are asked to select the terms that best describe the sensory properties of the product.

CATA data is collected and tabulated in a contingency table where the number of times a specific attribute is cited for a wine is counted. Correspondence analysis (CA) is typically performed to obtain a multivariate sensory map illustrating the sensory attributes associated with each product as well as the similarities and difference between products (Valentin et al., 2012).

When consumers are used there is typically no restriction on the number of terms that the participant could use. A variant called "pick-*k* attributes", where participants choose the "*k"* most important" attributes, has been used by Chollet and Valentin (2000) to profile wine with industry professionals as sensory judges. In this case the main characteristics of the products are highlighted rather than obtaining a full detailed sensory description.

The popularity of CATA as wine profiling tool increased recently. Alencar et al. (2018) used CATA to profile Syrah wine aged with oak chips using consumers. Vidal et al. (2017, 2018b) used CATA performed by trained panels to investigated the astringency of Tannat wines. Lezaeta et al. (2017, 2018) evaluated consumer perception of white wines, enriched chemically to enhance the aroma profile, using CATA. Coste et al. (2018) used CATA performed by consumers to distinguish between warm and cool climate dry red wine styles. Corsi et al. (2017) tested lexical equivalences between Chinese and Western consumers describing wine flavours. Ares et al. (2015) compared the differences in sensory profiles obtained for CATA data when consumers and trained assessors are used to profile white wine, as well as other products. Botha (2015) compared expert and consumer perception of Chenin Blanc wine subjected to different wood treatments.

This increase in popularity can be attributed to the fact that CATA is a fast and simple technique due to the fact that intensity is not rated and results compare well to DA (Ares et al., 2010; Dooley et al., 2010; Valentin et al, 2012). However, when analysing CATA data, the assumption is made that attributes that were cited frequently had higher intensities than attributes cited only a few times. In some cases, this assumption cannot be made. Reinbach et al. (2014) introduced a variant of CATA, called CATA with intensity, independently introduced by Ares et al. (2014) as rate-all-that-apply (RATA).

3.1.2 Rate-all-that-apply (RATA)

When RATA is conducted, a second step, where the intensities of the selected attributes are rated, is performed after CATA. Reinbach et al. (2014) used a 15-point scale to evaluate beers using consumers. Ares et al. (2014) used a 3-point scale (“low”, “medium” and “high”) to rate intensity when consumers evaluated bread samples and gummy lollies. These authors, in another study, used a 5-point scale to rate applicability on a line scale ranging from “slightly applicable” to “very applicable” when milk desserts and yoghurt labels were evaluated. Franco-Luesma et al. (2016) used a 7-point scale ranging from “not intense” (1) to “very intense” (7) to profile wine model solutions spiked with volatile sulphur compounds responsible for off-odours with an expert panel consisting of Oenology students.

A few studies compared CATA to RATA. Vidal et al. 2018 compared CATA to RATA when consumers evaluated fruits. Reinbach et al. (2014) compared CATA to RATA and Napping as performed by consumers when evaluating beers. In both studies, it was concluded that the CATA and RATA results obtained were similar. In addition, Oppermann et al. (2017) compared RATA to DA when evaluating model food emulsions and found that the results obtained with RATA were similar to those obtained with DA in terms of the different multivariate sensory map configurations.

The main advantage of CATA and RATA is that it is possible to aggregate data (combine data sets) captured over multiple sessions due to the monadic presentation of samples to the sensory judges. However, in some cases, the research question is better answered by presenting samples simultaneously and comparing the samples to each other to describe their similarities and differences. Sorting and projective mapping or Napping can be used for that purpose (Valentin et al., 2012; Varela & Ares, 2012).

3.1.3 Sorting

Lawless et al. (1995) used free sorting first in the field of sensory science while studying the sensory perception of different types of cheese. During free sorting sensory judges are asked to group products in terms of similarities and dissimilarities. Products with similar sensory characteristics are grouped together. A second step, called "verbalisation" (Chollet et al., 2011) or "labelling" (Bécue-Bertaut & Lê, 2011), can be conducted where descriptions are provided to explain the groupings. Descriptors are provided to describe the collective sensory characteristics of all the samples in the group.

Sorting data is captured by means of a distance or similarity matrices for each judge. The number of times each pair of samples are grouped together is counted to obtain a distance or similarity matrix with the data from all the judges. The most commonly used statistical methods to analyse the grouping data are DISTATIS performed directly on the distance matrices of the individual sensory judges or MDS performed on the sum over all the matrices. The descriptors can be projected onto the graphs, namely the MDS plot or DISTATIS compromise map, using correlations coefficients calculated for each product. These correlation coefficients are calculated between the sum of the citations of every attribute over all the judges and the coordinates of the graph (Cartier et al., 2006). If the experimenter wants to investigate the descriptors only, CA can be performed on a contingency table that is compiled in the same way that contingency tables are compiled when CATA is conducted (Picard et al., 2003; Soufflet et al., 2004; Valentin et al., 2012). Techniques such as multiple correspondence analysis (MCA) used by Cadoret et al. (2009) and hierarchical MFA by Bécue-Bertaut and Lê (2011) are also options for sorting data analysis.

Sorting has been used extensively to investigate the sensory properties of wine (Piombino et al., 2004; Ballester et al., 2005, 2013; Abdi & Valentin., 2007; Campo et al., 2008; Bécue-Bertaut & Lê, 2011; Johnson et al., 2013; Franco-Luesma et al., 2016; Honoré-Chedozeau et al., 2017).

3.1.4 Projective mapping and Napping

Projective mapping (PM) methods have been introduced under different names and variants. In 1983 Dun-Rankin introduced placing and in 1994 Goldstone the spatial arrangement procedure (SAP) to the field of psychology research. Risvik et al. (1994, 1997) used the name projective mapping (PM) and Pagès (2003, 2005) Napping in the field of sensory science. When PM or Napping is conducted sensory judges are asked to arrange the products on a piece of paper according to similarity. Similar samples should be placed close together and different samples far apart. A second step where sensory judges describe the arrangement of samples by providing descriptions for each sample is frequently used in the field of sensory science and was called ultra flash profile (UFP) by Perrin et al. (2008, 2009).

PM data are captured by tabulating the *X* and *Y* coordinates of each sample provided by each judge. The most frequently used statistical method to analyse PM data is MFA where the *X* and *Y* coordinates provided by each judge is taken into account. The data obtained from the descriptors can be compiled in a contingency table as for CATA data analysis. The descriptor data can be projected onto the MFA multivariate sensory map or CA can be conducted to visualise the descriptor data. As with sorting all samples are presented simultaneously and a holistic intuitive map of the similarities between the samples can be formed prior to verbalisation of the specific characteristics responsible for the difference and similarities. However, for PM methods each sample is described individually in contrast to sorting where the group of samples are described together. Pagès et al. (2010) introduced sorted Napping a technique combining sorting and Napping.

PM techniques were used to profile wines by Torri et al. (2013) comparing expert and consumer results. Hopfer and Heymann (2013) investigated the effect of the shape of the PM sheet and replicated tastings performed by the same judge when profiling wine. Louw et al. (2013, 2015) investigated and validated Napping as a profiling tool for high alcohol beverages using a trained panel. Vidal et al. (2014) investigated the number of consumers needed to perform Napping on various products including red wine and champagne. Liu et al. (2016) described small sample difference in model wine. In a recent study, Heatherly et al. (2019) investigated the relationship between colours, shapes and wine odours using PM. These studies showed that Napping has been applied and validated for sensory analysis of alcoholic beverages, including wine a number of times.

3.1.5 Comparison of rapid sensory methods

Rapid sensory methods have been compared, tested and validated for their suitability to profile various foodstuffs mainly using consumers as sensory judges (Ares et al, 2010; Cadena et al., 2014; Reinbach et al., 2014; Lezaeta et al., 2017; Oppermann et al., 2017; Liu et al., 2018). Dehlholm et al. (2012a) compared free multiple sorting, partial Napping and flash profile to conventional profiling when evaluating liver pâtés with a trained panel. Even though rapid methods have been tested for their suitability to profile complex matrices including wine no single study has been conducted that compare the frequently used methods against each other and DA with industry professionals performing the rapid sensory analysis.

The wine industry has an interest to profile wine using industry professionals as part of product development prior to profiling by consumers. Furthermore, a large number of wine aroma wheels are available that can be used by sensory experimenters and panel leaders as pre-determined lists during the verbalisation step of a rapid sensory method.

The aim of this study was to compare frequently used rapid methods to DA as sensory profiling tools for applications in the wine industry and wine research using resources available within the wine industry. Free sorting, CATA, RATA and Napping conducted by industry professionals were compared to DA performed by a trained panel. A pre-determined list of attributes was used as the verbalisation step for all the rapid methods. The similarity between multivariate sensory maps, the attributes used to describe products and how they relate to different Chenin Blanc styles, the easiness/difficulty and the time required to perform the method measured were here.

3.2 Materials and methods

3.2.1 Samples

In this study 10 Chenin Blanc wines were evaluated. The wines were chosen to span the sensory space of South African Chenin Blanc wine based on knowledge from previous studies (Bester, 2011; Hanekom, 2012; Van Antwerpen, 2012) and the recommendations of South African wine industry professionals. All the wines were produced in South Africa and certified by the South African Wine and Spirit Board (see Table 1). Wines were stored in the dark at 15^{o}C prior to sensory analysis.

3.2.2 Panels

Two separate panels, A and B, were recruited for this study. Panel A performed descriptive analysis (DA) and consisted of trained judges, two male and 10 female judges between the ages of 24 and 57 (average age: 40). All judges had at least two years of experience in white wine sensory analysis and were remunerated for their services. Judges were not screened prior to this study, but were only invited to participate if they produced repeatable results and were in consensus with other judges when Chenin Blanc wines were evaluated during previous studies.

Panel B evaluated the wines by means of rapid sensory analysis methods namely: check-all-that-apply (CATA), rate-all-that-apply (RATA), free sorting and Napping, the restricted version of projective mapping (PM, Dehlholm et al., 2012a). This panel consisted of 15 professional qualified winemakers, eight male and seven female judges between the age of 22 and 45 (average age: 35). The judges on panel B were not remunerated for tasting on the panel, they participates out of interest to gain more experience in wine sensory evaluation and wine tasting. Eight of the 10 judges had more than 10 years' experience working in the wine industry. The experience of the other five judges varied between 4 and 7 years of technical wine tasting experience, which included their training as students and work experience.

TABLE 1
Summary of the vinification parameters and chemical analysis after bottling of the wines subjected to PP and FC sensory analyses.

Wine code	Origin	Vintage	Alc[a] % (v/v)	RS[b] (g/L)	pH	TA[c] (g/L)	Vinification and aging
Chenin Blanc wines							
A	Paarl	2012	14.0	4.2	3.54	6.2	Partially barrel fermented, aged in older barrels.
B	Paarl	2012	12.5	3.5	3.36	6.5	Tank fermented to be consumed as a young wine while still fresh.
C	Stellenbosch	2012	13.5	3.0	3.69	5.4	Fermented from old bush vine grapes. Matured on the lees in Burgundian barrels for eight months.
D	Swartland	2012	12.5	2.9	3.45	6.0	Tank fermented fresh and fruity style.
E	Paarl	2012	13.5	2.7	3.29	6.4	Tank fermented at 12 – 15°C, kept on the lees for three months.
F	Paarl	2012	12.5	3.2	3.50	6.2	Cold tank fermentation. Fresh citrus and fruity aromas.
G	Western Cape	2012	12.5	6.5	3.30	6.2	Tank fermented fresh and crisp with fruity flavours.
H	Coastal	2012	13.5	2.7	3.37	5.7	Tank fermented unwooded Chenin Blanc.
I	Stellenbosch	2012	14.0	3.6	3.44	6.0	Extended skin contact was applied. Fermentation was started in tanks and completed in new (20%), second fill (40%) and third fill (40%) barrels.
J	Stellenbosch	2012	13.5	3.9	3.35	6.4	Blend of tank fermented (50%) and barrel fermented (50%) wine matured "sur lie" for seven months in barrel.

[a]Alcohol, [b]Residual sugar, [c]Titratable acidity

3.2.3 Sensory methodology

3.2.3.1 Descriptive analysis (DA)

Training. Panel training was conducted by means of the consensus method (Lawless & Heymann, 2010). A total of 10 sessions of two hours each with a 10-minute break after an hour was used for training. The panel attended three training sessions per week over four weeks. During the first two sessions all the wines of the specific product set were presented. The judges were instructed to generate as many attributes as they wanted to describe the sensory space highlighting similarities and differences between the samples. Reference standards were prepared from the consensus list of attributes obtained during the first two sessions. The wines and reference standards were presented to the panel. They could evaluate the reference standards and make new suggestions to better describe the attributes where necessary. During the next three sessions consensus on the attributes was achieved and the list of attributes was reduced and finalised (Table 2). The order in which the panel preferred to rate the attributes and the anchors of the scale were established. Rating of the attributes on an unstructured 10 cm line scale anchored at “none” to “intense” were practised and consensus were reached after three sessions.

TABLE 2.
Aroma reference standards presented during DA training representing the final attribute lists.

Descriptor	Reference standard	Amount
Chenin Blanc wines		
Pineapple	Pineapple (fresh)	2 - 4 cm^2 piece
Peach / apricot	Peach (fresh)	4 - 4 cm^2 piece
Citrus	Lemon, orange and grapefruit peel	2 - 2 cm^2 piece of each
Paw-paw	Paw-paw (fresh)	3 - 4 cm^2 piece
Passion fruit	Passion fruit (fresh)	4 pips and a 1cm^2 piece of skin
Stewed dried fruit	Cooked dried fruit (Safari)	1 dried apple, ½ prune, ½ dried peach, 1 dried pear
Honey	Acacia honey (Lune de Miel)	15 mL
Orange marmalade	Seville marmalade (Rhodes)	5 mL
Caramel / burnt sugar	Caramel syrup (St. Dalfour)	20 mL
Buttery toffee	Soft toffees (Toff-o-lux)	1 toffee in boiling water
Oaky	Medium toasted French oak chips (NT Bois, RX South Africa)	2 g
Cooked veg	Canned vegetable brine	10 mL canned bean brine (Rhodes), 10 mL canned asparagus brine (Goldcrest), 10 mL artichoke brine (Goldcrest)
Flinty / mineral	Flint stone	2 flintstones struck against each other
Floral	Honeysuckle essence (Ferminich)	5 drops on cotton wool
Green pepper	Green pepper (fresh)	2 cm^2
Litchi	Litchi (canned, KOO)	1 litchi and 10 mL syrup

Procedure. Sensory judges had to rate the intensities of all the attributes for all the wines on the unstructured 10 cm line scale provided that was anchored at “none” to “intense”. The attributes were listed on the tasting ballot in the order presented in Table 2. The panel evaluated the entire sample set three times on one day. Ten-minute breaks were enforced between the replicates. A monadic sample presentation procedure, presenting one sample at a time was followed.

3.2.3.2 Check-all-that-apply (CATA)

A pre-determined list of terms compiled from data obtained in previous studies (Campo et al., 2008; Campo et al., 2010; Bester, 2011; Hanekom, 2012; Van Antwerpen, 2012) and the help of industry professionals, were provided. The list was constructed to span the sensory space of South African Chenin Blanc wines. Only sensory attributes were used, no quantifiers, e.g. “high”, “medium”, “very”, hedonic or emotional terms or phrases were used. Sensory judges were asked to choose the three to five attributes from the list that best described the sensory characteristics of that specific sample. They were given the option to provide terms that were not on the list if they found the list insufficient. Samples were presented according to a monadic serving protocol, one at a time. This list was used for RATA, and during the verbalisation steps of the free sorting task and the PM exercise.

3.2.3.3 Rate-all-that-apply (RATA)

RATA was performed by first performing CATA followed by a second step where the intensities of the attributes selected were rated on an unstructured 10 cm line scale anchored at “none” to “intense”. Samples were presented according to a monadic serving procedure, one at a time.

3.2.3.4 Free sorting

During the free sorting task all the samples were presented simultaneously. The judges were asked to group samples with similar sensory characteristics together according to their own criteria. They could group as many samples together as they deemed necessary, creating at least two groups and grouping at least two samples together in one of the groups. In other words, each sample could not be in its own group and all the samples could not be in the same group (Chollet et al., 2011).

To explain the categorisation/grouping of the samples a “labelling” (Bécue-Bertaut & Lê, 2011) or “verbalisation” (Chollet et al., 2011) step followed where judges had to provide three to five terms per group. These descriptors had to be chosen from the provided list of attributes to simplify the task of the sensory judges, the data analysis (Lelièvre et al., 2008) and achieve uniformity between the procedures used for the different sensory methods. No quantifiers such as “very”, “medium” or “high” were provided or allowed.

3.2.3.5 Napping

The specific restricted version of projective mapping (Risvik et al., 1994, 1997; Pagès, 2003, 2005) called Napping was carried out using 60 x 40 cm white paper sheets in the “landscape” orientation. Sensory judges received all of the samples simultaneously.

Judges had to place similar samples close to each other and different samples far apart marking the desired positions of the samples with the sample’s three-digit code and an X on the white paper sheets. Sticky “Post-it” paper notes were provided to judges to make notes on and stick to the wine glasses to reduce the difficulty of the task of remembering the sensory characteristics of each sample during the positioning process.

Judges could move around the samples as many times as they wanted and take as much time as they deemed necessary to complete the task. Once a judge decided on the final configuration of the samples, three to five words from the provided list had to be provided. This verbalisation step was named ultra flash profiling (UFP) by Perrin et al. (2008).

3.2.4 Wine evaluation

Wine samples were presented in black tasting glasses (ISO NORM 3591, 1977) and covered with Petri-dishes as lids. Samples were labelled with random 3-digit codes. The serving order of the samples was randomised across sensory judges according to a Williams Latin-square design (Macfie et al., 1989). Therefore, each judge received the samples in a different order. A well-ventilated, temperature controlled, 20 ± 2^{o}C, odour free sensory lab secluded from extraneous noise equipped with separate off-white individual tasting booths and controlled lighting conditions were used for the evaluation of the wines. Each glass contained 25 mL of wine and was covered with a Petri-dish as lid. Wines were poured between 20 and 30 minutes before the sensory evaluation session in order to allow volatile compounds to reach equilibrium in the headspace of the glass.

Wines were evaluated orthonasally. All the wines were evaluated in triplicate for all the methods. Triplicates were evaluated on the same day with a 10-minute break in between to limit sensory fatigue. Panelists did not receive information about the style, vintage or cultivar of the samples and did not know that they evaluated the same wines twice.

3.2.5 Data analysis

3.2.5.1 Comparison of multivariate configurations from different methods

Multivariate statistical techniques were applied to the data obtained from the different sensory methods to create sensory maps illustrating the perceived sensorial similarities and difference of the samples relative to each other. Different statistical techniques were used, as proposed in literature, to accommodate the various data types.

The similarity between these sensory map configurations, corresponding to data from different sensory methods, were determined by calculating RV coefficients. An RV coefficient is a measure of the amount of variance shared between two matrices (Robert & Escoffier, 1976; Abdi et al., 2013; El Ghazir & Qannari, 2015). A schematic representation of the data analysis process can be seen in Fig. 1. In addition, hierarchical cluster analysis (HCA) was performed to assist with interpretation of the multivariate sensory maps. Ward's aggregation criteria and Euclidean distances were used.

Descriptive analysis. The performance of the DA panels was monitored according to the workflow suggested by Tomic et al. (2010) using PanelCheck V1.4.2 (www.panelcheck.com, Nofima) in order to determine when the panels were sufficiently trained and ready for data capturing. Once panel consensus and repeatability were confirmed by means of Tucker-1 and p*MSE plots further statistical analysis was conducted.

A 2-way mixed model ANOVA with judges, products and the judge*product interaction as factors were used to determine which of the sensory attributes were perceived significantly different for the different products. The judge and judge*product effects were assumed to be random. The product effect was tested using the regular F-test. Only attributes for which significant differences ($p \leq 0.05$) with regards to the product effect was found were kept for multivariate statistical analysis.

During the next step a multivariate statistical technique, standardised principal component analysis (PCA), was performed on the correlations matrix of the mean intensity ratings of the significant attributes. Confidence ellipses were added to the PCA graphs and were calculated by means of bootstrapping (Cadoret & Husson, 2013; Dehlholm et al., 2012b).

Check-all-that-apply (CATA). The number of attributes cited ("checked" on the CATA list) by the panel to describe the samples were reduced prior to statistical analysis using a protocol similar to the one describe by Campo et al. (2010). Attributes cited by less than 20% of the panel were combined with similar attributes. In cases where synonyms did not occur on the list, attributes were not used for further statistical analyses. Three sensory analysts combined similar attributes by means of semantic categorisation independently. Attributes combined differently by the sensory experts were discussed and consensus was reached on the matter prior to the final attribute reduction step.

The number of sensory judges that cited a specific attribute for a specific wine was counted. This procedure was followed for all the attributes and all the wines. A contingency table containing the sums of the citations over all the judges for each attribute for each wine was compiled. The citation frequency of an attribute for a wine was tabulated at the intersection of the corresponding row of that wine and column of that attribute. Correspondence analyses (CA) with confidence ellipses, calculated by means of bootstrapping, were performed on the contingency tables of the different data sets using.

Rate-all-that-apply (RATA). PCA was conducted on the correlations matrix of the mean intensity ratings. In addition, bootstrapping was used to construct confidence ellipses added to the PCA score plots (Cadoret & Husson, 2013; Dehlholm et al., 2012b).

Free sorting. The grouping of the samples by the different sensory judges during the sorting task was captured in individual distance matrices on which DISTATIS (Abdi et al., 2007) was performed.

Projective mapping. The *X* and *Y* coordinates, for each wine as placed on the A2 paper sheet by each judge, were tabulated using the left bottom corner as the origin. The *X* and *Y* coordinates were grouped for each judge. These individual data tables were analysed by means of Multiple factor analysis (MFA, Escofier & Pagès, 1990; Abdi & Valentin, 2013, 2014).

3.2.5.2 Comparison of the attributes used

Semantic data referring to sensory descriptors or attributes were provided as part of the sensory analysis when all the rapid methods were conducted. The descriptor list used during CATA was used for this purpose for all the rapid methods. During the RATA procedure, this data was captured using CATA prior to the intensity rating of the attributes. When projective mapping was performed judges had to write 3–5 descriptors from the CATA list onto the A2 sheet next to each sample. After the judges sorted samples into groups, "labelling" or "verbalisation" by citing sensory attributes from the list to explain the choice of samples grouped together was performed. In effect, a CATA step was embedded in each one of the rapid methods performed in this study an additional step where descriptors were provided.

These attributes obtained for the different methods were tabulated in contingency tables and analysed by means of MFA, performing CA on the separate data tables originating from the different rapid methods. The contingency tables were constructed using the same criteria as during the analysis of the CATA data. A schematic representation of the comparison of the attribute data obtained from the different sensory methods can be seen in Fig. 1.

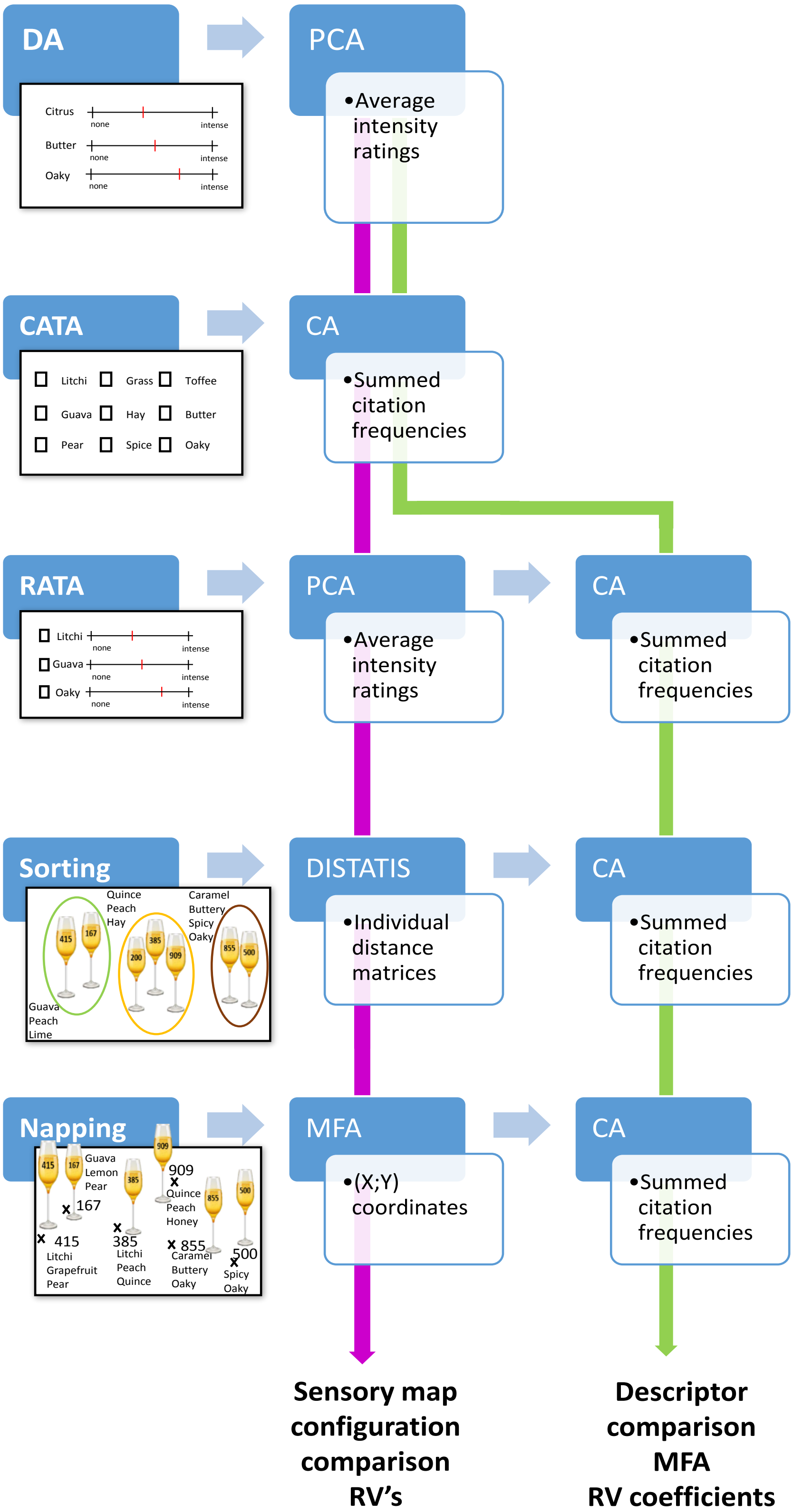


Fig. 1. Schematic overview of the comparison of the rapid methods: check-all-that-apply (CATA); rate-all-that-apply (RATA); sorting and Napping to each other and descriptive analysis (DA). The sample configurations obtained by principal component analysis (PCA), correspondence analysis (CA), DISTATIS and multiple factor analysis (MFA) were compared by means of RV coefficients. The descriptor data were compared by performing MFA with the descriptors obtained from the different methods as separate data tables.

3.2.5.3 Difficulty of the sensory task

After evaluating the wine samples by means of a specific sensory method, judges had to rate the easiness/difficulty of the task on a 9-point scale. The scale was derived from the 9-point hedonic liking scale (Peryam & Pilgrim, 1957) using the specific words: extremely easy, very easy, moderately easy, slightly easy, neither easy nor difficult, slightly difficult, moderately difficult, very difficult and extremely difficult. A value of one was tabulated when extremely easy was chosen and nine when extremely difficult was chosen. ANOVA was used to investigate significant differences between the difficulty of the sensory task as perceived by the judges for DA, CATA, RATA, sorting and projective mapping.

A three-way mixed model ANOVA, with method, sample set and the method*sample set interaction as fixed factors and judge as well as the judge interactions as random factors. The Fisher's LSD post hoc test was used to compute pairwise comparison when a significant ANOVA f-test result was found with $\alpha = 0.05$.

3.2.5.4 Data management and analyses

All data management and statistical analyses were conducted using Microsoft Excel (www.microsoft.com, Microsoft Corporation), XLSTAT 2017 (www.xlstat.com, Addinsoft), Statistica 13 (www.statsoft.com, Statsoft Inc.) and R version 3.4.0, packages "car", "cabootcrs" and DistatisR (www.R-project.org).

3.3 Results and discussion

3.3.1 Multivariate sensory map configuration comparison

The similarity between the sensory maps obtained from the different methods was assessed by calculating pairwise RV coefficients using the first two dimensions. The RV coefficients indicated that the multivariate sensory maps obtained from the different methods were similar with values ranging from 0.68 to 0.83 (Table 3), where an RV coefficient of 0.7 is regarded as good similarity between sensory maps. The CA plot constructed from the CATA data (Fig. 2E) was least similar to the PCA plot constructed from the DA data (Fig. 2A) with an RV coefficient of 0.69. The DISTATIS compromise map (Fig. 2G) and the MFA plot (Fig. I) obtained from the Napping data were most similar with an RV coefficient of 0.83 (Table 3). The PCA plot obtained from the RATA data (Fig. 2C) was more similar to the DA PCA plot (Fig. 2A), with an RV coefficient of 0.82, than to the CATA CA plot (Fig. 2E), with an RV coefficient of 0.68 (Table 3).

When evaluating the configurations by inspection it is clear that all the graphs show a similar pattern. Two distinct groups of samples can be identified. Samples I and C were clearly separated from the other samples along dimension1 or PC1. This is confirmed by the HCA performed on the multivariate maps (Fig. 2B, D, F, H and J).

TABLE 3
Rv coefficients used to compare different rapid methods and DA.

Sensory method	Rv coefficient pairwise comparisons				
Multivariate sensory map configurations					
	DA[a]	CATA[b]	Napping[c]	RATA[a]	Sorting[d]
DA[a]	1	0.69	0.82	0.82	0.82
CATA[b]	0.69	1	0.80	0.68	0.80
PM[c]	0.82	0.80	1	0.79	0.83
RATA[a]	0.82	0.68	0.79	1	0.78
Sorting[d]	0.82	0.80	0.83	0.78	1
Descriptors used					
	DA[a]	CATA[b]	Napping[b]	RATA[b]	Sorting[b]
DA[a]	1	0.86	0.95	0.93	0.92
CATA[b]	0.86	1	0.85	0.87	0.83
PM[b]	0.95	0.85	1	0.94	0.90
RATA[b]	0.93	0.87	0.94	1	0.88
Sorting[b]	0.92	0.83	0.90	0.88	1

[a]PCA was conducted on the correlations matrix,
[b]CA was conducted on the sum of the citation frequencies for all the attributes over all the judges' for all samples,
[c]MFA was performed on the individual judges' data,
[d]DISTATIS was performed on the similarity matrices of the individual sensory judges.

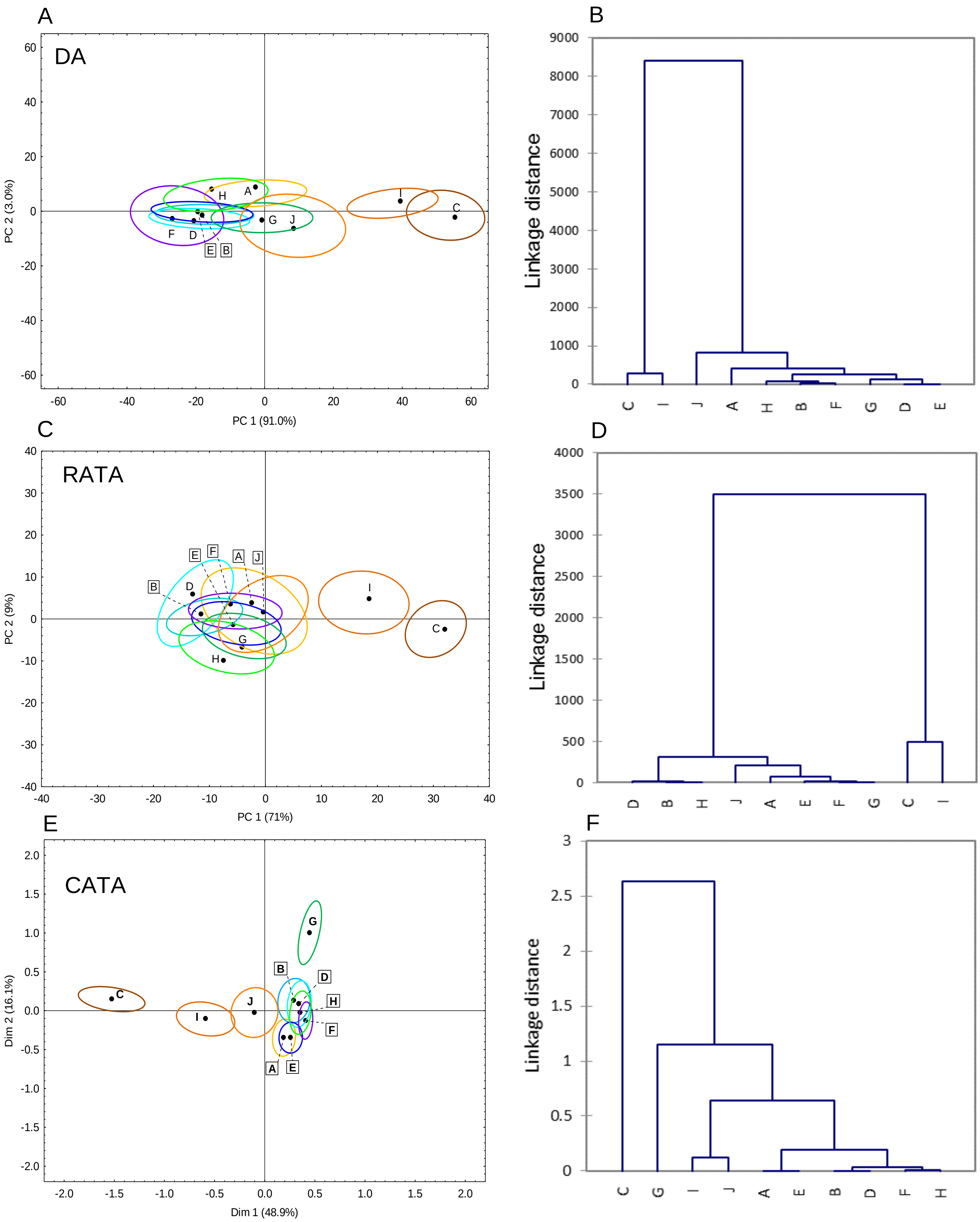


Fig. 2. Multivariate statistical analysis performed on the data obtained from the different sensory methods used: (A) principal component analysis (PCA) and (B) hierarchical cluster analysis (HCA) performed on descriptive analysis (DA) data; (C) PCA and (D) HCA performed on rate-all-that-apply (RATA) data; (E) correspondence analysis (CA) and (F) HCA performed on check-all-that-apply (CATA) data; (G) DISTATIS and (H) HCA performed on sorting data and (I) multiple factor analysis (MFA) and (J) HCA performed on Napping data.

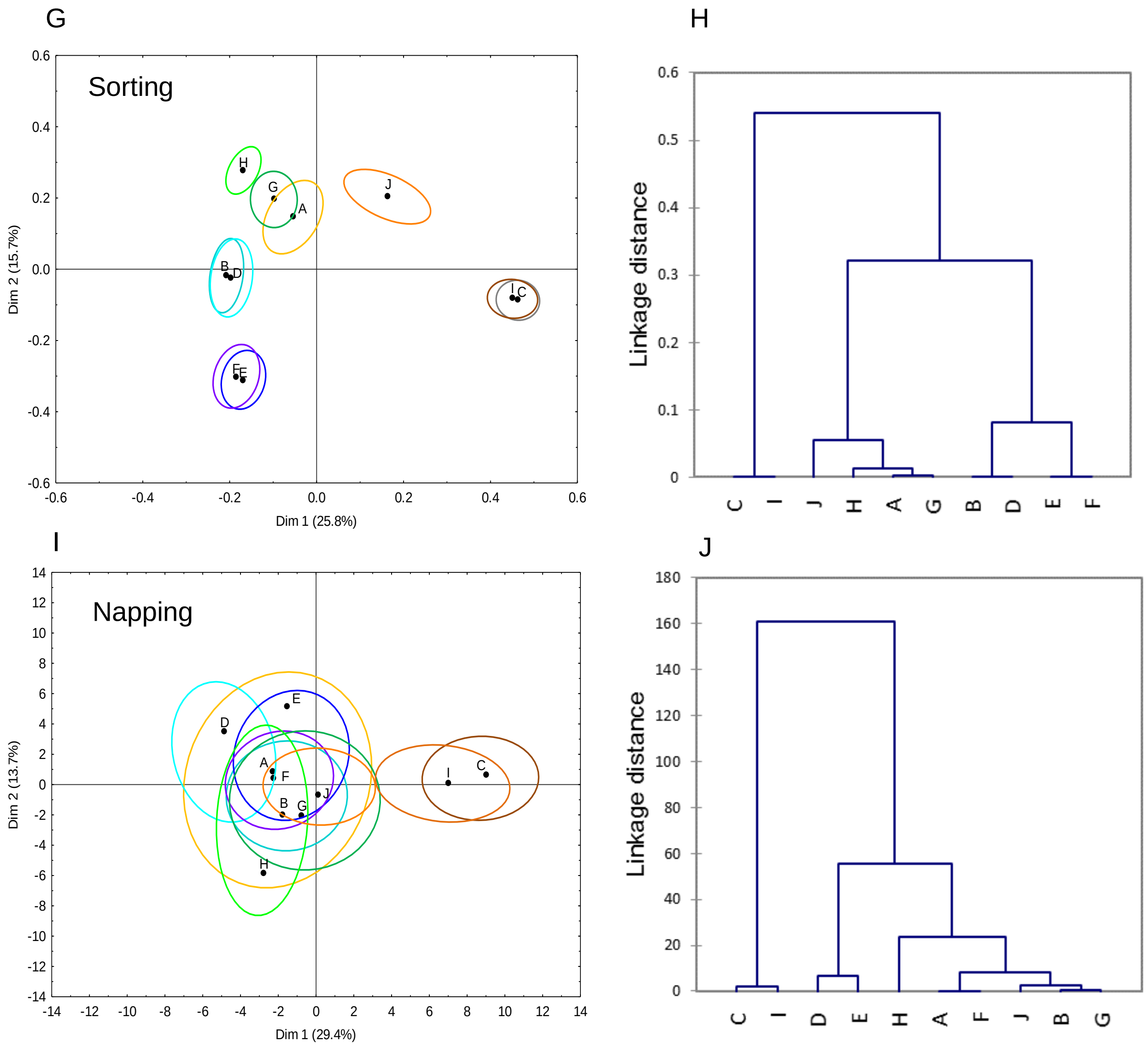


Fig. 2 cont. Multivariate statistical analysis performed on the data obtained from the different sensory methods used: (A) principal component analysis (PCA) and (B) hierarchical cluster analysis (HCA) performed on descriptive analysis (DA) data; (C) PCA and (D) HCA performed on rate-all-that-apply (RATA) data; (E) correspondence analysis (CA) and (F) HCA performed on check-all-that-apply (CATA) data; (G) DISTATIS and (H) HCA performed on sorting data and (I) multiple factor analysis (MFA) and (J) HCA performed on Napping data.

The MFA conducted on the Napping data (Fig. 2I) did not show clear differences between any of the other samples due to large overlapping confidence ellipses even though three separate groups were identified by the HCA (Fig. 2J) suggesting that samples D and E also formed a separate group. The overlap of confidence ellipses covered a larger area than on the DA PCA. A similar observation was made by Dehlholm et al (2012a) where the confidence ellipses for global Napping was larger and overlapped more frequently than for DA when liver pâtés were evaluated. It should also be said that the general variability measured by the explained variance is lower than for DA looking at the first two dimensions. However, it should be noted that the PCA was conducted on the average intensity scores over the entire panel, where the MFA was constructed from the individual data, which could have contributed to the lower explained variance of the first two factors of the MFA plot.

Both the PCA conducted on the RATA data (Fig. 2C) and the CA conducted on the CATA data (Fig. 2E) showed better separation between sample I and C with confidence ellipses that do not overlap, than the PCA plot obtained from the DA data (Fig. 2A) where the confidence ellipses overlap. In addition, the confidence ellipses around samples J and G do not overlap with those of the other samples on the CATA CA plot (Fig. 2E) indicating that this method could discriminate better between samples than DA, highlighting the differences between samples effectively.

Small, well-separated confidence ellipses could be seen on the DISTATIS compromise map constructed from the sorting data (Fig. 2G). However sample I and C overlap almost entirely indicating that these samples were perceived similarly and grouped together by many judges. The same can be said for sample E and F overlapping and B and D overlapping. Sample J appears separate as well as sample H. Sample G and A overlap but only partially. Comparing the DISTATIS map (Fig. 2G) to the HCA (Fig. 2H) similar groups of samples can be identified. Sorting seems to be able to discriminate better between the samples than DA in this case. Furthermore, comparing the sorting DISTATIS plot (Fig. 2G) to the CATA CA plot (Fig. 2E), sample similarity was highlighted rather than sample differences since the confidence ellipses around sample I and C overlap on the DISTATIS plot and not on the CA plot. It is interesting to note that the DA is results are almost one dimensional since over 90% of the variance is explained by PC1, where the DISTATIS graph is more balanced. Again it is important to note that DISTATIS was conducted on the individual data and PCA on the averages over all the judges.

For the data set analysed in this study, it can be said that the multivariate map configurations obtained with all the rapid methods were similar to DA. This observation is in-line with what was found in previous studies where one or two rapid methods were compared to DA at a time (Cartier et al., 2006; Perrin et al., 2008; Dehlholm et al., 2012a; Mielby et al., 2014).

To get a broad overview of the sensory space covered by the products and how they compare to each other, any one of the methods could be used. However, the best discrimination between samples was obtained by CATA and sorting taking the overlap of the confidence ellipses into account. Furthermore, the sorting DISTATIS map highlighted similarities between samples where the CATA CA plot highlighted slightly different product differences. This is expected since the sorting task entails the grouping of similar products while directly comparing them to each other, but when CATA is performed similarities between products are a result of common attributes used only, since products are never directly compared to each other.

3.3.2 Comparison of descriptors used

From the MFA partial projections plot of the descriptor data it can be seen that sample H, B and D showed the lowest variability since the data points corresponding to the different methods were closely grouped around the data point representing the sample. It is also clear that sample C, I and J showed the largest variability since the data points corresponding to the different methods were widely spread and far from the data point representing the sample (Fig. 3). It was shown that the sensory space of South African Chenin Blanc forms a continuum rather than distinct style categories in previous studies (Bester, 2011; Van Antwerpen, 2012; Hanekom, 2012) with the exception of wooded Chenin Blanc that is perceived as a separate category by trained panellists and industry professionals. Hence the large variability between the wooded samples, C and I, that was perceived as “oaky” with “vanilla” and “caramel” notes, and the rest of the products in the sample set. The low variability between sample H, B and D, which was described as “mineral” and fruity with aroma notes including: “yellow apple”, “citrus” and “pineapple”, could be attributed to the fact that these samples were not perceived differently (Fig. 4). It is also interesting to note that Ballester et al. (2013) and Parr et al. (2015) reported that the term “mineral” was not well understood by industry professionals.

In addition, the CATA and DA results for sample G and E are contradictory (Fig. 3). This could be attributed to the fact that sample G was described as herbaceous during CATA by many industry professionals. The assumption is made for CATA data that a high citation frequency indicates a high intensity, this is not necessarily true. When DA is conducted the intensity is rated and the average intensity is used during construction of the sensory map. It could also be possible that the term “herbaceous” was not understood in the same way by trained panellists and industry professionals and were not used similarly when DA was conducted by the trained panel and CATA by the industry professionals.

The confidence ellipses on the CA graphs constructed from the Napping (Fig. 4D) and RATA (Fig. 4A) data overlapped more frequently than those of the CATA (Fig. 4C) and sorting data (Fig. 4C) indicating that it is harder to distinguish between “noise” and what is “data”. This is not surprising and was seen when the different configurational plots were compared. However, it should be kept in mind that capturing data as RATA data and analysing it as CATA data was not recommended by the authors of any of the studies since discrimination ability between the samples are lost when RATA data is analysed as CATA data (Vidal et al., 2017; Oppermann 2017).

It is interesting to note that slightly better separation between products is obtained with the CA plot computed from the descriptors than the MFA plot computed from the coordinates of the Napping results. These two plots are almost identical, and it could be argued that the coordinate data do not provide extra information on the similarity and dissimilarity between samples. Furthermore, comparing the CATA CA (Fig. 4B) to the Napping (Fig. 4D) graphs, samples were separated better with fewer overlapping confidence ellipses on the CATA CA (Fig. 4B) indicating that the differences between the samples were described in more detail when CATA was performed than when Napping was performed. This could be due to the difficulty of the task since Napping was perceived a significantly more difficult than CATA (Table 4). It would be interesting to compare CATA and Napping evaluating sample sets with different within set variability and in addition investigate the contribution of the coordinates and descriptors to the discrimination between samples separately.

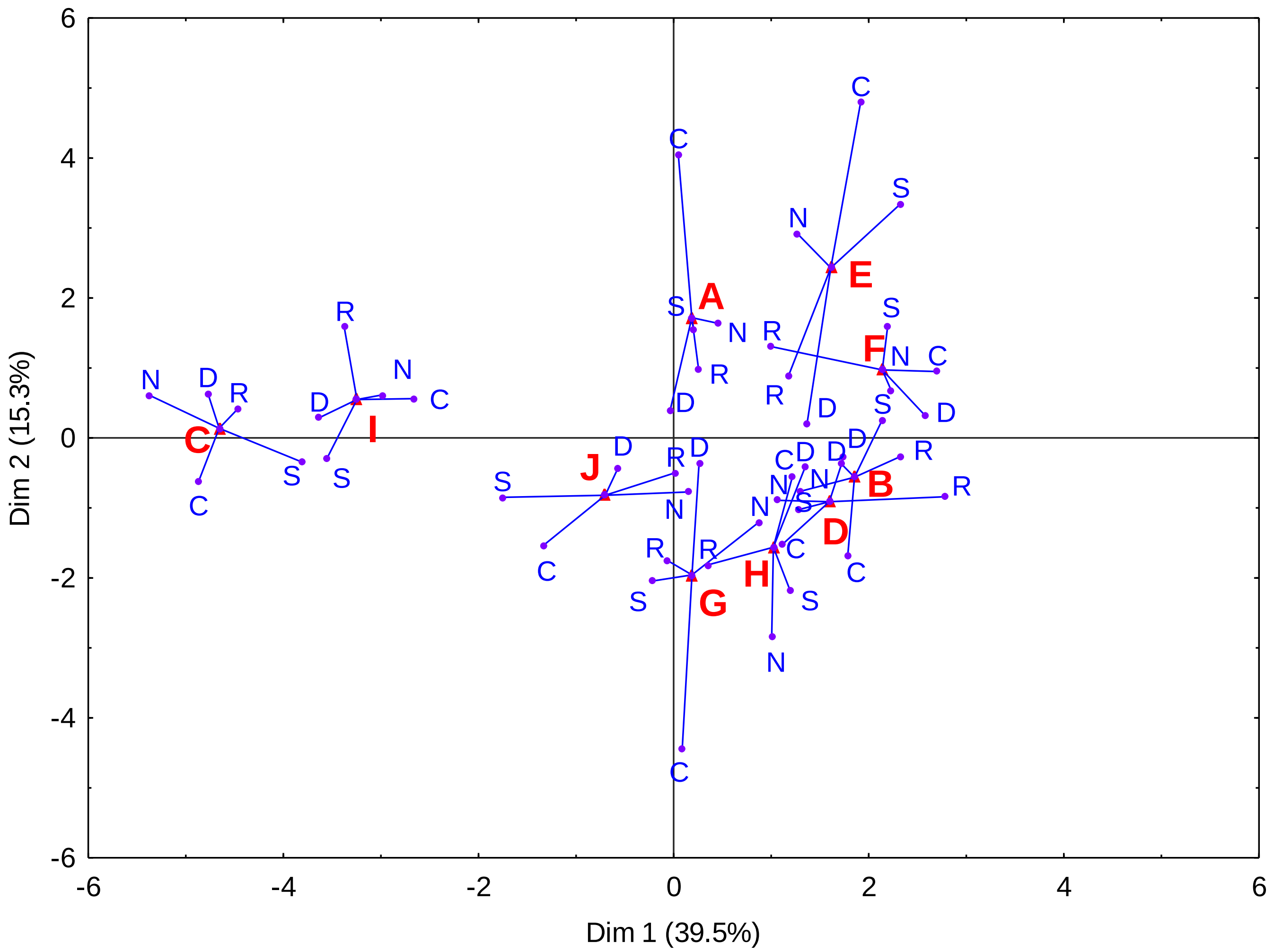


Fig. 3. Partial projections of the multiple factor analysis (MFA) conducted on the descriptor data obtained from the different methods where: D represents descriptive analysis (DA); C check-all-that-apply (CATA); N Napping or projective mapping; R rate-all-that-apply (RATA) and S sorting.

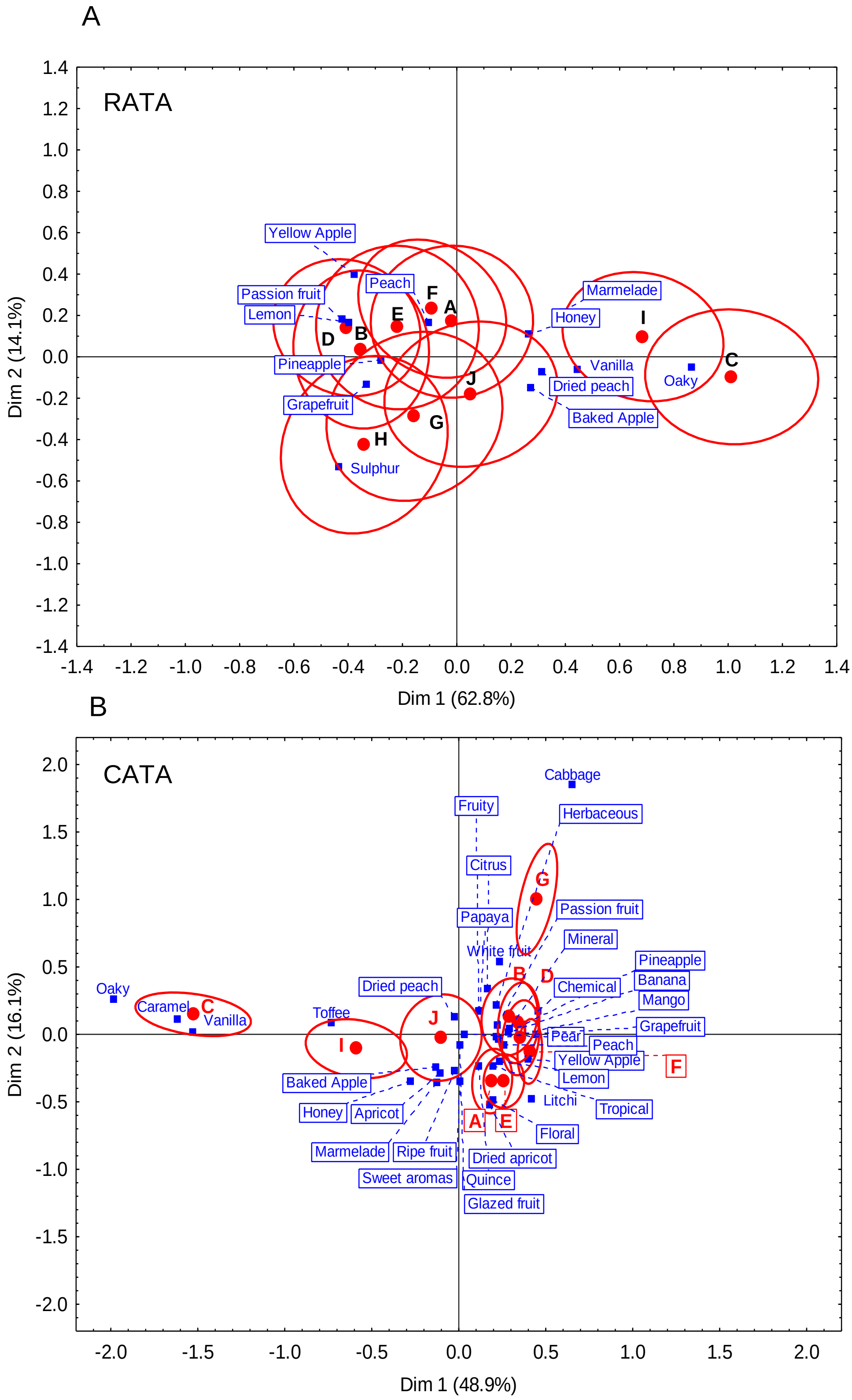
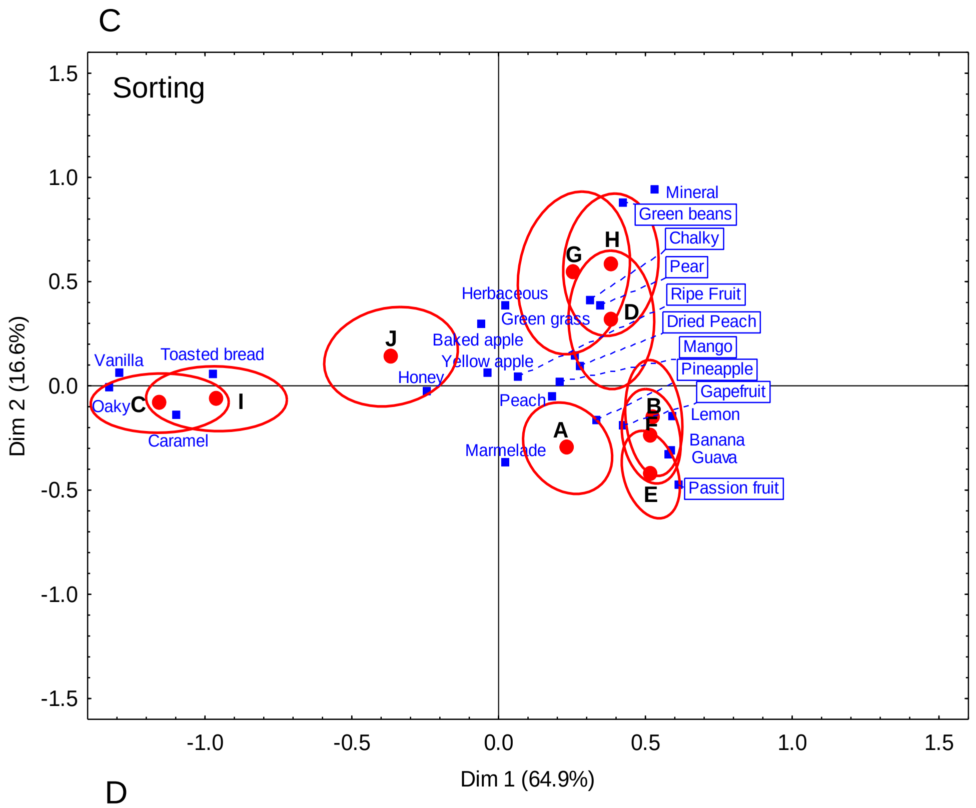


Fig. 4. CA conducted on the descriptor data of the respective rapid sensory methods (A) rate-all-that-apply (RATA), (B) check-all-that-apply (CATA), (C) sorting and (D) Napping descriptor data. (E) Principal component analysis (PCA) biplot of the DA results.



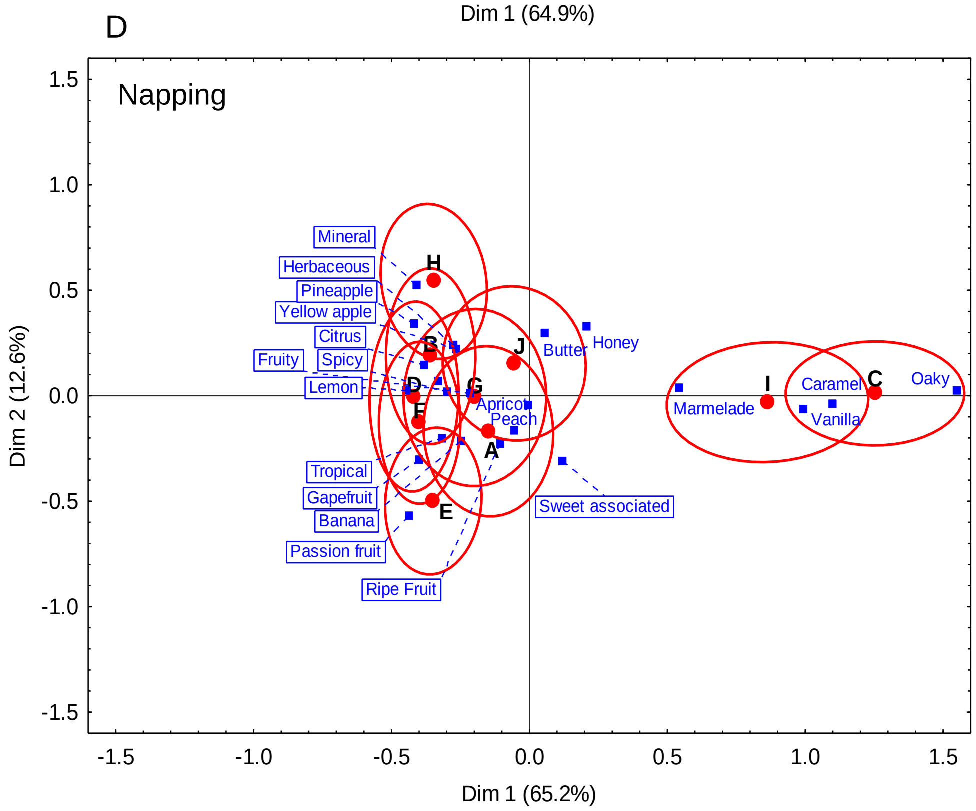


Fig. 4. CA conducted on the descriptor data of the respective rapid sensory methods (A) rate-all-that-apply (RATA), (B) check-all-that-apply (CATA), (C) sorting and (D) Napping descriptor data. (E) Principal component analysis (PCA) biplot of the DA results.

E

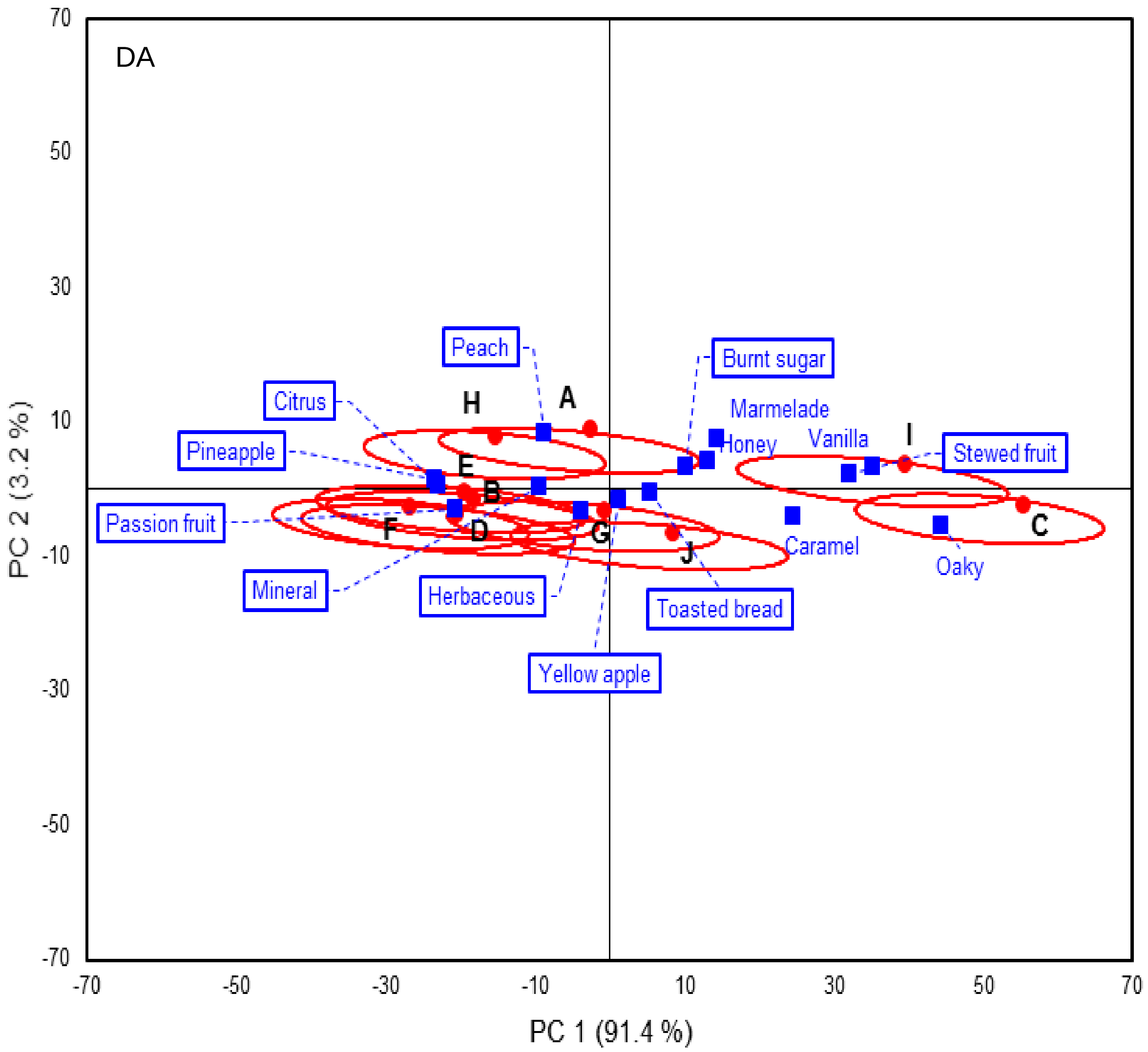


Fig. 4. CA conducted on the descriptor data of the respective rapid sensory methods (A) rate-all-that-apply (RATA), (B) check-all-that-apply (CATA), (C) sorting and (D) Napping descriptor data. (E) Principal component analysis (PCA) biplot of the DA results.

3.3.2.1 Describing and differentiating between different Chenin Blanc styles

Similar wine style descriptions were obtained with all the rapid methods. Taking both the configurations (Fig. 3)) and descriptors (Fig. 4) into account, it can be said that all the methods were able to separate between wooded and unwooded wines. Wooded wines were described as “oaky”, “vanilla” and “caramel”. A style known as “rich and ripe” could be distinguished from the “fresh and fruity” style only with sorting and CATA. In addition, the herbaceous sample, G was also separated from the other samples assessing the CATA CA (Fig. 3E, Fig. 4B). DA (Fig 4E), RATA (Fig. 4A) and Napping (Fig. 4D) provided a sensory map with the “rich and ripe” wines forming a continuum with the “fresh and fruity” wines. The “rich and ripe” wines were characterised by attributes including “honey”, “marmalade”, “baked apple”, “ripe fruit” and “quince”. It was reported by Bester (2011) that the Chenin Blanc sensory space is a continuum with the wooded wines forming a separate group by inspecting sorting and DA data. Bester (2011), however, did not apply bootstrapping to draw confidence ellipses when interpreting multivariate results. Revisiting those data sets applying confidence ellipses might provide additional insights into identifying South African Chenin Blanc styles on multivariate sensory maps.

3.3.3 Practical considerations

All the rapid sensory methods evaluated in this study can be performed in a single sensory evaluation session where DA can take up to 6 weeks to complete due to the extensive training period (Lawless & Heymann, 2010). Once the DA panel was trained the judges rated the intensities of the wine in less than 30 minutes. CATA and sorting were the fastest to perform with Napping that took the most time to complete (Table 4). Napping was, in addition experienced as the most difficult task for wine industry professionals to perform with sorting and CATA being significantly easier than the other rapid methods (Table 4). It is interesting to note that the easiest and fastest methods, sorting and CATA, provided sensory maps explaining the variation between the different samples the best with the least overlap of confidence ellipses.

TABLE 4.
Comparison of different sensory methods in terms of difficulty and the time required to complete the task.

Category	Difficulty of the task		Evaluation time		Total lab time
	Mean score out of 9	Significant letters	Mean evaluation time per replicate[d] in minutes	Significant letters	
Napping	7.33 ± 0.13	A	48.07 ± 1.02	A	1 session
DA[a]	5.86 ± 0.15	B	26.61 ± 1.14	D	11 sessions
RATA[b]	4.97 ± 0.15	C	35.67 ± 1.19	B	1 session
Sorting	4.02 ± 0.13	D	30.36 ± 1.02	C	1 session
CATA[c]	3.44 ± 0.13	E	27.36 ± 1.02	D	1 session

[a]Descriptive analysis (DA), [b]rate-all-that-apply, [c]check-all-that-apply.
[d]Three replicates of each sample set were evaluated for each method by each judge.
Alternatively, for rapid sensory analysis one replicate could be evaluated using 30 sensory judges.

3.4 Conclusions

The multivariate sensory maps obtained from DA data, conducted by a trained panel, and rapid methods, performed using industry professionals, were similar with RV coefficients higher or close to 0.7. All the sensory methods evaluated could discriminate between the "wooded" and "unwooded" style which was also shown by (Bester, 2011; Van Antwerpen 2012; Hanekom 2012). Taking the overlap of bootstrap confidence ellipses on the multivariate sensory maps into account, CATA and sorting were able to explain the difference between the samples better than DA being able to separate the "rich and ripe" from the "fresh and fruity" wines. CATA highlighted small differences between wines more effectively than sorting and provided richer descriptions with a wider vocabulary where sorting highlighted similarities and provided fewer descriptors. Due to the different pros and cons of these two methods the best strategy to use will depend on the purpose of the sensory evaluation session and the question to be answered. Sorting can for example be used as a quick profiling tool where the experimenter wants to gain information on the similarities and differences between the samples and focusses less on the individual sample's characteristics. CATA can be used if a more detailed profile for each sample is required.

CATA and sorting were rated as the easiest methods and took the industry professionals the shortest time to complete. These two methods are particularly suitable for sensory evaluation of wine as cost-effective alternatives for DA. However, the statistical analysis of the data obtained from rapid methods can be tedious to the experimenter and can be prone to bias since the attributes are condensed by the experimenter and co-workers and not the sensory judge. Future work in the field of rapid sensory analysis is required to optimise the condensing of the attributes.

3.5 References

Abdi, H., Valentin, D., Chrea, C., & Chollet, S. (2007). Analyzing assessors and products in sorting tasks: DISTATIS, theory and application. *Food Quality and Preference*, *18*, 627–640.

Abdi, H., & Valentin, D. (2014). Encyclopedia of Measurement and Statistics Multiple Factor Analysis. *Encyclodepia of Measurement and Statistics*, 658–664.

Abdi, H., Williams, L. J., & Valentin, D. (2013). Multiple factor analysis: Principal component analysis for multitable and multiblock data sets. *Wiley Interdisciplinary Reviews: Computational Statistics*, *5*, 149–179.

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. *Poster presented at the 7th Pangborn Sensory Science Symposium.* 16 August 2007, Minneapolis, USA. Delegate Manual, pp. 12–418.

Alencar, N. M. M., Ribeiro, T. G., Barone, B., Barros, A. P. A., Marques, A.T.B., & Behrens, H. J. (2018). Sensory profile and check-all-that-apply (cata) as tools for evaluating and characterizing syrah wines aged with oak chips. *Food Research International*, (March), 1–9.

Ares, G., Bruzzone, F., Vidal, L., Cadena, R. S., Giménez, A., Pineau, B., Hunter, D. C., Paisley, A.G., & Jaeger, S. R. (2014). Evaluation of a rating-based variant of check-all-that-apply questions: Rate-all-that-apply (RATA). *Food Quality and Preference*, *36*, 87–95.

Ares, G., Deliza, R., Barreiro, C., Giménez, A., & Gámbaro, A. (2010). Comparison of two sensory profiling techniques based on consumer perception. *Food Quality and Preference*, *21*(4), 417–426.

Ares, G., Antúnez, L., Bruzzone, F., Vidal, L., Giménez, A., Pineau, B., Beresford, M. K., Jin, D., Paisley, A. G., Chheang, S. L., Roigard, C. M., & Jaeger, S. R. (2015). Comparison of sensory product profiles generated by trained assessors and consumers using CATA questions : Four case studies with complex and / or similar samples, *Food Quality and Preference 45*, 75–86.

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

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

Bester, I., (2012). Classifying South African Chenin blanc wine styles. MSc thesis, Stellenbosch University. http://hdl.handle.net/10019.1/18048

Botha, A., (2015). The use of different oak products during the fermentation and ageing of Chenin Blanc : sensory properties, perceived quality, and consumer preference. MSc thesis, Stellenbosch University. http://hdl.handle.net/10019.1/98006

Cadena, R. S., Caimi, D., Jaunarena, I., Lorenzo, I., Vidal, L., Ares, G., Deliza, R., & Giménez, A. (2014). Comparison of rapid sensory characterization methodologies for the development of functional yogurts. *Food Research International*, *64*, 446–455.

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

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

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.

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.

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. (2000). Le degré d'expertise a-t-il une influence sur la perception olfactive? Quelques éléments de réponse dans le domaine du vin. L'Anne´e Psychologique, 100, 11–36.

Chollet, S., Lelie`vre, 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.

Coomb, C.H. (1964). *The Theory of Data*. New York: Wiley.

Corsi, A. M., Cohen, J., Lockshin, L., & Williamson, P. (2017). Testing lexical equivalences for wine flavours in emerging markets : Do hawthorns taste like blackberries ? *Food Quality and Preference*, *62*, 296–306.

Coste, A., Sousa, P., & Malfeito-Ferreira, M. (2018). Wine tasting based on emotional responses : An expedite approach to distinguish between warm and cool climate dry red wine styles. *Food Research International*, *106*, 11-21.

Dehlholm, C., Brockhoff, P. B., Meinert, L., Aaslyng, M. D., & Bredie, W. L. P. (2012a). 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. (2012b). 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.

Dooley, L., Lee, Y.S., & 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.

El Ghaziri, A., & Qannari, E. M. (2015). Measures of association between two datasets; Application to sensory data. *Food Quality and Preference*, *40*, 116–124.

Escofier, B., & Pagès, J. (1990). Multiple factor analysis. *Computational Statistics & Data Analysis, 18*, 121–140.

Franco-Luesma, E., Sáenz-navajas, M., Valentin, D., & Ballester, J. (2016). Study of the effect of H_2S , MeSH and DMS on the sensory profile of wine model solutions by Rate-All-That-Apply (RATA). *FRIN*, *87*, 152–160.

Hanekom, E., (2012). Chemical, sensory and consumer profiling of a selection of South African Chenin blanc wines produced from bush vines. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/71812

Heatherly, M., Dein, M., Munafo, J. P., & Luckett, C. R. (2019). Crossmodal correspondence between color , shapes , and wine odors. *Food Quality and Preference*, *71*, 395–405.

Honoré-Chedozeau, C., Lelièvre-desmas, M., Ballester, J., Chollet, S., & Valentin, D. (2017). Knowledge representation among assessors through free hierarchical sorting and a semi-directed interview : Exploring Beaujolais wines, *Food Quality and Preference, 57*, 17–31.

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.

Jaeger, S. R., Kim, K., Lee, S. M., Hunter, D. C., Kam, K., Chheang, S. L., Jin, D., Lee, P-Y., Xia, Y., & Ares, G. (2017). Concurrent elicitation of hedonic and CATA / RATA responses with Chinese and Korean consumers : Hedonic bias is unlikely to occur. *Food Quality and Preference*, *56*, 130–137.

Jaeger, S. R., Lee, P., & Ares, G. (2018). Product involvement and consumer food-elicited emotional associations : Insights from emoji questionnaires. *Food Research International*, *106*, 999–1011.

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.

Kostov, B., Bécue-Bertaut, M., & Husson, F. (2014). An original methodology for the analysis and interpretation of word-count based methods: Multiple factor analysis for contingency tables complemented by consensual words. *Food Quality and Preference*, *32*, 35–40.

Lawless, H. T., & Heymann, H. (2010). *Sensory evaluation of food. Principles and practices*. New York: Springer.

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.

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.

Lezaeta, A., Bordeu, E., Næs, T., & Varela, P. (2017). Exploration of consumer perception of Sauvignon Blanc wines with enhanced aroma properties using two different descriptive methods. *Food Research International*, *99*, 186–197.

Lezaeta, A., Bordeu, E., Agosin, E., Pérez-correa, J. R., & Varela, P. (2018). White wines aroma recovery and enrichment : Sensory-led aroma selection and consumer perception. *Food Research International*, *108*, 595–603.

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.

Liu et al., 2016 Liu, J., Schou, M., Di, R., Giacalone, D., & Bredie, W. L. P. (2016). Performance of Flash Profile and Napping with and without training for describing small sensory differences in a model wine. Food Quality and Preference, 48, 41–49.

Liu, J., Bredie, W. L. P., Sherman, E., Harbertson, J. F., & Heymann, H. (2018). Comparison of rapid descriptive sensory methodologies : Free-Choice Profiling , Flash Profile and modified Flash Profile. *Food Research International*, *106*, 892–900.

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., Oelofse, S., Næs, T., Lambrechts, M., Van Rensburg, P., & Nieuwoudt, H. (2015). Optimisation of the partial napping approach for the successful capturing of mouthfeel differentiation between brandy products. *Food Quality and Preference*, *41*, 245–253.

MacFie, H. J., Bratchell, N., Greenhoff, K., & Vallis, L. V. (1989). Designs to balance the effect of order of presentation and first-order carry-over effects in hall tests. *Journal of sensory studies*, *4*, 129-148

Mielby, L. H., Hopfer, H., Jensen, S., Thybo, A. K., & Heymann, H. (2014). Comparison of descriptive analysis, projective mapping and sorting performed on pictures of fruit and vegetable mixes. *Food Quality and Preference*, *35*, 86–94.

Oppermann, A. K. L., Graaf, C. De, Scholten, E., Stieger, M., & Piqueras-fiszman, B. (2017). Comparison of Rate-All-That-Apply (RATA) and Descriptive sensory Analysis (DA) of model double emulsions with subtle perceptual differences. *Food Quality and Preference*, *56*, 55–68.

Parr, W. V., Ballester, J., Peyron, D., Grose, C., & Valentin, D. (2015). Perceived minerality in Sauvignon wines: Influence of culture and perception mode. *Food Quality and Preference*, *41*, 121–132.

Pagès, J. (2003). Recueil direct de distances sensorielles: application à l'évaluation de dix vins blancs de Val de Loire. *Science des Aliments, 23*, 679–888.

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*(5), 637–658.

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.

Perrin, L., Symoneaux, R., Maître, I., Asselin, C., Jourjon, F., & Pagès, 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.

Peryam, D. R., & Pilgrim, F.J. (1957). Hedonic scale method of measuring food preference. *Food Technology*. *11*, 9-14

Picard, D., Dacremont, C., Valentin, D., & Giboreau, A. (2003). Perceptual dimensions of tactile textures. *Acta Psychologica*, *114*, 165–184.

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

Reinbach, H. C., Giacalone, D., Ribeiro, L. M., Bredie, W. L. P., & Frøst, M. B. (2014). Comparison of three sensory profiling methods based on consumer perception : CATA , CATA with intensity and Napping Ò. *Food Quality and Preference*, *32*, 160–166.

Risvik, E., McEwan, J.A,, Colwill, J.S., Colwill, & 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.

Soufflet, I., Calonnier, M., & Dacremont, C. (2004). A comparison between industrial experts' and novices' haptic perception organization: a tool to identify descriptors of handle of fabrics. Food Quality and Preference, 15, 689–699.

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., 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.

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*, 1563–1578.

Van Antwerpen, L., (2012). Chemical and sensory profiling of dry and semi-dry South African Chenin blanc wines. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/71853

Vidal, L., Antúnez, L., Giménez, A., Medina, K., Boido, E., & Ares, G. (2017). Sensory characterization of the astringency of commercial Uruguayan Tannat wines. *Food Research International*, *102*, 425–434.

Vidal, L., Antúnez, L., Rodríguez-Haralambides, A., Giménez, A., Medina, K., Boido, E., & Ares, G. (2018a). Relationship between astringency and phenolic composition of commercial Uruguayan Tannat wines : Application of boosted regression trees. *Food Research International*, *112*, 25–37.

Vidal, L., Ares, G., Hedderley, D. I., Meyners, M., & Jaeger, S. R. (2018b). Comparison of rate-all-that-apply (RATA) and check-all-that-apply (CATA) questions across seven consumer studies. *Food Quality and Preference*, *67*, 49–58.

Vidal, L., Cadena, R. S., Antúnez, L., Giménez, A., Varela, P., & Ares, G. (2014). Stability of sample configurations from projective mapping: How many consumers are necessary? *Food Quality and Preference*, *34*, 79–87.

Williams, A.A., & Langron, S.P. (1984). The use of free-choice profiling for the evaluation of commercial ports. *Journal of the Science of Food and Agriculture, 35*, 558–568.

Chapter 4

Research results

Validating Pivot Profile by means of comparison to Frequency of attribute Citation: Analysing complex products with trained assessors

This manuscript was submitted for publication in **Food Quality and Preference**

4. Research results

Validating Pivot Profile by means of comparison to Frequency of attribute Citation: Analysing complex products with trained assessors

J. Brand[a,b,*], D. Valentin[c], M. Kidd[d], M. A. Vivier[b], T. Næs[e], H.H. Nieuwoudt[b]

[a]*Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[b]*Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[c]*Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, CNRS, INRA, Univ. Bourgogne, Franche-Comté, F-2100, Dijon, France*

[d]*Centre for Statistical Consultation, Department of Statistics and Actuarial Sciences, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[e]*Nofima AS, Raw material and Process Optimization, Oslovegen 1, P.O. Box 210, N-1431, Ås, Norway*

Keywords: Pivot profile, frequency of attribute citation, pick-*K* attributes, check-all-that-apply, trained panel, correspondence analysis, confidence ellipses

Abstract

Pivot profile (PP), a rapid frequency-based method, is receiving progressively more attention due to its promising potential to profile complex matrices. When performing PP each sample is compared to a reference, the pivot, in an ordinal manner. Comparing results obtained from PP to descriptive techniques, that include panel training, has not been done. In addition, presentation of individual sensory judge's data on multivariate sensory maps was highlighted as an aspect that requires attention. This study aimed at validating and comparing PP, as profiling tool for complex wine matrices, against frequency of attribute citation (FC), by considering individual judges' data. Three sets, of six wines each, with varying within-set product similarity were analysed by a trained panel. The stability of the PP sensory space was tested by changing the pivot sample. The results were compared to the FC sensory space using RV coefficients. Bootstrapping, represented by confidence ellipses on the Correspondence Analysis (CA) plot, was applied to consider individual sensory judges' data. CA plots constructed from PP data, changing the pivot, were less similar to each other, with lower RV coefficients, than to CA plots constructed from FC data. The most profound differences between RV coefficients were observed for the sample set with extreme within-set variations. Higher explained variance was

obtained with PP than FC. However, confidence ellipses covered larger areas and overlapped more frequently indicating fewer significant differences between samples for PP than FC data. PP and FC data were comparable for the sample set with medium within-set variation.

4.1 Introduction

Describing the intrinsic properties of food products to obtain sensory profiles is a primary need within the food industry and plays an important role during product development, production, quality control, advertising and marketing. Due to increased pressure from the food and beverage industry to profile products faster and more cost–effectively, new sensory methods and optimised statistical tools are continually being developed. These include rapid sensory methods where no training of the panel is required and where the evaluation can be performed by experts, as well as trained or naïve consumer panels (Valentin et al., 2012; Varela & Ares, 2012).

One of the recent additions to the rapid methods category is pivot profile (PP), proposed by Thuillier et al. in 2015. When PP is performed, each sample is compared to a reference, the pivot sample. The attributes perceived as respectively, less or more intense, in the sample than in the pivot, are listed by the panel. PP, therefore, provides an estimation of the intensity of attributes in the samples relative to the reference, which is not the case with other rapid methods. A direct comparison between the samples and the pivot is made during the tasting of the samples. When other frequency-based methods, for example, check-all-that-apply (CATA, Adams et al., 2007), free choice profiling (Williams & Langron, 1984) and Ultra Flash profile (Valentin et al., 2012) are used, an estimate of attribute intensities is obtained through the assumption that attributes mentioned by more judges are more intense. PP could, therefore, be more suitable than other frequency-based methods for applications such as benchmarking and as a sensory profiling tool for homogeneous sensory spaces and complex matrices such as wine (Thuillier et al., 2015) and beer (Lelièvre-Desmas et al., 2017).

Thuillier et al. (2015) profiled Champagne using product experts as sensory judges when the method was first introduced. Lelièvre-Desmas et al. (2017) used PP to profile beers and showed that the choice of pivot did not have a large effect on the product positioning on the correspondence analysis (CA) plot for that specific data set. In the field of dairy research, Fonseca et al. (2016) compared PP to a frequency-based method, comment analysis (Symoneaux et al., 2012), and demonstrated that consumers could profile chocolate ice cream products efficiently using both methods. In a study by Esmerino et al. (2017), focussing again on consumer perception, PP was compared to CATA and projective mapping (PM, Risvik et al., 1994) when profiling Greek yoghurt samples. It was found that the three rapid methods provided similar results of sufficient quality to profile the evaluated products.

In a recent study, Deneulin et al. (2018) used PP to profile a large number of honey samples from all over the world. These studies showed that PP is a valuable asset to the rapid sensory method toolbox.

As with all new methods, however, further studies are needed to investigate and understand the strengths and limitations of a method such as PP. A number of specific aspects of PP that require further investigation were highlighted during previous studies and included the choice of the pivot and the within set similarity (Thuillier et al., 2015). Lelièvre-Desmas et al. (2017) evaluated the effect of these factors and reported that within set similarity impacted the results more than the choice of the pivot. However, in that study, the discrimination power of PP was not studied. Calculating confidence ellipses could provide a way to test significant differences between samples within data sets subjected to PP and better panel performance measures including repeatability and consensus are also needed (Lelièvre-Desmas et al., 2017). In the studies by Deneulin et al. (2018) and Fonseca et al. (2016) panel performance was not measured. Deneulin et al. (2018) concluded that the vocabulary used required more attention, calculating panel repeatability and consensus could shed light on these matters. Although Fonseca et al. (2016) used consumers as sensory judges and repeatability could not be measured, investigating segmentation could be interesting and contribute to understanding the sensitivity of PP as sensory method.

Thuillier et al. (2015) noted that descriptive analysis (DA) might be more suitable than PP for a detailed description of sample sets. However, to date, no study was conducted to test PP against traditional sensory methods for profiling complex products, such as beer and wine. DA has the limitation that, when assessing complex matrices, sensory judges could have difficulty differentiating different odours by using a line scale (Lawless, 1999). Training sensory judges on a list of attributes using reference standards and asking them to provide attribute names to describe products are easier.

Campo et al. (2008) called this strategy frequency of attribute citation (FC). FC is an adapted CATA procedure with specific adaptations and restrictions: (1) The list contains only sensory attributes, no phrases, emotional or hedonic terms are allowed; (2) The sensory attributes are organised into categories for example odour or aroma families; (3) Judges are trained to use the CATA list by means of reference standards; (4) Judges can re-organise the CATA list during training through consensus; and 5) Panel repeatability is measured to ensure quality data. FC was used to analyse wine (Campo et al. 2008) and compared to DA in a later study, obtaining similar results (Campo et al., 2010).

To compare continuous data, obtained from using a DA line scale, to ordinal data, obtained from PP, might add extra variation. To avoid that it would be better to rather compare two ordinal data sets by comparing PP with FC, rather than comparing PP with DA.

Although PP was proposed as a method that can be used with product experts as sensory judges (Thuillier et al., 2015), a trained panel was used in this study for both PP and FC to eliminate the panel effect when comparing the two methods. Another advantage of using a trained panel is to limit heterogeneity through training.

The aim of this study was to validate PP for the profiling of complex matrices, specifically wine, using FC, an established and trusted method, as reference. Three specific objectives were formulated. The first objective was to investigate the significance of the differences between wine samples in a set by applying bootstrapping to PP data to plot confidence ellipses on the CA plots. The second objective was to test the robustness of PP against FC, by changing both the pivot sample and the complexity of the sample set, defining complexity as within set variability. The third objective was to compare panel repeatability, consensus and perceived difficulty of the method for PP and FC.

4.2 Materials and methods

4.2.1 Samples

The wines used for this study were chosen based on the knowledge acquired on their sensory properties from previous studies (Bester, 2011; Hanekom, 2012; Van Antwerpen, 2012), the knowledge of expert tasters, wine industry professionals and sensory researchers and analysts. Three sets with different with-in set variation were analysed since it was noted by Lelièvre-Desmas et al. (2017) that the stability of the sensory space not only depends on the pivot sample chosen but also on the with-in set variation between samples. The following three sets with different within-set variation of samples were subjected to sensory analysis: (i) Six wooded Pinotage wines with similar characteristics; (ii) six wooded Chenin Blanc wines representing a sample set with medium within-set variation; and (ii) six Sauvignon Blanc wines with extreme style differences. The three sets of wines were profiled, using the same sensory methodology and workflow, resulting in three separate data sets.

Each set was analysed by means of FC and PP. Three different PP experiments were conducted for each set using different pivot samples, P1, P2 and P3. P1 was chosen to be a sample with extreme sensory characteristics for that particular sample set. P2 was chosen to be an average sample with no extreme characteristics and P3 was a blend of all the samples in the sample set using equal volumes.

All wines were commercially available, produced in South Africa and certified by the South African wine and spirits board (see Table 1).

TABLE 1
Summary of the vinification parameters and chemical analysis after bottling of the wines subjected to PP and FC sensory analyses.

Wine code	Wine	Vintage	Alcohol (% v/v)	RS[c] (g/L)	pH	TA[d] (g/L)	Vinification	Aging
Chenin Blanc wines								
ALC	Anura Limited Release Chenin Blanc	2012	14.5	3.7	3.49	6.4	Barrel fermented	Matured on lees for 12 months in French oak.
BHC	Graham Beck Bowed Head Chenin Blanc	2012	13.9	4.0	3.40	6.4	Barrel fermented	Matured for 9 months in French oak
CQC[b]	Welmoed cellar Credo Quattour	2010	14.8	4.8	3.50	3.5	Barrel fermented	Matured for 28 months in French oak
MSC	Mulderbosch Steen op Hout	2012	13.8	3.1	3.50	5.4	Barrel and tank fermented	Aged on lees for 6 months and barrel aged
HBC	Remhoogte Honeybunch Chenin Blanc	2013	14.0	2.4	3.36	5.8	Maceration on skins for 12 hours, fermented wild	Matured in French oak for 12 months
SBC[a]	Stellenbosch vineyards bush vine Chenin Blanc	2014	14.1	5.1	3.60	6.4	Barrel fermented with natural yeast.	Barrel aged for 12 months in French oak.
Pinotage wines								
AGP	Altydgedacht Pinotage	2014	14.4	2.9	3.57	5.7	MLF[e] in barrels	Matured for 12 months in 50% French oak and 50% American oak
BKP[a]	Beyerskloof Pinotage	2014	14.5	2.9	3.86	5.4	MLF[e]	Treated with oak
NHP	Neethlingshof Pinotage	2014	14.6	3.9	3.63	5.5		Matured for 9 months in 60% French oak and 40% American oak
LCP[b]	La Cave Pinotage	2014	14.5	3.5	3.50	5.6	MLF[e] in French oak.	Matured for 18 months in French oak
RCP	Riebeek cellars Pinotage	2014	14.3	4.3	3.66	5.5	French oak chip treatment during MLF[e].	
SBP	Stellenbosch Vineyards Bush vine Pinotage	2014	14.6	2.4	3.53	5.4		Matured for 16 months in French oak barrels.

[a]Pivot 1 indicated as P1; [b]Pivot 2 indicated as P2
[c]Residual sugar; [d]Titratable acidity expressed as tartaric acid equivalent
[e]Malolactic fermentation

TABLE 1 Cont.
Summary of the vinification parameters and chemical analysis after bottling of the wines subjected to PP and FC sensory analyses.

Wine code	Wine	Vintage	Alcohol (% v/v)	RS[c] (g/L)	pH	TA[d] (g/L)	Vinification	Aging
Sauvignon Blanc wines								
DGK	De Grendel Koetshuis	2014	13.0	1.7	3.28	6.3	Skin contact and cold settling was applied.	Extended lees contact.
GBP	Graham Beck Pheasants' run	2014	14.1	1.9	3.37	6.7	Skin contact for 16 hours, reductive conditions and cold fermentation.	Lees contact for 5 months prior to bottling.
GCS	Groot Constantia Sauvignon Blanc	2014	13.4	1.6	3.27	5.7	Cold fermentation.	Matured on the lees for 3 months.
HVS	Hidden Valley Sauvignon Blanc	2014	13.4	2.4	3.15	7.1	Cold fermentation and reductive methods.	Lees contact for 3 months.
JTO[b]	Jordan The Outlier	2014	13.0	2.1	3.40	5.1	Barrel fermented, 60% and tank fermented, 40%.	Aged for 8 months in barrel.
TSL[a]	Thelema Sutherland	2014	13.1	1.4	3.34	5.6	Fermented in stainless steel tanks.	

[a]Pivot 1 indicated as P1
[b]Pivot 2 indicated as P2
[c]Residual sugar
[d]Titratable acidity expressed as tartaric acid equivalent

4.2.2 Panel

The panel of sensory judges consisted of three males and 12 females all between the ages of 24 and 65 years (average age: 32). All judges were trained sensory assessors with more than two years of experience in wine sensory analysis performing DA, and were paid for their participation. The same panel participated in both the PP and FC experiments. Sensory evaluation sessions of a specific set of wines, for example, all the Chenin Blanc wines, were conducted at least two weeks apart, to ensure that the panel did not remember the wines, but less than a month apart, to ensure that wine ageing did not play a role. This protocol was followed for all three data sets represented by the three different cultivars.

4.2.3 Sensory methodology

4.2.3.1 Frequency of attribute citation (FC) and pivot profile (PP) methodology

Training. Panel training consisted of 15 sessions of one hour each over six weeks. Ballot training on 134 wine aroma attributes using reference standards (Table 2) was conducted according to the frequency of attribute citation training procedure (Campo et al., 2008 and Campo et al., 2010). The list of terms given to the panel of sensory judges was subdivided into aroma categories according to literature (Noble et al., 1987; Campo et al., 2010; Bester, 2011; Hanekom, 2012; Van Antwerpen, 2012). During each training session, judges were presented with 10 to 15 aroma standards to familiarise themselves with the terms on the list (ballot). Two to three wines were presented per session. Attributes used by the panel to describe the wines were discussed and the most frequently cited attributes were highlighted by the panel leader.

The training consisted of two phases, a general training phase, used to train the panel on the initial list of terms, and a specific training phase, where the panel was trained to profile wines similar to the wines presented during the final evaluation. During the specific training, judges could add terms to the initial list and change the categorisation of the terms in the separate aroma families in order to describe the sensory properties of the wines accurately. The final wine aroma attribute list with aroma standards is shown in Table 2 and consisted of 103 aroma attributes. Two specific training sessions, evaluating and discussing wines from the relevant cultivar and vintages, were performed per data set.

TABLE 2.
Aroma reference standards presented during training, representing the final aromas listed during FC and PP training and evaluation.

Aroma Family	Descriptor	Reference standard	Amount
Red berries	Raspberry	Raspberry sauce (Vahiné)	10 mL
	Redcurrant	Redcurrants (Hillcrest, frozen and thawed)	5 berries
	Strawberry	Strawberries (Hillcrest, frozen and thawed)	3 strawberries
Black berries	Blackberry	Blackberries (Hillcrest, frozen and thawed)	10 berries
	Blackcurrant	Blackcurrant syrup (Ribena)	20 mL
	Blueberry	Blueberries (Hillcrest, frozen and thawed)	15 berries
	Cherry	Cherry syrup (Védrenne)	10 mL
	Mulberry	Mulberries (Hillcrest, frozen and thawed)	6 berries
Tropical	Pineapple	Pineapple (fresh)	2 - 4 cm^2 pieces
	Passion fruit	Passion fruit (fresh)	4 pips and 1 cm^2 piece of skin
	Guava	Guava juice (Sir Fruit)	20 mL
	Litchi	Litchi (canned, Pot'O Gold)	1 litchi and 10 mL syrup
	Melon	Melon (fresh)	4 - 4 cm^2 pieces
	Mango	Mango (fresh)	3 - 4 cm^2 pieces
	Gooseberry	Gooseberry (frozen and thawed)	5 berries
	Banana	Banana (fresh)	3 disks
Stone fruit	Peach	Peach (fresh)	4 - 4 cm^2 pieces
	Apricot	Apricot juice (Ceres)	30 mL
White fruit	Pear	Pear (fresh)	4 - 4 cm^2 pieces
	Yellow apple	Yellow apple (fresh golden delicious)	2 wedges
Citrus	Lemon	Extract (Vahiné)	15 drops in 20 mL of water
	Grapefruit	Grapefruit peel	2 - 2 cm^2 pieces
	Orange	Orange peel	2 cm^2 piece

TABLE 2. Cont.
Aroma reference standards presented during training, representing the final aromas listed during FC and PP training and evaluation.

Aroma Family	Descriptor	Reference standard	Amount
Floral	Honeysuckle	Perfume essence (Ferminich)	5 drops on cotton wool
	Elderflower	Elderflower syrup (Bottlegreen)	20 mL
	Linden tree flower	Linden tea (Twinings)	1 tea bag prepared in 125 mL boiling water
	Violet	Violet syrup (Védrenne)	10 mL
	Rose	Rose water (Woolworths)	10 mL
	Geranium	Geranium petals (fresh and crushed)	2 petals
Dried fruit	Prune	Dried prune (Safari)	1 prune cut to pieces
	Raisin	Raisins (Safari)	8 raisins crushed
	Date	Dates (fresh, local supermarket)	1 date cut to pieces
	Stewed fruit	Cooked dried fruit (Safari)	1 dried apple, ½ prune, ½ dried peach, 1 dried pear
Sweet associated	Baked apple	Cooked fresh apple (golden delicious)	Cooked puree, 15 mL
	Quince	Quince jam (local farm stall)	5 mL
	Jammy	Mixed fruit jam (Rhodes)	5 mL
	Ripe fruit	Verbal decription	Intense sweet fruity aroma
	Marmalade	Seville orange marmalade (All Gold)	5 mL
	Honey	Acacia honey (Lune de Miel)	15 mL
	Glazed fruit	Glazed fruit (Moir’s)	1 cherry, ¼ orange, ¼ pineapple
	Muscat	Le Nez du Vin standard	5 drops on cotton wool
	Candy floss	Candy floss (local supermarket)	2 g
	Vanilla	Vanilla pods (Woolworths)	1 pod
	Caramel	Caramel syrup (St. Dalfour)	20 mL
	Toffee	Soft toffees (Toff-o-lux)	1 toffee in boiling water
	Chocolate	Chocolate sauce (Hersey’s)	30 mL

TABLE 2.Cont.

Aroma reference standards presented during training, representing the final aromas listed during FC and PP training and evaluation.

Aroma Family	Descriptor	Reference standard	Amount
Toasted	Coffee	Roasted coffee beans (LavAzza)	5 grinded coffee beans
	Toasted bread	Toasted bread (local supermarket)	2 g
	Smoky	Le Nez du Vin standard	5 drops on cotton wool
Wooded	Oaky	Medium toasted French oak chips	2 g (NT Bois, RX South Africa)
	Planky	Pine wood shavings (local carpenter)	2 g
	Pencil shavings	Pencil shavings (Staedtler HB pencil)	1 g
Mineral	Flinty	Flint stone	5 flintstones struck against each other
	Salty	Sea water	50 mL sea water (local beach)
Savoury	Meaty	Meat stock (Knorr)	1 cube dissolved in 50 mL boiling water
	Soy	Soy sauce (Vital)	20 mL
	Bacon	Cooked bacon (Enterprise)	5 - 1cm^2 pieces
Fresh green	Green grass	Green grass (fresh)	Cut grass pieces 3 g
	Green pepper	Green pepper (fresh)	1 piece 2 cm^2
	Celery	Celery (fresh)	2 pieces of 1 cm^2
	Minty	Mint leaves (fresh)	2 leaves crushed
	Bay leaf	Bay leaves (dried, Roberston spice)	3 leaves broken into pieces
	Tomato leaf	Tomato leaves (fresh)	2 leaves crushed
	Eucalyptus	Eucalyptus leaves (fresh)	3 leaves broken into pieces
Canned green	Asparagus	Canned white asparagus (Pot'O Gold)	1 spear and 10 mL brine
	Green beans	Canned green beans (Rhodes)	2 beans and 20 mL brine
	Canned peas	Canned peas (KOO)	5 peas and 20 mL brine
	Olive	Olives in brine (Darling)	2 pitted olives and 10 mL brine
	Cooked veg	Beans, asparagus, artichoke brine (KOO)	10 mL
	Gherkin	Pickled gherkin (KOO)	1 gherkin and 20 mL brine
Dried green	Tea	Black tea leaves (Glen)	5 mL
	Hay/Straw	Straw (local farmer)	1 g
	Dried grass	Pet shop grass (local petshop)	1 g
	Tobacco	Tobacco (Boxer)	5 mL
	Dried herbs	Mixed dried herbs (Robertson spice)	15 mL

TABLE 2.Cont.

Aroma reference standards presented during training, representing the final aromas listed during FC and PP training and evaluation.

Aroma Family	Descriptor	Reference standard	Amount
Spicy	Nutmeg	Nutmeg powder (Robertson spice)	5 mL
	Clove	Cloves (Robertson spice)	5 mL
	Cinnamon	Cinnamon sticks (Robertson spice)	1 stick
	Aniseed/liquorice	Liquorice (Mister sweet)	1 stick (2 cm x 1 cm)
	Black pepper	Black pepper (grinded, Robertson spice)	5 mL
Earthy	Dusty	Wet slate	Slate stone with water
	Earthy	Geosmine (Merck)	1 ng/L[a]
	Mouldy	Moulded white bread	3 - 1 cm^2 pieces
	Forest floor	Soil (collected from local nature reserve)	15 mL
	Mushroom	Brown mushrooms (fresh)	1 mushroom crushed
Animal	Cat pee	3MH (Merck)	6 000 ng/L[a]
	Barnyard	4-EP (Merck)	800 mg/L[a]
	Leather	New leather strip	1 piece (4 cm^2)
Chemical	Acetone	Acetone (Merck)	50% v/v[a]
	Alcohol	Alcohol (Merck)	10% v/v[a]
	Nail polish remover	Nail polish remover (Cutex)	10 mL diluted 1:5
	Vinegar	White wine vinegar (Safari)	10 mL dilutes 1:3
	Oxidised apple	Grated apple left to oxidise (Granny Smith)	15 mL
	Sherry	Old Brown Sherry (Sedgwick's)	30 mL
	Sulphur	Sulphur (SO_2 solution)	40 ppm[a]
Nutty	Coconut	Coconut milk (Mayfair)	40 mL
	Almond	Almond essence (Moir's)	5 drops in 30 mL water
	Walnut	Walnut butter (local farm stall)	15 mL
	Hazelnut	Hazelnut spread (local farm stall)	15 mL
	Pistacio	Pistachio spread (local farm stall)	15 mL
Lactic	Buttery	Butter (melted, unsalted, Clover)	15 mL
	Cheesy	Matured cheddar and blue cheese (Fairview)	1 cm^2 piece of blue cheese and two 1 cm^2 piece of cheddar
	Rancid	Nez du Vin standard	5 drops on cotton wool

Procedure. Judges had to provide three to five terms from the list to describe the aroma of each wine when FC was conducted. Campo et al. (2010) suggested to specify the required number of attributes each judge should use to describe products when the frequency of attribute citation method is used. The reason for this suggestion was to avoid the use of too few or too many descriptors. People have a limited capacity to discriminate between and describe odours in complex samples, but using too few descriptors can lead to difficulty to achieve accurate descriptions of samples (Laing & Glemarec, 1992). On the other hand, the opposite scenario could occur where large numbers of attributes, including many synonyms, are used to describe wines adding noise to the data, complicating and adding biases to the statistical analysis workflow of the data.

During PP judges were requested to write down the attributes that they perceived less intense and more intense in the sample than the pivot from the list of attributes (Fig. 1). The same list as provided for FC was used. Judges were limited in terms of the number of attributes that they could use during PP in order to achieve a degree of standardisation between the instructions for PP and FC. No more than five attributes per sample were allowed to describe the aromas that they perceived less intense in the sample than the pivot. The same rule applied to the attributes perceived more intense than the pivot. Finally, judges had to provide at least three attributes in total per sample.

Wine evaluation. A well-ventilated, temperature controlled, 20 ± 2°C, odour free sensory lab secluded from extraneous noise equipped with separate off-white individual tasting booths and controlled lighting conditions were used for the evaluation of the wines. Monadic samples presentation was applied for the FC method. For PP, samples were presented in pairs. Each sample was presented together with a fresh pivot.

Black (ISO NORM 3591, 1977) tasting glasses labelled with random 3-digit codes were used. Sample randomisation across judges, according to a Williams Latin-square design (MacFie et al., 1988). Each glass contained 25 mL of wine and was covered with a Petri-dish as a lid. Wines were poured between 20 and 30 minutes before the sensory evaluation session in order to allow volatile compounds to reach equilibrium in the headspace of the glass.

Wines were evaluated orthonasally in duplicate for both methods. Duplicates were evaluated on the same day with an enforced 10-minute break between the two duplicates to limit sensory fatigue. The panel did not receive information on the nature of the wines in terms of style, vintage or cultivar and did not know that they evaluated the same wines twice. Data were collected using Compusense cloud software (www.compusense.com, Compusense).

4.2.4 Data analysis

4.2.4.1 Panel performance

Repeatability. Panel repeatability was calculated for the individual judges by means of the reproducibility index (R_i) proposed by Campo et al. (2008). In addition, a global reproducibility index (R_i) was calculated by computing the average across all judges. This measure ranges from 0 to 1. If all the attributes cited during the first and second repeated evaluation session are the same, then R_i will be 1. If completely different attributes were cited then R_i will be 0. A minimum R_i of 0.2 was proposed by Campo et al. (2008) to deem a sensory judge as repeatable enough to record the response as data.

$$R_i = \frac{1}{n}\sum \frac{2 \times des_{com}}{(des_{rep1} + des_{rep2})}$$

Where: $n = number\ of\ wines$

$des_{com} = number\ of\ identical\ descriptors\ chosen\ by\ the\ judge\ in\ both\ replicates$

$des_{rep1} = number\ of\ descriptors\ chosen\ by\ the\ judge\ in\ replicate\ 1$

$des_{rep2} = number\ of\ descriptors\ chosen\ by\ the\ judge\ in\ replicate\ 2$

R_i values were calculated for both the FC and PP methods for all the data sets. Each PP set obtained from using a different pivot sample was treated as a separate data set.

A 3-way mixed model ANOVA with cultivar, method and the cultivar*methods interaction as fixed factors and sensory judges as random factors was computed . The ANOVA was used to study the differences between repeatability of the panel in terms of Ri values computed when (1) sample sets with different within-set variation was evaluated and (2) different sensory methods (PP and FC) and pivot samples were used. Sample sets from different cultivars represented sets with different within-set variability as explained before. Pinotage represented low, Chenin Blanc medium and Sauvignon Blanc large within-sample variability. The different methods used were FC and PP using different pivot samples, P1, P2 and P3. The REML estimation method was used. When significant ANOVA results were found pairwise comparisons were calculated using the Fisher's LSD post-hoc test with α set at 5%.

Consensus Panel consensus was measured calculating Cohen's Kappa coefficients for each pair of judges. Cohen's kappa coefficient is a measure of the similarity or agreement between the ratings provided by two individuals. It is commonly used on nominal data as an interrater reliability measure in the field of medical and educational surveying (Cohen, 1960; Altman 1991;

McHugh, 2012; Gisev et al., 2013). In this study, Cohen's kappa coefficients (K) were calculated using the mathematical equation below.

$$\kappa = \frac{p_o - p_e}{1 - p_e}$$

Where:

$$p_o = the\ relative\ observed\ agreement\ among\ raters\ (sensory\ judges\ in\ this\ case)$$

$$p_e = the\ hypothetical\ probability\ of\ chance\ agreement$$

In addition, the average panel consensus was calculated for each data set by computing the average of all the Cohen's kappa coefficients across all the judges. Each PP sample set obtained from using a different sample as pivot was treated as a separate data set. A 3-way mixed model ANOVA similar to the ANOVA computed on the Ri values was computed on the Cohen's kappa coefficients.

Difficulty of the sensory task. Sensory judges were asked to give a score out of 9 on an easiness scale derived from the 9-point hedonic liking scale (Peryam & Pilgrim, 1957). The specific words used were: extremely easy (1); very easy (2); moderately easy (3); slightly easy (4); neither easy nor difficult (5); slightly difficult (6); moderately difficult (7); very difficult (8); and extremely difficult (9). A 3-way mixed model ANOVA, similar to the ANOVA's applied to assess panel consensus and repeatability, was performed to investigate significant differences between the perceived difficulty of the different FC and PP data sets.

4.2.4.2 Product characterisation

The descriptors generated to describe each group of wines in the verbalisation phase were captured by constructing a contingency table. The number of attributes used was reduced prior to statistical analysis. Attributes cited by less than 20% of the panel were combined with similar terms or discarded. Three sensory experts combined similar terms independently by means of lemmatisation and semantic categorisation. Attributes combined differently by the sensory experts were discussed, and consensus was reached on the matter prior to the final attribute reduction step. A schematic representation of the data organisation and analysis can be seen in Fig. 1A.

A multivariate approach was used to visualise the sensory space spanned by the different wines within a data set. Correspondence analyses (CA) with confidence ellipses, calculated by means of bootstrapping (Cadoret & Husson, 2013; Dehlholm et al., 2012), were performed on the contingency tables.

Contingency tables were constructed from FC and PP data in different ways. For FC data the total number of citations over all the judges for each descriptor for each wine was tabulated with the attributes as variables in the columns and the wines as objects in the rows. The number of judges who cited an attribute for a specific wine was tabulated at the intersection of the corresponding column (representing the attribute) and row (representing the wine). This procedure is the same as for standard CATA.

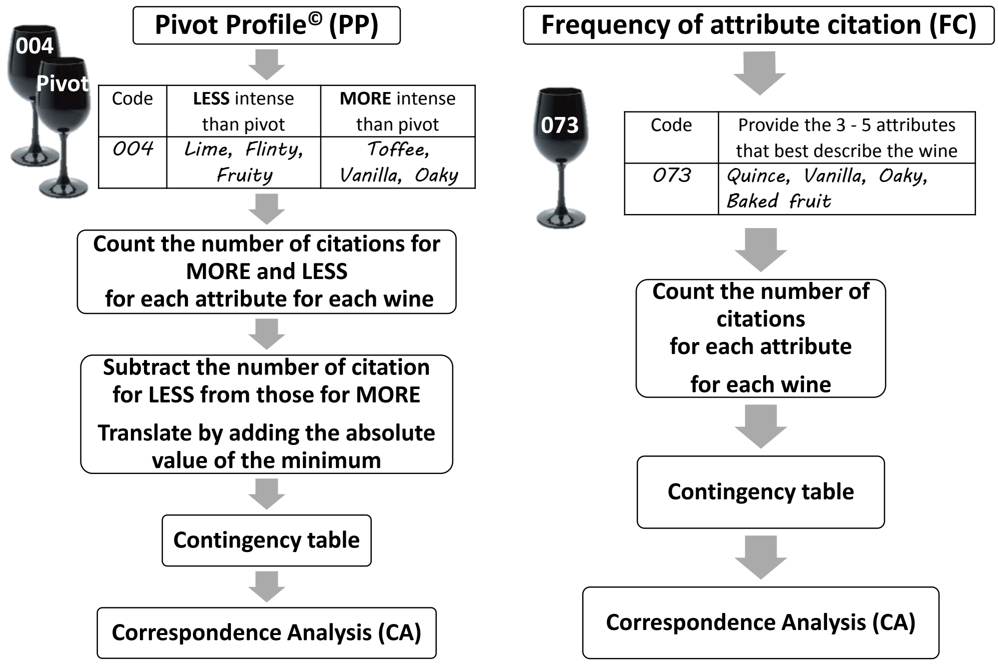
PP data sets were compiled by subtracting the citation frequency of "less" from "more" for each attribute for each wine. This procedure produced both positive and negative values. Since CA cannot be conducted on a table containing negative values, translation had to be performed to obtain a contingency table consisting of positive values. To obtain the contingency table the absolute value of the minimum was added to all the values as translation step. This procedure is described in detail by Thuillier et al. (2015) and summarised in Fig.1. In order to apply bootstrapping on the PP data, the contingency table was converted into an appropriate data set for CA by repeating each combination of wine and descriptor n_{ij} times where n_{ij} is the frequency of the *i*-th wine and the *j*-th descriptor in the contingency table.

4.2.4.3 Comparison of methods and testing the stability of the PP sensory space

The similarities between multivariate plots were assessed by calculating RV coefficients. RV coefficients are used to measure the similarity between two matrices or data sets by measuring the amount of variance shared (Robert & Escouffier, 1976; Abdi et al., 2013; Abdi et al., 2014). CA plots, generated from PP data sets where different samples were used as the pivot were compared to each other and to the CA plot constructed from FC data (Fig. 1B). This procedure was followed for all three sets, the set with the low (Pinotage), with medium (Chenin Blanc) and with large (Sauvignon Blanc) within-set variability, separately.

In addition, the repeatability, panel consensus and difficulty as perceived by the panellists performing the two different sensory methods were compared using ANOVA, as mentioned above. All data organisation and analyses were conducted using Microsoft Excel 2016 (www.microsoft.com, Microsoft), XLSTAT (www.XLSTAT.com, Addinsoft SARL.), Statistica 13 (www.statsoft.com, Statsoft Inc.) and R version 3.4.0, packages "car" and "cabootcrs" (www.R-project.org).

A SENSORY METHODOLOGY SCHEMATIC REPRESENTATION



B MULTIVARIATE MAP CONFIGURATION COMPARISON

SCHEMATIC REPRESENTATION

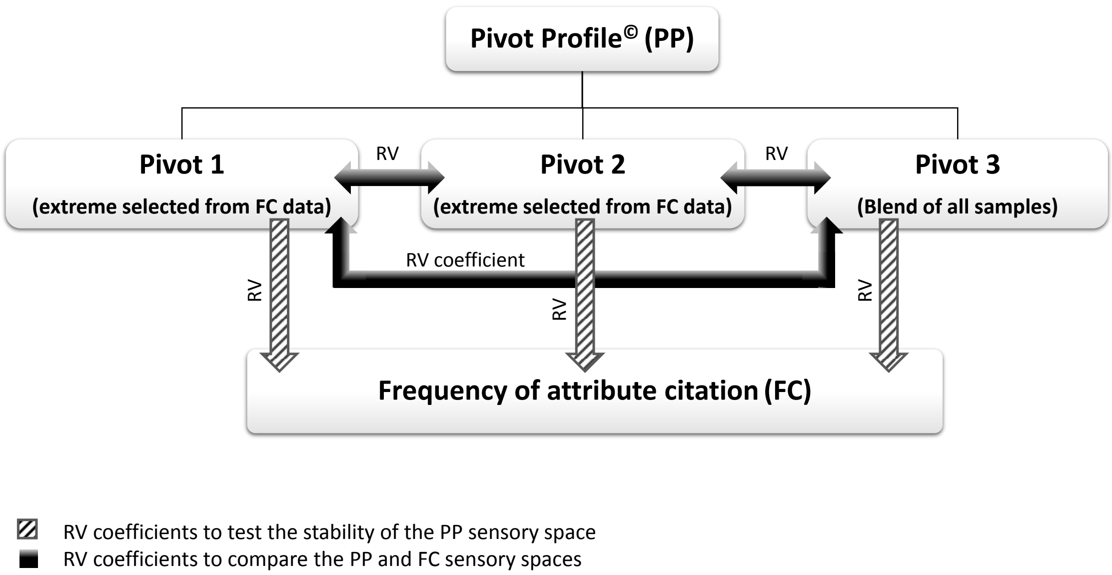


Fig. 1. Schematic representation highlighting the differences between pivot profile (PP) and frequency of attribute citation (FC) in terms of the (A) sensory methodology, data capturing, data analysis and (B) showing the strategy used to compare PP data obtained from changing the pivot sample to FC data.

4.3 Results

4.3.1 Panel performance

4.3.1.1 Repeatability

The individual repeatability indices for all the sensory judges were above 0.2 for both FC and PP, irrespective of which samples were used as the pivot. The highest R_i value was 0.86 and the lowest 0.26. All the judges produced repeatable results considering that the scale ranges from 0 to 1 and Campo et al. (2008) proposed 0.2 as the lowest acceptable value.

It is clear from the 2-way ANOVA results (Fig. 2A), with cultivar, representing data set complexity, and method as fixed factors that the repeatability of the sensory judges varied depending on the complexity of the data set analysed. Sensory judges were slightly less repeatable when conducting FC than PP for the data set with medium within-set variation (Chenin Blanc wines). A significant difference between FC and PP with P2 and P3 was seen. In addition judges were less repeatable when P1 was used than when P2 was used. No significant difference in repeatability was seen when P1 and P3 (the blend of all the samples) and P2 and P3 were used. Only a significant difference between using P2 and P1 as pivot sample could be seen for the data set with extreme within-set variability (Sauvignon Blanc wines). In addition, no significant differences between PP, when changing the pivot or between PP and FC was observed for the data sets with small (Pinotage wines) within-set variation.

However, the average panel repeatability was the lowest for the sample set with the lowest within sample variations (Pinotage wines) and differed significantly from that of the set with the extreme within-set variation (Sauvignon Blanc wines).

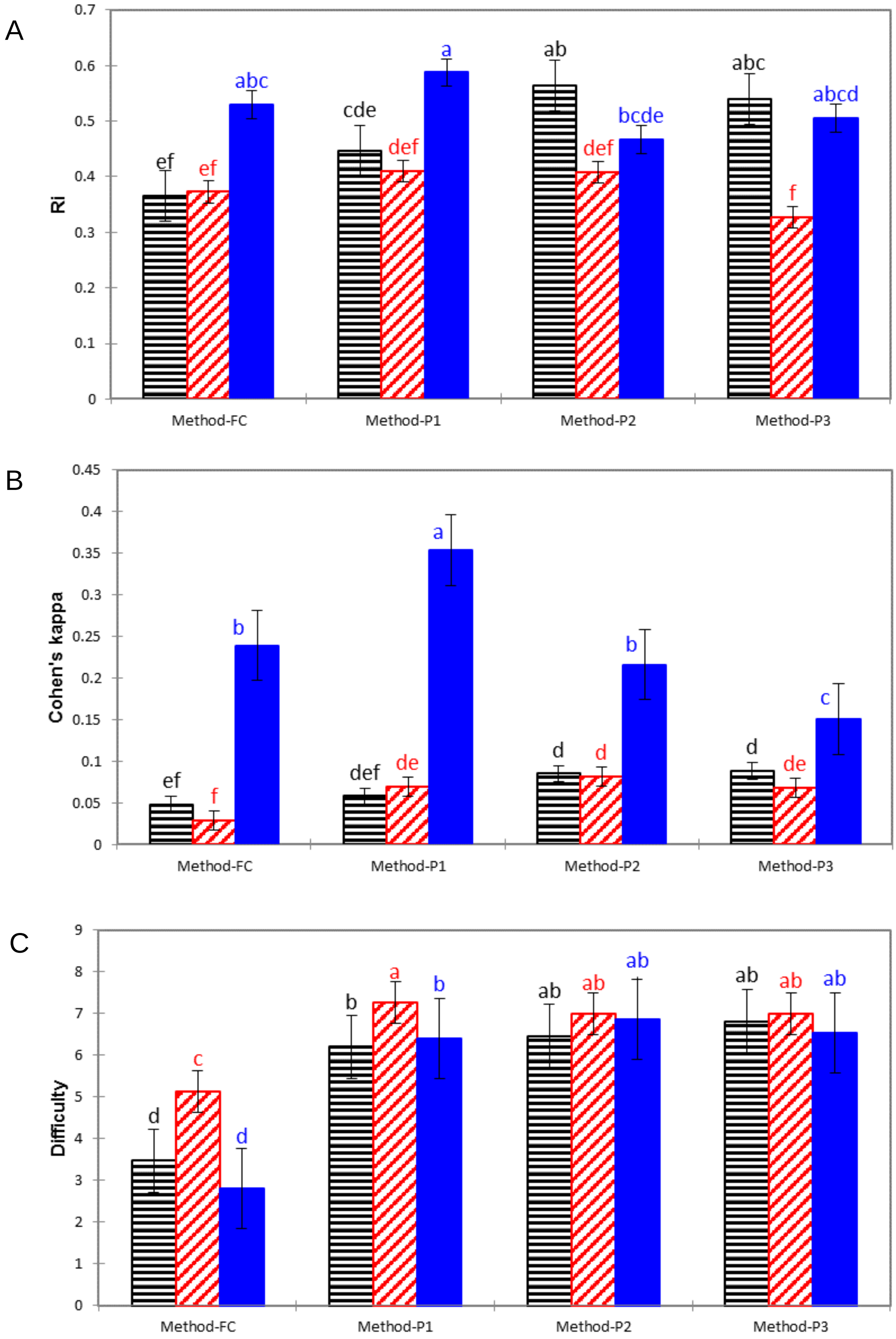


Fig. 2. ANOVA results showing significant differences in panel performance and task difficulty between frequency of attribute citation (FC) and pivot profile (PP) when different pivots were used to analyse three different wine sets with varying within-set variability. (A) Consensus measured with Cohen's kappa coefficients, (B) repeatability measured by means of R_i coefficients and (C) difficulty measured on a 9-point scale. Sample sets are indicated on the bar graph as follows: Pinotage wines with low within sample set variation, Chenin Blanc wines with medium within-set variation and Sauvignon Blanc wine with extreme within-set variation.

4.3.1.2 Consensus

Panel consensus measured by means of Cohen's kappa coefficients ranged from poor to moderate, where values below 0.2 are considered poor, between 0.2 and 0.4 fair and between 0.4 and 0.6 moderate (Altman, 1991). The highest value measured was 0.55 and the lowest value 0.02.

The ANOVA results (Fig. 2B) clearly show that different trends were observed for the sample sets with different within-set variability in terms of average panel consensus. The panel consensus for the set with the small (Pinotage wines) and the set with medium (Chenin Blanc wines) within-set variation was poor with average Cohen's kappa coefficient of the panel below 0.2. Interpreting significant differences with such low values would be unwise.

It is interesting to note that the only data set with acceptable average panel consensus coefficients, above 0.2, were the set with extreme within-set variation. Cohen's kappa coefficients above 0.2 were observed for FC and PP except when the blend of the samples was used as pivot for which a significantly lower value of 0.17 was observed. The best consensus was achieved when P1 was used and was significantly higher than when FC was performed and when other pivot samples were used.

PP was experienced by the sensory judges as significantly more difficult to perform when compared to FC, irrespective of the complexity of the data set and the pivot sample used (Fig. 2C).

4.3.2 Product description and comparison of methods

When analysing the set with the lowest within-set variability (Pinotage wines), the following observations were made. The RV coefficients calculated between the CA configurations, constructed from the different PP data sets where the pivot sample was changed indicated acceptable similarity (Table 3) ranging from 0.52 to 0.83. However, the similarity between the FC configuration and PP configurations, corresponding to P1 (Fig. 3A) and P2 (Fig. 3B) as pivot samples, indicated low similarity with RV coefficients below 0.35 (Table 3). When a blend of all the samples was used as pivot sample, P3 (Fig. 3C), creating a centre sample, better similarity was observed with an RV coefficient of 0.60.

TABLE 3.
RV coefficients used to compare correspondence analysis (CA) plots obtained from frequency of attribute citation (FC) and pivot profile (PP) experiments where the pivot was changed (P1, P2 and P3).

Cultivar	RV coefficient pairwise comparisons			
Pinotage				
	P1	P2	P3	FC
P1	-	0.83	0.70	0.34
P2	0.83	-	0.52	0.28
P3	0.70	0.52	-	0.60
FC	0.34	0.28	0.60	-
Chenin Blanc				
	P1	P2	P3	FC
P1	-	0.44	0.75	0.66
P2	0.44	-	0.51	0.88
P3	0.75	0.51	-	0.68
FC	0.66	0.88	0.68	-
Sauvignon Blanc				
	P1	P2	P3	FC
P1	-	0.68	0.51	0.86
P2	0.68	-	0.28	0.95
P3	0.51	0.28	-	0.36
FC	0.86	0.95	0.36	-

Furthermore, large overlapping confidence ellipses indicated that no significant difference between samples could be observed when PP was conducted on this sample set (Fig 3A, B and C), despite the fact that the explained variance for the first two factors was well above 60% (Fig. 3). Confidence ellipses on the CA plot of the FC configuration were smaller and indicated that two of the samples were perceived significantly different from the other four samples (Fig. 3D). It is, however, interesting to note that the explained variances of the CA plots were higher for PP (Fig. 3A, B and C) than for FC (Fig. 3D).

Descriptors belonging to the same aroma families appeared more scattered on the CA plot, and showed less positive correlation with each other, for PP data than FC data. The most obvious and prominent cases occurred when extreme samples, P1 and P2, were used as pivot samples (Fig. 3A and B). When the blend, P3 (Fig. 3C) was used as pivot, aroma attributes belonging to the same aroma family grouped well together indicating acceptable positive correlation. Examples were: (1) “oaky”, “wooded”, “pencil shavings”, “toasted” and “burnt wood” belonging to the “wooded” aroma family and (2) “blackberry”, “blackcurrant”, “black fruit” (including all dark berries except “blackberry” and “blackcurrant”, “cherry”, “raspberry” and “strawberry” belonging to the “berry” aroma family.

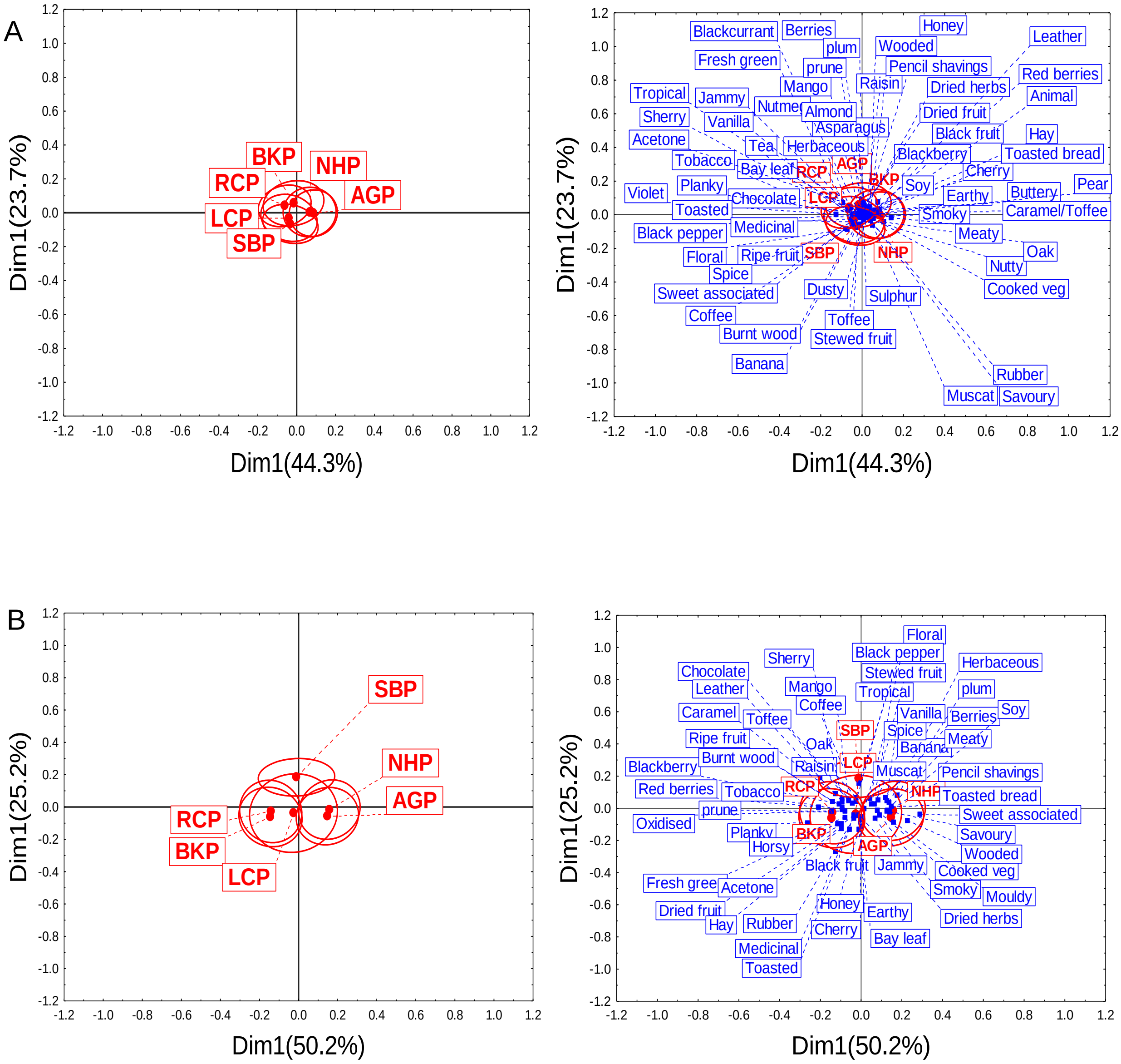


Fig. 3. Pinotage pivot profile (PP) and frequency of attribute citation (FC) correspondence analysis (CA) plots showing confidence ellipses. (A) PP with P1 as pivot sample, (B) PP with P2 as pivot sample, (C) PP with P3 as pivot sample and (D) FC data.

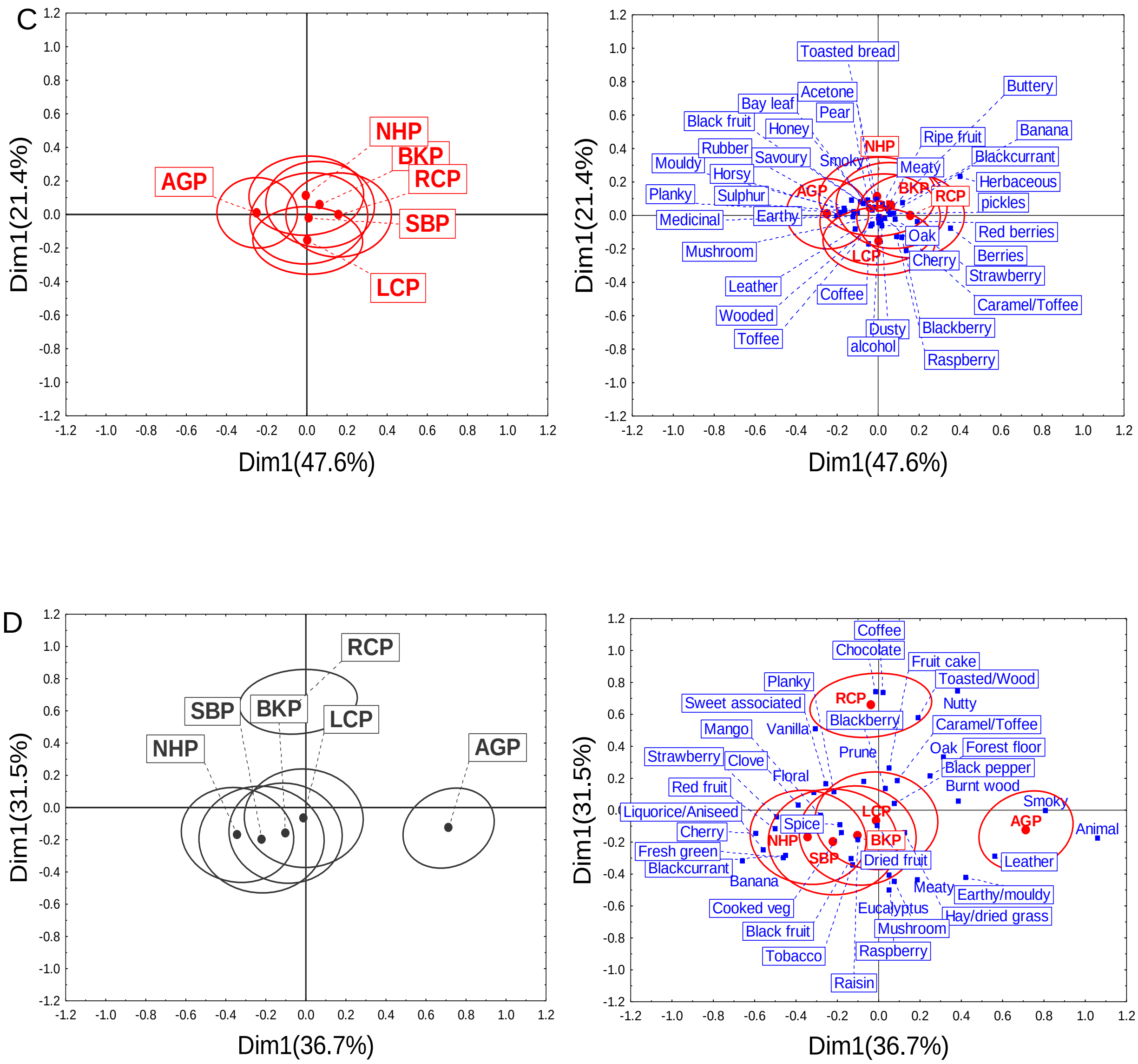


Fig. 3. Pinotage pivot profile (PP) and frequency of attribute citation (FC) correspondence analysis (CA) plots showing confidence ellipses. (A) PP with P1 as pivot sample, (B) PP with P2 as pivot sample, (C) PP with P3 as pivot sample and (D) FC data.

The data set with medium within-set variability (Chenin Blanc) produced CA plots (Fig. 4) with similar configurations for the PP (Fig. 4A, B and C) and FC data sets (Fig. 4D) with RV coefficients ranging from 0.66 to 0.88 (Table 3). In general, the differences between CA plots from PP data, where different pivot samples were used, were larger, with lower RV coefficients, than the differences between PP and FC. The RV coefficient between the CA plots constructed using P1 and P2 was 0.44 indicating dissimilarity. P2 had aroma characteristics that could overshadow other aroma nuances since it was described by words such as “vanilla”, “wooded”, “oaky”, “buttery” and “caramel” by many of the judges (Fig. 4B).

The confidence ellipses on this CA plot shows large overlap between samples. A possible explanation could be that it was difficult for the sensory judges to detect differences between the other samples when comparing samples to P2 which had intense extreme sensory characteristics. Confidence ellipses overlapped less frequently when a blend between the samples was used as pivot (P3), indicating clearer significant differences between samples (Fig. 4C). It is interesting to note that descriptors from the same aroma family are grouped well together on all CA plots obtained for this set. Examples were: (1) “sweet associated” characteristics such as “vanilla”, “caramel”, “honey” and “toffee” and (2) “oak”, “wooded” and “planky” which were positively correlated.

From the CA plots constructed for the data set with extreme within-set variability (Sauvignon Blanc) it can be seen that the variation explained by dimension 1 and 2 is above 75% (Fig. 5) which is regarded as high for sensory data. Clear separation between the confidence ellipse of the pivot sample and the other samples were visible, but the overlapping confidence ellipses of the other samples indicated similarity and an inability of the panel to discriminate between those samples. It is possible that the uniqueness of the pivot sample caused the high explained variance and over–shadowed the variation between other samples causing a loss of separation between them.

The RV coefficients between the different sample sets varied from 0.28 to 0.95. Even though the effect of the pivot is overshadowing sensory characteristics, the RV coefficients between the CA maps when the extreme samples were used as pivots, P1 (Fig. 5A) and P2, (Fig. 5B) and the FC CA map are above 0.86 (Table 3). The low RV coefficient of 0.28 between CA maps constructed from P3 (Fig. 5D) and P2 (Fig. 5C), 0.51 between P1 and P3 and 0.36 between FC and P3 originates from the fact that one of the samples, TSL, was profiled differently when P3 was used as pivot sample.

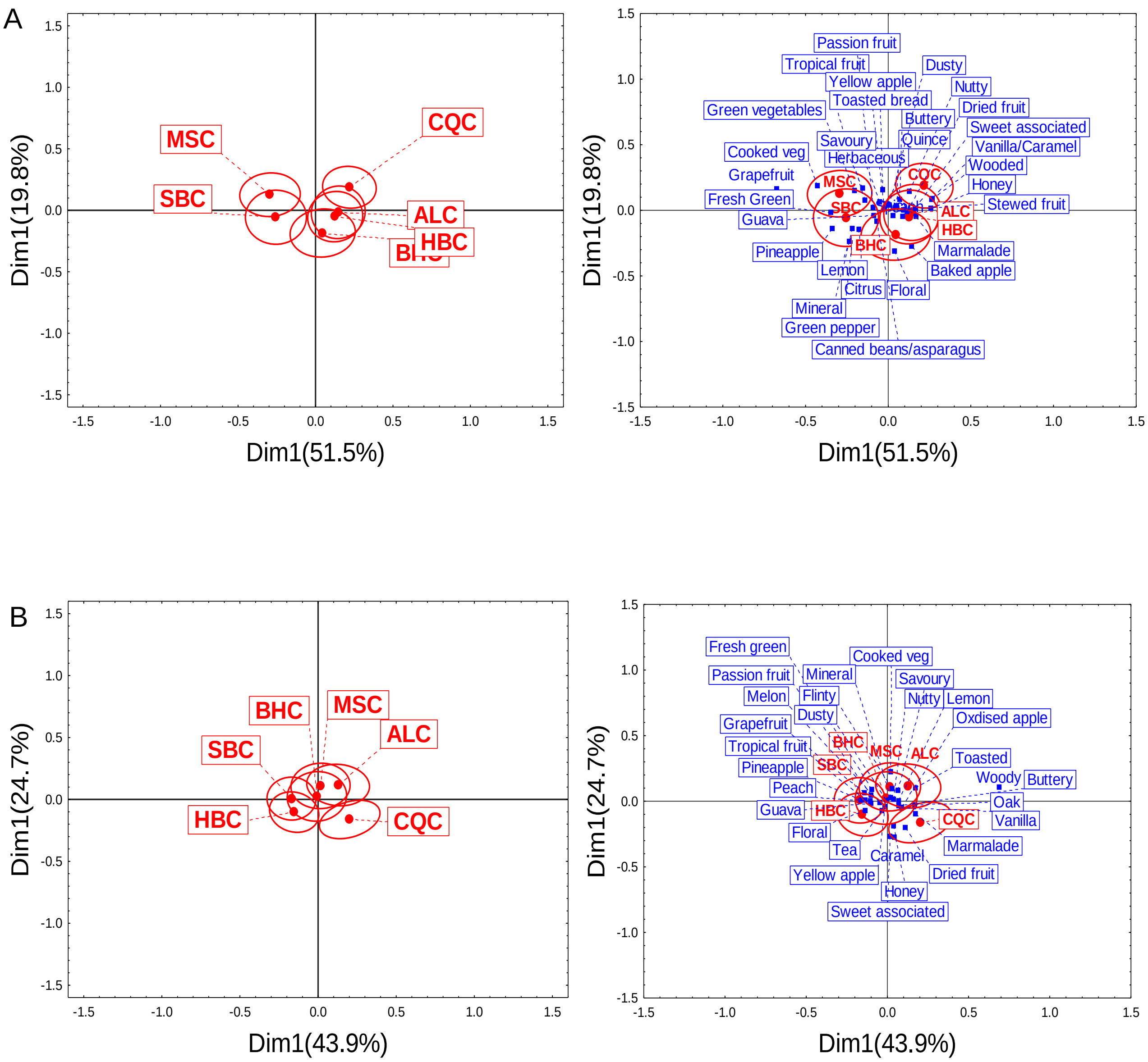


Fig. 4. Chenin Blanc pivot profile (PP) and frequency of attribute citation (FC) correspondence analysis (CA) plots showing confidence ellipses. (A) PP with P1 as pivot sample, (B) PP with P2 as pivot sample, (C) PP with P3 as pivot sample and (D) FC data.

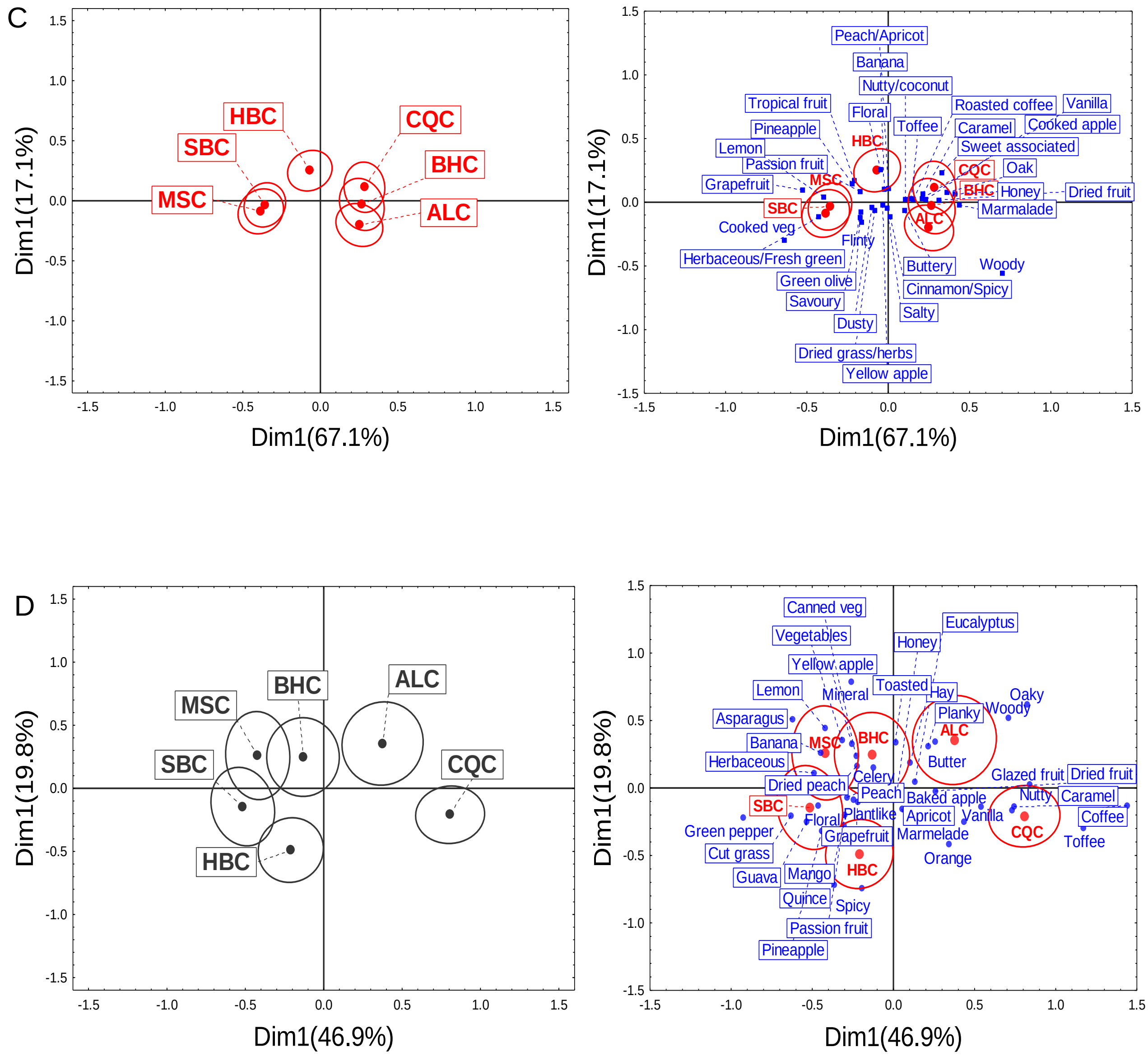


Fig. 4. Chenin Blanc pivot profile (PP) and frequency of attribute citation (FC) correspondence analysis (CA) plots showing confidence ellipses. (A) PP with P1 as pivot sample, (B) PP with P2 as pivot sample, (C) PP with P3 as pivot sample and (D) FC data.

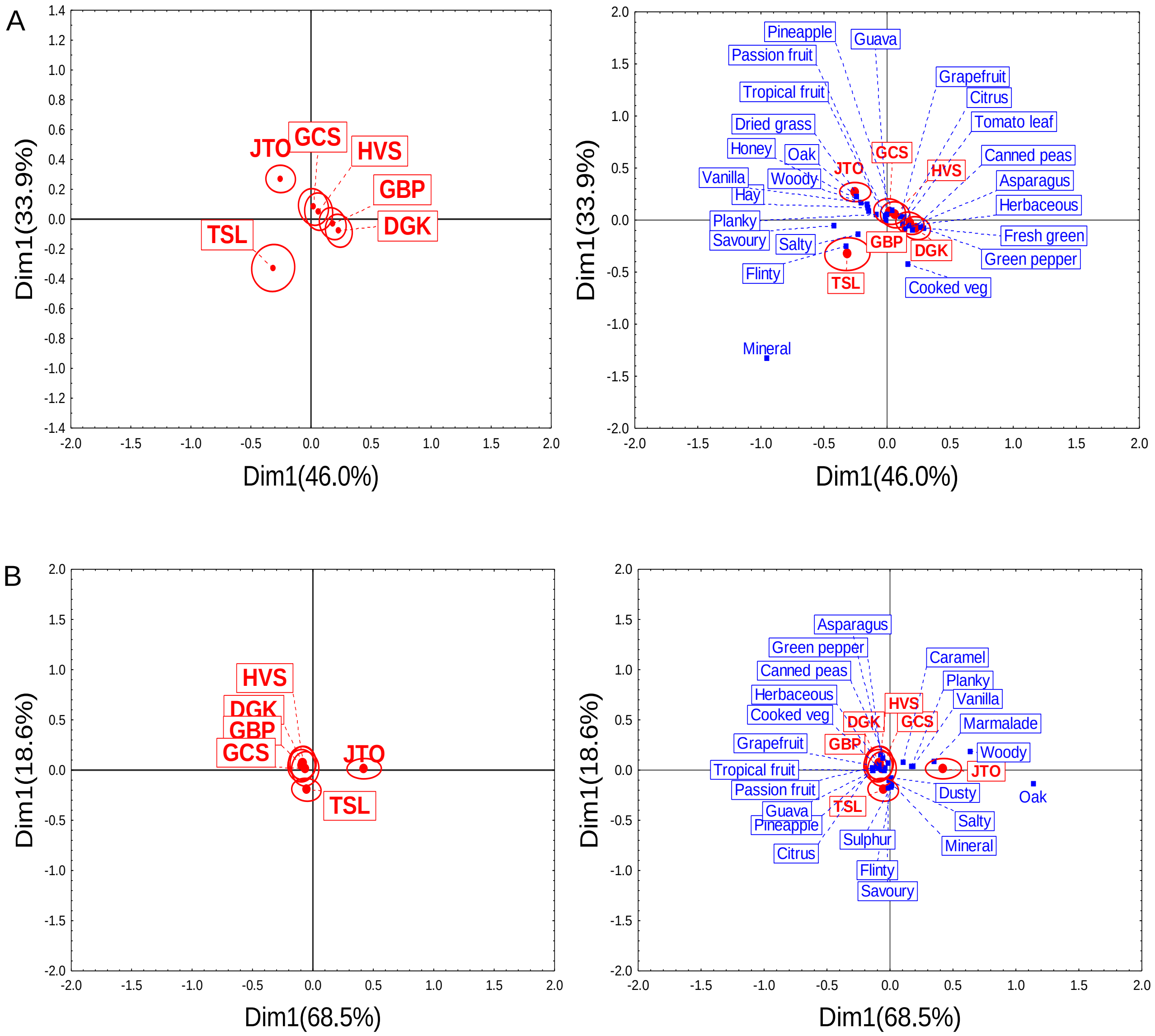


Fig. 5. Sauvignon Blanc pivot profile (PP) and frequency of attribute citation (FC) correspondence analysis (CA) plots showing confidence ellipses. (A) PP with P1 as pivot sample, (B) PP with P2 as pivot sample, (C) PP with P3 as pivot sample and (D) FC data.

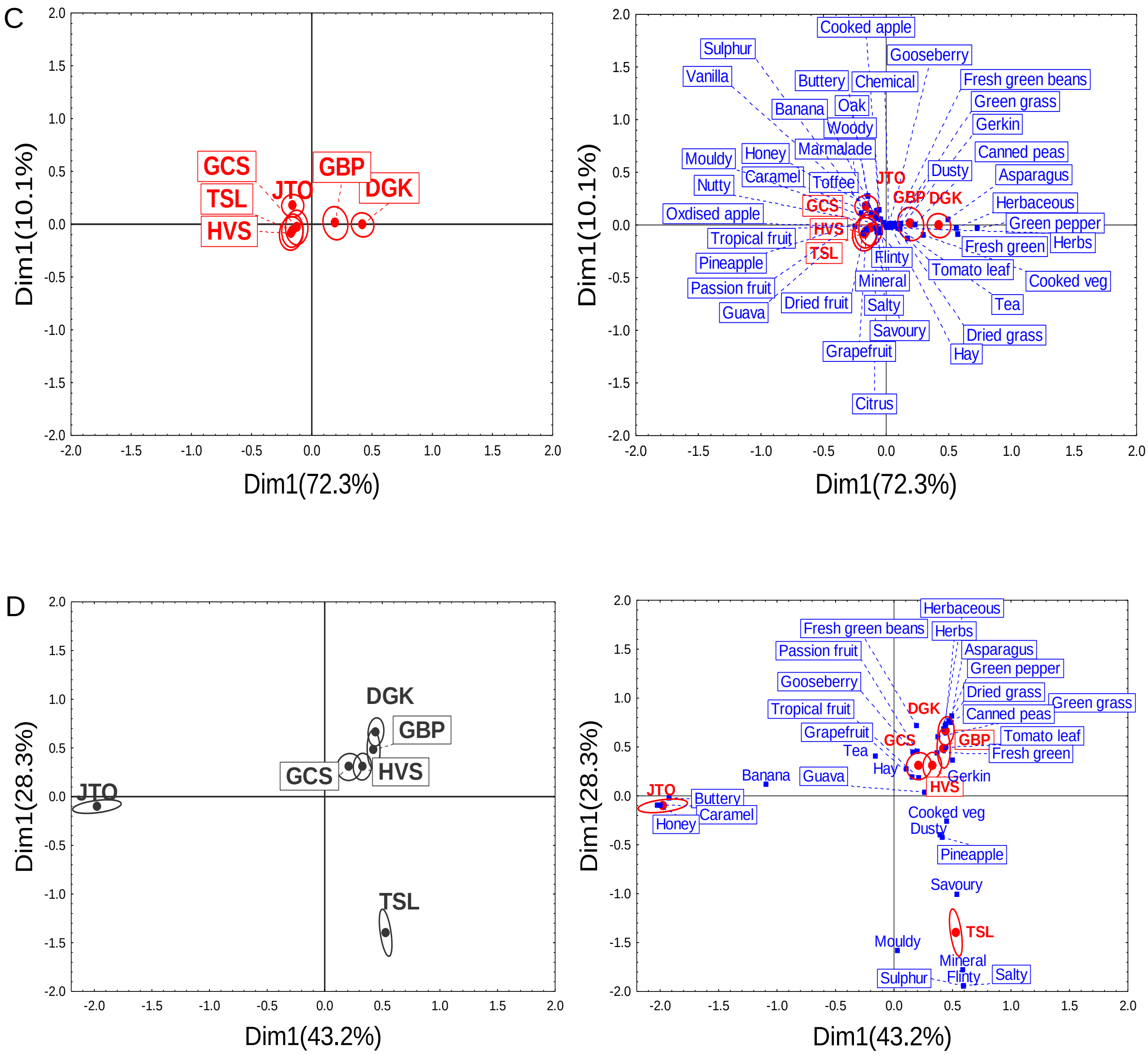


Fig. 5. Sauvignon Blanc pivot profile (PP) and frequency of attribute citation (FC) correspondence analysis (CA) plots showing confidence ellipses. (A) PP with P1 as pivot sample, (B) PP with P2 as pivot sample, (C) PP with P3 as pivot sample and (D) FC data.

4.4 Discussion

Pivot profile can be a useful technique to use for the profiling of complex products such as wine (Thuillier et al., 2015) and beer (Lelièvre-Desmas et al., 2017). The objective of this study was to evaluate PP critically, for the profiling of complex matrices, comparing PP to FC, a well-established descriptive method (Campo et al., 2008). More specifically, to determine whether one of these techniques offer better discrimination between samples than the other one. To investigate these aspects thoroughly three data sets with different levels of within-set variability were analysed using a trained panel and CA was performed to obtain multivariate sensory maps.

Inspecting these CA plots the following conclusions were made. The variance explained by the first two factors were higher for PP than FC regardless of the complexity of the data set or the choice of pivot, indicating that differences between samples were described well when PP was performed. However, descriptors belonging to the same aroma family appeared more scattered on the PP CA plots than the FC CA plots (Fig. 3, Fig. 4 and Fig. 5). Confidence ellipses, calculated by means of bootstrapping, were added to the CA results as suggested by Lelièvre-Desmas et al. (2017) to understand the significance of product differences described by PP and FC. The size of the confidence ellipses covered larger areas for PP than FC, showing that fewer samples were perceived significantly different when PP was performed than when FC was performed.

In addition, confidence ellipses shed light on perceived product differences when within-set product variability was varied. It is clear that the smaller the within-set variation between samples, the larger and the more severe the overlap of confidence ellipses. Due to severe overlap of large confidence ellipses for the data set with small within-set variation, it is not recommended to use PP to analyses such a set of products, even though it was suggested by Lelièvre-Desmas et al. (2017) that PP might be better suited for more homogenous spaces. However, for the sets with medium and large within-set variability, the confidence ellipses overlapped less frequently when a blend of the samples was used as pivot sample rather than other samples from the set. It can, therefore, be concluded that more samples were perceived significantly different when the blend was used as the pivot and the within-set variation was medium or extreme.

The similarity between sample configurations on the CA plots was tested by means of RV coefficients. Similarity between the different PP configurations, when the pivot sample was changed, and FC configurations differed for data sets with different degrees of within set variation. Similar product configurations were obtained when the pivot was changed for the data set with small within-set variation, indicating that the choice of the pivot was not crucial. This observation is in line with observations made by Thuillier et al. (2015) when proposing pivot profile and Lelièvre-Desmas et al. (2017) when testing the stability of the product space by varying the pivot sample used as well as the within-set variability. However, similarity between PP configurations and the FC configuration was poor, except when a blend of all the samples was used as pivot. Thuillier et al. (2015) proposed using the blend as the pivot to create a centre sample, containing a wide range of sensory properties that span the sensory space, to which other samples are compared. Lelièvre-Desmas et al. (2017) noted that the idea of using a blend as pivot might be well adapted for profiling of homogeneous spaces which is confirmed in this study.

It is important to keep in mind that few significant differences between samples were observed for this set when PP was conducted. Even though Lelièvre-Desmas et al. (2017) found that PP might be more suited for homogenous spaces than heterogeneous spaces, this set was probably too homogeneous for profiling using pivot profile. Lelièvre-Desmas et al. (2017) however did not compute confidence ellipses by means of bootstrapping to validate product discrimination. Furthermore, the lack of quantification of the degree of within-set similarity of a sample set causes subjective interpretation of what small, medium and extreme within-set variability is.

If the set, regarded by Lelièvre-Desmas et al. (2017) as the set with small within-set variation, is compared to the set defined in this study as the set with medium within-set variation then remarkably similar results were obtained.

Similarity between FC and PP data sets was good, with RV coefficients above or close to 0.7, regardless of the pivot used when the sample set with medium within-set variation was subjected to MFA. It is interesting to note that higher RV coefficients, indicating better similarity, were observed between the different PP data sets where different pivot samples were used and FC data than when these PP data sets were compared to each other. This was observed for the data set with large within-set variation as well with an exception when a blend of all the samples was used as pivot. In that case, poor similarity, with low RV coefficients, was observed with the FC CA configuration and PP CA configurations, originating from different pivot samples. Visual inspecting of the CA plots revealed that one sample, in particular, was described differently and was consequently located differently relative to the other samples. It was noted by El Ghaziri and Qannari (2015) that RV coefficients will not provide a good estimate of the similarity of two spaces if one sample is configurationally not in the same position on both maps. In other words, if one sample is perceived differently the RV coefficient will be low even though all the other samples were perceived similarly and will not provide a good estimate of the overall similarity between two configurations, in this case sensory spaces.

The question, however, remains why this sample was perceived differently. There could be two factors playing a role here: a physiological perception factor and a methodological limitation to use vocabulary that will distinguish wines from each other. It was noted by Lelièvre-Desmas et al. (2017) that the vocabulary might change when a different pivot is used and therefore, suggested that PP might not always be the best method to obtain detailed sensory characterisation of samples but should rather be used to compare samples.

An aspect of PP that still require attention is the testing of panel performance. In previous studies where PP was used as profiling technique the measurement of panel performance did not receive sufficient attention. Thuillier et al. (2015) proposed the method, but did not propose a strategy to measure panel performance since the focus of that study was on simulation where panel heterogeneity was set as a parameter. It would, therefore, not make sense to test panel performance on the simulation data. Fonseca et al. (2016) and Esmerino et al. (2017) performed PP using consumers as panellists without investigating possible segmentation or testing the performance of individuals. Testing panel repeatability was not possible with the data obtained during the consumer studies, as judges did not repeat the test. Testing panel performance, however, when consumers perform the test is not common and deemed irrelevant due to the large number of participants that increase the statistical power of the experiment. However, investigating panel segmentation and individual differences could provide valuable insights on how consumers profile product when performing PP. Lelièvre-Desmas et al. (2017) proposed a strategy to evaluate global panel consensus and repeatability when performing PP but the authors also acknowledged that more work needs to be done in this field.

In this study panel repeatability was measured using the Ri value and consensus using Cohen's kappa coefficients. Both these measures provide useful insights in panel performance, but are probably too strict since they only take exact matches of attributes as good consensus between two judges. It could make sense to penalise judge less, or not at all when two judges use slightly different attributes, but still belong to the same odour family. In order to incorporate this idea into panel performance testing, more work is required in the field of sensometrics.

Critical investigations of panel performance measurements and a proposed workflow to measure consensus and repeatability for PP and FC, similar to the work published by Tomic et al. (2007) and Tomic et al. (2010) for DA, could be valuable additions to the methodology development of rapid methods.

4.5 Conclusions

The sensory space generated using PP for wine sample sets with medium within-set variability using a central sample, as the pivot, is comparable to results obtained with FC. For these type of sample sets, PP was robust for profiling wines using a trained panel. The choice of the pivot sample was not crucial in terms of the stability of the sensory space. PP can be used as an alternative for FC and can be particularly useful to use as a benchmarking tool.

From this study, it could be observed that when sample sets with very low variability between samples are tested, FC is a more sensitive technique to use due to the large overlap between confidence ellipses of different samples that can occur on the CA plots of PP data.

Sample sets containing samples with large within-set variation might be less suited for analysis by PP and FC results will probably be more stable. Re-testing of this hypothesis is required to confirm findings from this study since it is unclear why one sample in particular was perceived different when the pivot sample was changed.

It would be interesting to compare PP to other rapid sensory methods such as sorting and particularly reference-based rapid sensory methods such as polarised sensory positioning (PSP, Teillet et al., 2010) and polarised projective mapping (PPM, Ares et al., 2013) particularly as benchmarking tools. The panel repeatability was comparable and good for both PP and FC. A workflow to test panel consensus and repeatability will add value to the PP methodology. Panel performance testing is currently a shortcoming of the PP methodology.

4.6 References

Abdi, H., & Valentin, D. (2014). Encyclopedia of Measurement and Statistics Multiple Factor Analysis. *Encyclopeia of Measurement and Statistics*, 658–664.

Abdi, H., Williams, L. J., & Valentin, D. (2013). Multiple factor analysis: Principal component analysis for multitable and multiblock data sets. *Wiley Interdisciplinary Reviews: Computational Statistics*, *5*, 149–179.

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. *Poster presented at the 7th Pangborn Sensory Science Symposium.* 16 August 2007, Minneapolis, USA. Delegate Manual, pp. 12–418.

Altman, D.G. (1991). *Practical statistics for medical research*. London England: Chapman and Hall, p404.

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.

Bester, I., (2012). Classifying South African Chenin blanc wine styles. MSc thesis, Stellenbosch University. http://hdl.handle.net/10019.1/18048

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

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.

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.

Cohen, J. A. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement. 20*, 37–46.

Dehlholm, C., Brockhoff, P. B., & Bredie, W. L. P. (2012). 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.

Deneulin, P., Reverdy, C., Rébénaque, P., Danthe, E., & Mulhauser, B. (2018). Evaluation of the Pivot Profile©, a new method to characterize a large variety of a single product: Case study on honeys from around the world. *Food Research International*, *106,* 29-37.

El Ghaziri, A., & Qannari, E. M. (2015). Measures of association between two datasets; Application to sensory data. *Food Quality and Preference*, *40*, 116–124.

Esmerino, E. A., Tavares Filho, E. R., Thomas Carr, B., Ferraz, J. P., Silva, H. L. A., Pinto, L. P. F., Freitas, M. Q., Gruz, A. G., Bolini, H. M. A. (2017). Consumer-based product characterization using

Pivot Profile, Projective Mapping and Check-all-that-apply (CATA): A comparative case with Greek yogurt samples. *Food Research International*, 99, 375-384.

Fonseca, F. G. A., Esmerino, E. A., Filho, E. R. T., Ferraz, J. P., Cruz, A. G., & Bolini, H. M. A. (2016). Novel and successful free comments method for sensory characterization of chocolate ice cream : A comparative study between pivot profile and comment analysis. *Journal of Dairy Science, 99*, 3408–3420.

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability : Key concepts, approaches and applications. *Research in Social and Administrative Pharmacy, 9*, 330–338.

Hanekom, E., (2012). Chemical, sensory and consumer profiling of a selection of South African Chenin blanc wines produced from bush vines. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/71812

Laing, D. G., & Glemarec, A. (1992). Selective attention and the perceptual attention of odor mixtures. *Physiology and Behavior, 52*, 1047–1053.

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

Lelièvre-Desmas, M., Valentin, D., & Chollet, S. (2017). Pivot profile method : What is the influence of the pivot and product space? *61*, 6–14.

McFie H. J. H., & Thomson D. M. H. 1988. Preference mapping and multidimensional scaling. In: J. R. Piggott (ed.), *Sensory Analysis of Foods*. Elsevier Applied Science, New York, pp. 381–409.

McHugh, M.L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*, 276 – 282.

Noble, A.C., Arnold, R.A., Buechsenstein, J., Leach, E.J., Schmidt, J.O. & Stern, P.M. (1987). Modification of a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture, 38*, 143- 146.

Peryam, D. R., & Pilgrim, F.J. (1957). Hedonic scale method of measuring food preference. *Food Technology*. *11*, 9-14

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

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

Symoneaux, R., Galmarini, M.V., & Mehinagic, E. (2012). Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples. *Food Quality and Preference*, *24*, 59–66.

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.

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.

Tomic, O., Nilsen, A., Martens, M., & Næs, T. (2007). Visualization of sensory profiling data for performance monitoring. *LWT - Food Science and Technology*, *40*, 262–269.

Thuillier, B., Valentin, D., Marchal, R., & Dacremont, C. (2015). Pivot© profile: A new descriptive method based on free description. *Food Quality and Preference*, *42*, 66–77.

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*, 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 Antwerpen, L., (2012). Chemical and sensory profiling of dry and semi-dry South African Chenin blanc wines. MSc thesis, Stellenbosch University, http://hdl.handle.net/10019.1/71853

Williams, A.A. & Langron, S.P. (1984). The use of free-choice profiling for the evaluation of commercial ports. *Journal of the Science of Food and Agriculture*, 35, 558–568.

Chapter 5

Research results

Sorting in Combination with Quality Scoring: A Tool for Industry Professionals to Identify Drivers of Wine Quality Rapidly

This manuscript was published in **The South African Journal of Enology and Viticulture**

5. Research results

Sorting in Combination with Quality Scoring: A Tool for Industry Professionals to Identify Drivers of Wine Quality Rapidly

J. Brand[a,b*], M. Kidd[c], L. van Antwerpen[d], D. Valentin[e], T. Næs[f], H.H. Nieuwoudt[b]

[a]*Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[b]*Institute for Wine Biotechnology, Department of Viticulture and Oenology, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[c]*Centre for Statistical Consultation, Department of Statistics and Actuarial Sciences, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[d]*Graham Beck Winery, R60, Robertson, 6705, South Africa*

[e]*Centre des Sciences du Goût et de l'Alimentation, AgroSup Dijon, CNRS, INRA, University of Bourgogne, Franche-Comté, F-2100, Dijon, France*

[f]*Nofima AS, Raw Material and Process Optimisation, Oslovegen 1, P.O. Box 210, N-1431, Ås, Norway*

Keywords: Rapid sensory analysis, sorting, quality scoring, drivers of quality, DISTATIS, Sauvignon Blanc wine, industry professionals

Condensed title: *Combining Sorting with Quality Scoring to Determine Drivers of Wine Quality*

Abstract

Quality plays an important role in the criteria directing wine product development. The evaluation of sensory characteristics associated with wine quality, as perceived by industry professionals, is therefore, important. We investigated the suitability of the free-sorting sensory evaluation method, in combination with wine quality scoring using a 20-point scoring system, to determine the drivers of quality. Eight commercial South African Sauvignon Blanc wines were assessed by a panel of 24 wine industry professionals. Free sorting with a verbalisation step to describe the groups, followed by quality scoring using score sheets routinely used in the wine industry, was performed. A multivariate sensory map was constructed using DISTATIS to explain the similarities and differences amongst the set of wines. Correspondence analysis (CA) was applied to the group descriptors, and CA deviates were calculated. Pearson's correlation coefficients between CA deviates and the quality scores were calculated to identify the drivers of quality. Significant differences in quality were observed between the wines. The sensory

attributes “passion fruit”, “green pepper”, “peas”, “asparagus” and “green” were frequently cited by the panel for the wines that received the highest average quality scores, and these attributes were identified as drivers of quality. In this study, a workflow is presented that combines sorting and quality scoring to investigate the relationship between sensory attributes and quality scores to identify the drivers of wine quality. Industry professionals and research environments can use this workflow to determine drivers of wine quality in a single evaluation session.

5.1 Introduction

Sensory characteristics are important intrinsic factors that influence the perceived quality of wine and play a crucial role in product development. During the blending process of wine production, for example, the sensory drivers of quality can be decisive factors guiding the process. Identifying the consumer target market and target price for a product depends largely on intrinsic characteristics such as colour, taste, mouthfeel, odour, aroma and flavour which contribute to the sensory dimension of the quality. As a result, industry experts, such as winemakers and brand managers, routinely conduct evaluations of wine quality, an activity that is especially important for high-quality wines. During these evaluations wines are typically rated for quality only and description are rarely provided. Occasionally, during informal tasting group social or training events, words or phrases describing the sensory characteristics are provided additionally. Several studies have been conducted to better understand the dimensions of wine quality and improve the methods used. The aim is to evaluate the effectiveness of the methods used by the wine industry and to propose new strategies to measure wine quality (Botonaki & Tsakiridou, 2004; Verdú Jover et al., 2004; Parr et al., 2006; Charters & Pettigrew, 2007; Torri et al., 2013; Sáenz-Navajas et al., 2015, 2016; Valentin et al., 2016). It is clear that wine quality assessment requires extensive attention and optimisation.

5.1.1 Quality assessment of wine

Quality is an abstract concept that is difficult to define (Ziethalm et al., 1988). Various methods have been proposed and tested in the last 10 to 15 years to measure wine quality. The majority of these methods focused on the perception of quality by the consumer. A few examples are papers by Botonaki and Tsakiridou (2004) and Charters and Pettigrew (2007).

Botonaki and Tsakiridou (2004) used self-administered questionnaires to obtain insights into Greek consumers’ attitudes towards quality attributes by taking their general knowledge of the Greek quality certification system and “destination of origin” into account. Consumers’ willingness to pay more for quality certified wines was also investigated. Verdú Jover et al. (2004) proposed and validated a 21-point scale to measure the dimensions of wine quality by proposing two different scales for intrinsic and extrinsic quality measurements. It was concluded that the methods were suitable to evaluate wine quality using both novice consumers and connoisseurs (expert consumers). Charters and Pettigrew (2007) used qualitative data obtained from questionnaires and focus groups to understand the complexity and dimensions of the quality perception of Australian consumers based on their level of involvement with the product. From these studies, it became clear that consumers' quality perception is a multidimensional concept and is a crucial measurement, since the translation of consumer demands into product specification leads to the development of products accepted by the consumers (Bredahl et al., 1998; Verdú Jover et al., 2004).

Measuring quality as perceived by consumers is not always possible, due to logistical matters such as cost implications and the phase of production when the quality measure is needed, for example during product development. In such cases, experts' opinions of product quality can be used as a first measure. Few studies have investigated the perception of quality by wine industry experts. Parr et al. (2006) compared a 20-point scale to a 100-point scale. Both scales are routinely used in the industry for judging at wine competitions. No significant differences were found in the data obtained from the two scales concerning product variation or judge variability. Lattey et al. (2009) used a 20-point scale to capture the quality perceptions of Australian winemakers and compared these to consumers' acceptance of the wines. It was shown that the winemakers and consumers used different criteria when evaluating quality. Interestingly, wines that obtained higher average consumer liking scores also obtained higher average quality scores from the experts. Sáenz-Navajas et al. (2013) came to a similar conclusion when the effect of consumers' culture and their levels of expertise on the sensory drivers of the quality of red wines were investigated. These authors concluded that experts and consumers do not rate quality the same, and reported that the quality rating was dependent on the level of expertise of the judge.

Torri et al. (2013) adapted the nine-point hedonic liking scale and proposed a quality scale ranging from "very poor quality" to "excellent quality" to assess wine quality using experts (oenologists and wine producers) as sensory judges. The experts' quality measurements were compared to the consumers' liking of the products. Both groups' abilities to differentiate between wines using a rapid sensory profiling method, projective mapping (Napping), were tested. The results showed that consensus amongst the consumers was low regarding perceived differences and similarities between wines and was driven by liking. The authors postulated that experts use a common language to describe samples based on their previous experiences of high-quality products and thus differentiate between products based on quality.

Sáenz-Navajas et al. (2016) used an unstructured line scale to assess the effect of tasting conditions referred to global perception of all modalities simultaneously versus isolating the modalities. The three modalities, namely visual, orthonasal (aroma) and in-mouth perception (flavour, taste and mouthfeel) were evaluated separately. Quality perception was found to be dependent on the evaluation conditions. It was concluded that the global quality rating was based on the perceived quality during tasting as well as cognitive information obtained during technical training of the winemaking process, rather than the sum of the in-mouth, odour and visual perception of the sample. For example, white wine with a yellow colour was believed to be of lower quality than white wine with a green tint due to the cognitive information namely the technical knowledge that winemakers have of the production process rather than the perception of the wine during tasting. Lastly, it was concluded that the olfactory properties (the “nose”) of the wines had a stronger and more important influence on the overall global quality than the visual and in-mouth perceptions.

In another study the relationship between wine quality and colour in Pinot Noir wine was investigated (Valentin et al., 2016). These authors used a 10-point scale anchored at “poor” and “good” to assess overall wine quality. In addition, sensory attributes describing the “nose”, the “palate” (including “balance” and “structure”) and “typicality” was measured in the same session after the quality rating. During the final assessment the colour of the wines were evaluated by rating “colour/hue”, “colour intensity” and “brightness”, to relate wine quality to colour. It was found that wine colour was not a major contributor towards Pinot Noir wine quality, while the perceived “balance” and “structure” of the wines was important. The perceived “balance” and “structure” was correlated with chemical parameters such as pH, ethanol, sugar content, astringency and acidity. Presenting wines in clear glasses as opposed to black glasses lead to higher quality scores.

From literature, it is clear that wine quality is a complex and abstract concept. It has many dimensions and can be approached from different angles, for example from a consumer viewpoint or an industry professional perspective. The methods used previously in research to measure quality were selected based on the specific research question asked and the aspects of quality measured. When measuring quality, a complex and abstract concept, it is essential to choose a method that is fit-for-purpose. It is rarely necessary or practical to evaluate all the dimensions of quality and from both the perspectives of the consumers and industry professionals. Testing the quality perception from a consumer viewpoint can be used to direct product development and supplement the development of a marketing strategy knowing what the target consumer wants. On the other hand, an expert’s initial quality assessment during product development, from a production perspective, can be useful, for example, as a benchmarking tool. In this study the sensory dimension of quality as perceived by industry professionals was studied. Considering the methodologies discussed in literature, the 20-point scale was chosen for this study for two reasons. (1) This method is familiar to the South-African wine industry professional and (2) no proof could be found that other methodologies will provide better results, e.g. Parr et al. (2006) found no significant differences between results when using the 20-point and 100-point scales.

It is, however of interest for wine industry professionals to, in addition to quality scoring, describe the sensory characteristics of the evaluated samples due to the fact that quality is based partially on those characteristics.

5.1.2 Rapid sensory profiling methods for alcoholic beverages

In addition to quality assessment, industry professionals can describe the intrinsic sensory properties of wine products, due to constant practice gained from frequent participation in informal and formal wine tasting events, as a result of their work experience. During informal tastings winemakers, discuss wines amongst themselves providing a few words to describe each wine, not following a specific sensory method or applying statistical data analysis to their descriptions. When guided during sensory evaluation sessions industry professional can describe wines using formal sensory analysis methods, for example rapid methods such as sorting.

Since the late 1990's and early 2000's, several rapid sensory profiling methods were proposed, as recently reviewed (Valentin et al., 2012; Varela & Ares, 2012). A number of these methods were tested, adapted and developed specifically for profiling of wine and alcoholic beverages using industry experts as sensory judges. These methods include check-all-that-apply (CATA) and it's variants "Pick-k-attributes" (McCloskey et al., 1996; Chollet & Valentin, 2000); pivot profile (PP) proposed by Thuillier et al. (2015); projective mapping (PM) including Napping (Pagès, 2003, 2005; Perrin et al., 2008; Perrin & Pàges, 2009; Torri et al., 2013) and sorting (Piombino et al., 2004; Ballester et al., 2005; Ballester et al., 2008; Bécue-Bertaut & Lê, 2011; Parr et al., 2015; Honoré-Chedozeau et al. 2017).

CATA is referred to as a verbal-based method (Valentin et al., 2012) and sensory judges select terms from a predetermined list to describe the test samples. CATA variants were successfully used to discriminate between wine samples (Chollet & Valentin, 2000). However, the number of terms on the lists must be carefully considered to prevent the list from being too long and tedious to use, or too short thereby excluding terms crucial to represent the sensory judges' perception of the products accurately. During CATA, samples are presented to the judges in a monadic way, i.e., one at a time. Obtaining a global, intuitive picture in one's mind of how the samples relate to each other is, however, difficult and for most judges impossible.
Another method, PP (proposed by Thuillier et al. 2015) was used to profile wine using industry experts as sensory judges. Wine samples are evaluated in pairs, one test sample and one reference called the pivot sample. Each wine is compared to the same pivot, thereby allowing for interpretation of differences between samples relative to a common reference. The main drawback of this methods is the choice of the pivot. Nevertheless, Lelièvre-Desmas et al. (2017) stated that: "the choice of the pivot has less influence than the within-set similarity between samples". This method, however, requires verbal expression of sensory perceptions, which can be difficult and is not an intuitive task.

Projective mapping (PM) techniques rely on the ability of the sensory judges to identify similarities and differences between samples in an intuitive manner prior to naming the sensory attributes (Pagès, 2003, 2005; Perrin et al., 2008; Perrin &Pagès 2009; Torri et el., 2013).

This is an easier task than verbalising sensory attributes as the first step which is the case for CATA and PP. The first step of a PM requires sensory judges to place samples on a surface, often an A2 or A3 sheet, in such a way that similar samples are placed close to each other and different samples far apart. During the second step, the sensory properties of each sample are described by assigning a few sensory words next to each sample to explain its position on the sheet.

The free sorting sensory method is based on the psychological theory that human beings routinely organise their environments intuitively and as part of daily life, by categorising objects according to similarity and dissimilarity (Neisser, 1987; Rosch, 1973). During the free sorting task, sensory judges receive all the samples simultaneously. They are asked to group similar products together and organise the groups in such a way that dissimilar samples appear in separate groups. They are allowed to create as many groups as they see fit to explain the similarities and dissimilarities of the samples presented to them. In order to obtain more information about the odour, aroma, taste and mouthfeel attributes responsible for the groupings, a verbalisation step is included and performed after sorting of the samples. Typically, judges are asked to write down the three to five attributes that best describe each group of wines in the verbalisation step (Faye et al., 2004; Chollet et al., 2011).

Research has shown that product descriptions using rapid methods generated results comparable to DA results (Valentin et al., 2012; Varela & Ares, 2012). Rapid methods are , therefore, suitable for the profiling purposes. Cartier et al. (2006) showed that sorting rendered similar results to descriptive analysis (DA) in a study where consumers were used as sensory judges. Industry experts are frequently used to perform sorting tasks as well. It is interesting to note that Ballester et al. (2008) reported that consumers and experts did not sort wines precisely the same. Experts distinguished between different cultivars better than consumers did by sorting the wines clearly into separate groups, suggesting that sorting performed by experts could provide results even more similar to DA than sorting performed by consumers.

Sorting is seen as an intuitive, rapid sensory analysis method, and it is faster to perform than DA and other techniques that involve panel training. It is, therefore, more cost-effective and time efficient. Sorting has been applied successfully to investigate the sensory characteristics of wine to investigate product similarities and differences. A number of studies have successfully used sorting to profile wines using industry professionals as sensory judges (Piombino et al., 2004; Ballester et al., 2005; Abdi et al., 2007; Parr et al., 2007; Campo et al., 2008; Bécue-Bertaut & Lê, 2011; Valentin et al., 2012; Varela & Ares, 2012).

Sorting was, therefore, the chosen method to describe the sensory characteristics of wine in combination with quality scoring to determine drivers of quality. The aim of this study was to develop a "ready-to-use" procedure for industry professionals to explore the sensory dimensions of quality, since quality is partially based on sensory characteristics. The proposed methodology consists of a descriptive step, sorting and a quality scoring step conducted during a single sensory evaluation session. The 20-point scoring method with which the industry professionals were familiar, was used. This study is the first, to our knowledge, where free sorting in combination with quality scoring has been proposed for this task.

5.2 Materials and methods

5.2.1 Wines

Eight commercial South African Sauvignon Blanc wines from the 2014 vintage and from different production regions were subjected to sensory evaluation (Table 1). All the wines were commercially available at the time of analysis and were certified by the South African Wine and Spirits Board.

Wines were chosen by industry professionals that regularly serve on Sauvignon Blanc tasting panels, including competition judging. The wines represented premium quality and unwooded South African Sauvignon Blanc wines. The industry professionals who selected the wines did not serve as judges during the evaluation of the wines tested in this study.

TABLE 1
Summary of the chemical analysis of the 8 commercial South African Sauvignon Blanc wines.

Wine code	Origin of grapes	Location of producer	Alc[a] % (v/v)	RS[b] (g/L)	pH	TA[c] (g/L)
A	Robertson	Robertson	12.5	1.5	3.50	8.5
B	Robertson	Robertson	12.6	3.3	3.39	7.5
C	Robertson	Robertson	13.4	1.9	3.04	6.9
D	Robertson	Robertson	12.9	1.8	3.24	6.7
E	Unknown	Franschhoek	14.0	1.9	3.37	6.7
F	Cederberg	Cederberg	13.0	2.4	3.40	6.8
G	West coast	West coast	13.6	1.3	3.35	6.7
H	Western cape: Stellenbosch, Elgin, Walker bay	Stellenbosch	13.6	2.1	3.41	7.0

[a]Alcohol, [b]Residual sugar, [c]Titratable acidity.

5.2.2 Panel

The sensory panel consisted of 24 judges, all wine industry professionals, of whom 67% were male and 33% female. The judges were between the ages of 23 and 60 years (average age: 35, median of ages: 34). The participants attended regular tasting events as part of their occupation. All the judges, except two individuals, had more than five years' experience as tasters in the wine industry. No training was provided before the sensory analysis, and the judges were not paid for their participation.

5.2.3 Methodology

The sensory evaluation was conducted in a well-ventilated, odour free and temperature-controlled tasting room of a local wine cellar. Wines were presented at an ambient temperature of 20°C in standardised international tasting glasses (ISO NORM 3591, 1977). Glasses were coded with random three-digit codes and covered with Petri dishes. Judges received 25 mL of each wine in a different order according to a Williams Latin square design. Participants were not allowed to communicate with each other during the session, and only received information at the end of the session.

Evaluation of the wines was conducted during a single session in a two-step process with a 10 min break in between. During the first step, a free sorting task with verbalisation, as described by Chollet et al. (2011), was performed to investigate similarities and differences between the sensory profiles of the wines. Judges had to freely describe their groups of wines using three to five words. They were requested not to use phrases, negative forms or intensity words; for example, phrases such as "not fruity" and "very fruity" were not permitted. Judges were not informed during the sorting task that quality scoring would follow. Quality scoring of the wines was performed in the second step..

Quality was scored separately for three different criteria, namely the appearance (total score 3); "the nose" (total score 7); and "the palate" (total score 10).

The criteria used for the evaluation of the appearance were clarity, colour depth and hue. The "nose" was explained to judges as orthonasal odour. "The palate" of the wine was defined as (1) the flavour, including retronasal aroma and the perception of basic tastes (sweet, sour, bitter) and (2) mouthfeel (concerning astringency, body) and (3) the length of the aftertaste.

The sum of the three individual scores was computed and represented the overall quality score out of 20.

The 20-point quality score scale was chosen to score quality since professionals commonly use this in the South African wine industry. The judges were, therefore, familiar with the method due to their work experience gained from judging at wine competition and other wine evaluation panels. Furthermore, since no training was provided to the judges, the assumption was made that the judges had the ability to use the quality scale in a similar way and were familiar with using the 20-point scale. It was noted in literature that in most cases experts score wine quality similarly (Torri et al., 2013; Sáenz-Navajas et al., 2016), therefore, justifying the above mentioned assumption.

5.2.4 Statistical analysis and visualisation of data

The analysis of the sensory data was conducted in three steps. Significant quality differences were identified as the first step, secondly similarities and differences between samples were investigated and lastly drivers of quality were identified.

Workflow of the statistical analysis of free sorting and quality scoring data to determine drivers of quality

Establishing significance of quality differences between samples

Step 1.	One way ANOVA	Quality scores

Investigating similarity and differences between samples

Step 2a.	DISTATIS and AHC	Sorting data similarity matrices
Step 2b.	Attribute projection onto DISTATIS compromise map	Attribute citations

Identifying drivers of quality

Step 3.	Pearson's correlation coefficients between Correspondence Analysis (CA) deviates and quality scores	Attribute citations and quality scores

Fig. 1. Schematic presentation of the three-step statistical workflow process. One-way ANOVA on the quality scores was performed as the first step to determine whether samples differed significantly. During the second step the sorting configuration was determined using DISTATIS, followed by AHC to identify groups of samples. Lastly sensory attributes that drove the quality scoring of the wines were determined correlating the correspondence analysis (CA) deviates with the quality scores.

5.2.4.1 Establishing significance of quality differences between samples

The first step of the data analysis process (Fig. 1) entailed the significance testing of the quality scores by means of one-way ANOVA using Statistica 13 (www.statsoft.com, Statsoft Inc.). Four separate ANOVA's were performed, respectively for the appearance, the "nose", the "palate" and overall quality. Pairwise comparisons were calculated, using the Fisher's LSD post-hoc test when the ANOVA results were significant.

5.2.4.2 Investigating similarities and differences between wines from the sorting configurations

Individual distance matrices showing the grouping of the wines by each judge in the sorting task was compiled. DISTATIS (Abdi et al., 2007) was performed directly on the distance matrices of the individual judges using R version 3.4.0 DistatisR (www.R-project.org), as shown in step 2a (Fig. 1). DISTATIS is a statistical method that takes many similarity or distance matrices, into account when analysing the similarity relationships between samples. It provides a visual representation where samples that appear close to each other on the plot are similar. Therefore, wines that were sorted into the same groups by many of the judges will appear close to each other and wines that were not grouped together will appear far from each other on the DISTATIS plot where the wine samples are represented. This plot is called the DISTATIS compromise map. DISTATIS has the advantage over other similar techniques that differences between the individual judges' data are represented on the compromise map by means of a confidence ellipse drawn around each sample (Abdi et al., 2007). In this study, the compromise map was used to: (1) analyse the data to investigate differences and similarities between wine samples and (2) to analyse the differences between the individual judges' data complimented by STATIS analysis performed in PanelCheck V1.4.2 (www.panelcheck.com, Nofima).

Agglomerative hierarchical cluster (AHC) analysis was performed using Statistica 13 (www.statsoft.com, Statsoft Inc.). AHC, using Ward's linkages and Euclidean distances, was applied to the coordinates of the first two principal components (PC1 and PC2) of the DISTATIS compromise map to visualise grouping of wine samples due to similarity. Differences between the sorting data of the individual judges were visualised by means of the confidence ellipses on the DISTATIS compromise map, as well as a STATIS analysis performed in PanelCheck V1.4.2 (www.panelcheck.com, Nofima).

5.2.4.3 Investigating similarity and differences between samples from the descriptors used to describe groups

The descriptors generated to describe each group of wines in the verbalisation phase were captured by constructing a contingency table (Step 2b, Fig. 1). The attributes were reduced by combining similar descriptors that were used by less than 20% of the panel. When no synonyms for a particular descriptor could be identified, that descriptor was not used for further data analysis, similar to the strategy used by Campo et al. (2008) and Chollet et al. (2011). An example of similar descriptors was "grass", "cut grass" and "fresh green" notes. Two sensory research scientists and one industry professional reduced the descriptors independently, and discussed the outcomes of the descriptor reductions. The criteria for disagreement on descriptor reduction was to reach consensus through discussion. Where consensus could not be reached through discussion the opinion of a fourth person, a researcher in oenology and viticulture who frequently worked in collaboration with the sensory team was acquired. The industry professional was a member of the tasting group who performed the sensory analysis but did not take part in this experiment as a judge.

The number of times a descriptor was used to describe a wine was counted. This was done for all the descriptors for all the wines for the reduced set of descriptors. The sum of the citations over all the judges for each descriptor for each wine, was compiled with the wines in the rows and the attributes in the columns of the contingency table. The number of judges who cited an attribute for a wine was tabulated at the intersection between the row of that wine and column of that attribute.

Pearson's correlation coefficients between the attributes and the DISTATIS product coordinates were calculated using the contingency table and the data of the first two dimensions of the DISTATIS analysis. Two Pearson's correlation coefficients were thus calculated for each descriptor; one between each attribute frequencies and the coordinates of the first dimension, and the second between the attribute frequencies and the coordinates of the second dimension. These correlation coefficients were projected as the x- and y-coordinates of the descriptors onto the DISTATIS compromise map to obtain a plot representing the similarity information from the sorting exercise as well as the descriptors assigned to the groups (Faye et al., 2004; Cartier et al., 2006, Abdi et al., 2007). Pearson's correlations coefficients and projections onto the DISTATIS space were executed using Microsoft Excel (Microsoft Corporation, wwwmicrosoft.com) and XLSTAT 2017 (www.xlstat.com, Addinsoft).

5.2.4.4 Identifying drivers of quality

During step three of the data analysis the drivers of quality were identified (Fig. 1). For this purpose, correspondence analysis (CA) was performed on the contingency table that contained the descriptors used to describe the groups that were created during the free sorting step. CA was conducted to obtain a descriptor-based sensory space which represents the relationship between the sensory characteristics of the samples. This space was used to correlate to the sensory characteristics to the perceived quality.

Standardised deviates (also called Pearson residuals) were calculated for each descriptor from the formula provided below. These deviates indicate the magnitude of deviation from independence between wines and descriptors. Negative deviates indicate less occurrence of a descriptor with a wine as would be expected under independence, and positive deviates more occurrence of a descriptor with a wine as expected under independence. The mathematical equation used to calculate standardised CA deviates was:

$$Standaridised\ deviates = \frac{observed\ frequency - expected\ frequency}{\sqrt{expected\ frequency}}$$

Pearson's correlation coefficients between the standardised CA deviates of attributes and quality scores could, therefore, be used to indicate the sensory drivers of quality. A positive correlation between a descriptor's standardised deviates and the quality score would indicate that the descriptor tend to co-occur more with better quality wines and less with worse quality wines. Attributes corresponding to Pearson's correlation coefficients larger than 0.6 were considered as drivers of quality. Statistica 13 (Statsoft Inc., www.statsoft.com) was used to perform CA, calculate standardised CA deviates and Pearson's correlation coefficients between the standardised CA deviates and quality scores.

5.3 Results and discussion

5.3.1 Quality differences between premium Sauvignon Blanc wines

Significant differences, at a 95% confidence level, between the wines in terms of the "nose", the "palate", the appearance and overall quality were obtained from the one-way ANOVA results.

Among the set of wines used in this study, wines A and E were scored the highest and wine D the lowest for overall quality. The same trend was seen for the quality related to the “nose”, although no significant difference between wines D and H was found. The quality differences related to the “palate” also showed the same trend, with the difference that wines C, G, and D did not differ significantly from each other and wine H was rated higher than wines C and D for the quality perceived on the “palate”. Only one wine, wine D, was scored significantly lower than the other wines for appearance (Fig. 2). Wine colour was, therefore, not considered as an important contributor to the overall quality differences between wines. Similar findings by Valentin et al. (2016) showed that wine colour was not a major contributor toward Pinot Noir wine quality as measured on a 10-point scale. In the present research, more wines differed significantly from each other in terms of quality related to the “the nose”, than for quality related to the “palate” (Fig. 2). Interestingly, Sáenz-Navajas et al. (2016) observed that olfactory quality perception, which relates to the quality of the “the nose” of the wine, was found to be the most important aspect of overall wine quality. To summarise, Wines A and E scored the highest for overall quality; wines F and B did not score significantly different from each other, as was the case for wines B and G, while wine F scored higher than wine G (Fig. 2D).

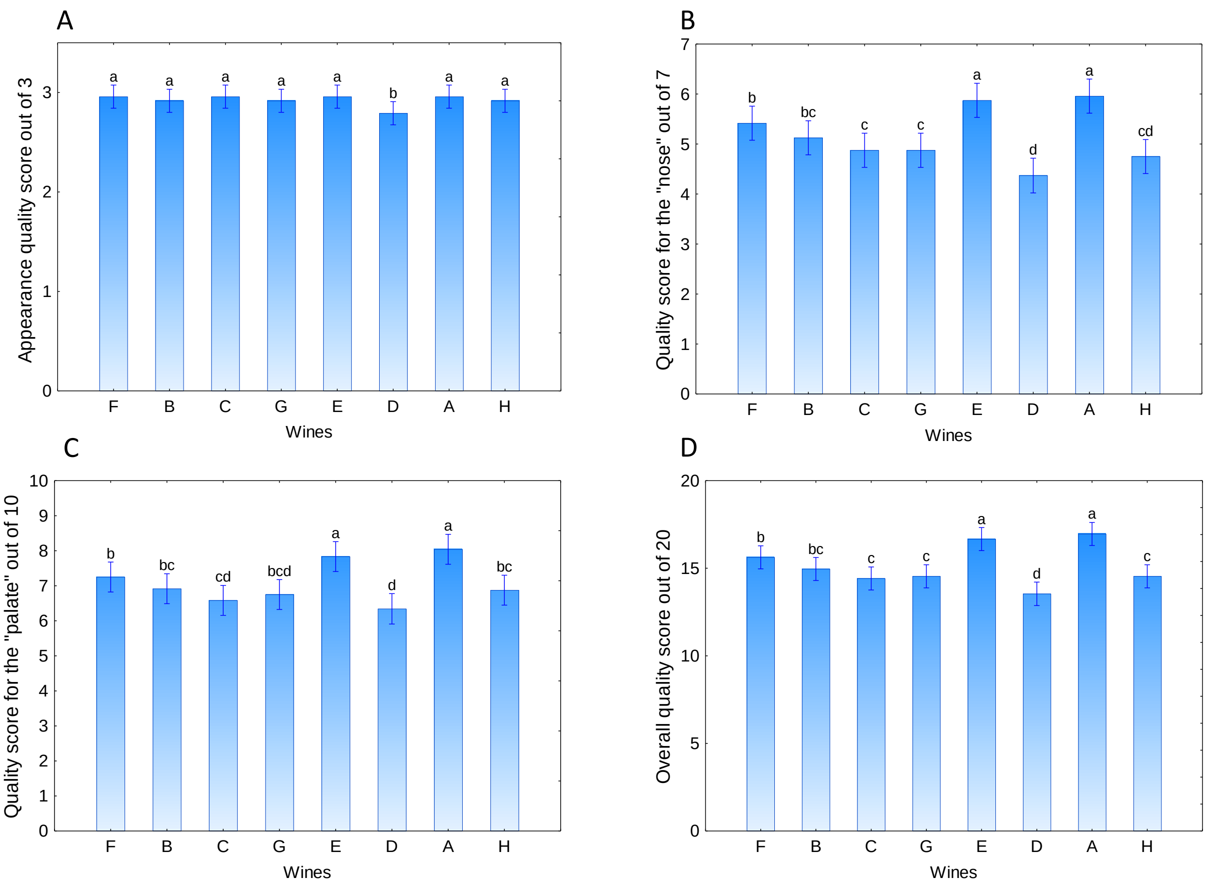


Fig. 2. Average quality scores for the appearance (A), odour indicated as the “nose” (B), the “palate” (C) and overall quality (D) for 8 premium quality Sauvignon Blanc wines analysed by one-way mixed model ANOVA and Fisher’s LSD post-hoc pairwise comparison test. The letters a - d indicate significant quality differences, $p < 0.05$, between the different wines, A – H.

5.3.2 Panel consensus and differences between individual judges' sorting data

The consensus among the individual judges was investigated by inspection of the DISTATIS plot that showed the judge configuration (Fig. 3A). It is evident that the panel consensus was good. This is in line with the findings from the study by Torri et al. (2013) who inferred that the good consensus observed between experts' description of wines could be ascribed to their use of a common language that stemmed from experience gained from evaluating good quality wines. In the present study, judge 11 could be considered as an outlier. This was confirmed by performing a STATIS Principal Component Analysis (PCA) shown in Fig. 3B. It is clear that this judge's calculated weight (Fig. 3B) was lower than that of the other judges. The data of Judge 11 were not removed from the final data analyses, since analyses performed with and without this judge's data provided the same results (data analysis excluding judge 11 is not shown).

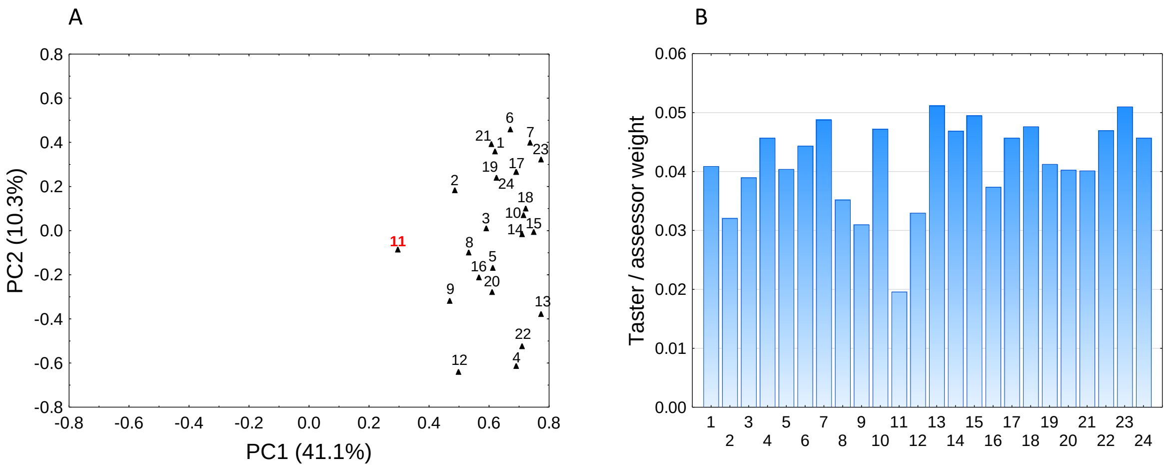


Fig. 3. DISTATIS judges plot (A) and STATIS analysis (B) performed on the individual judge distance matrices.

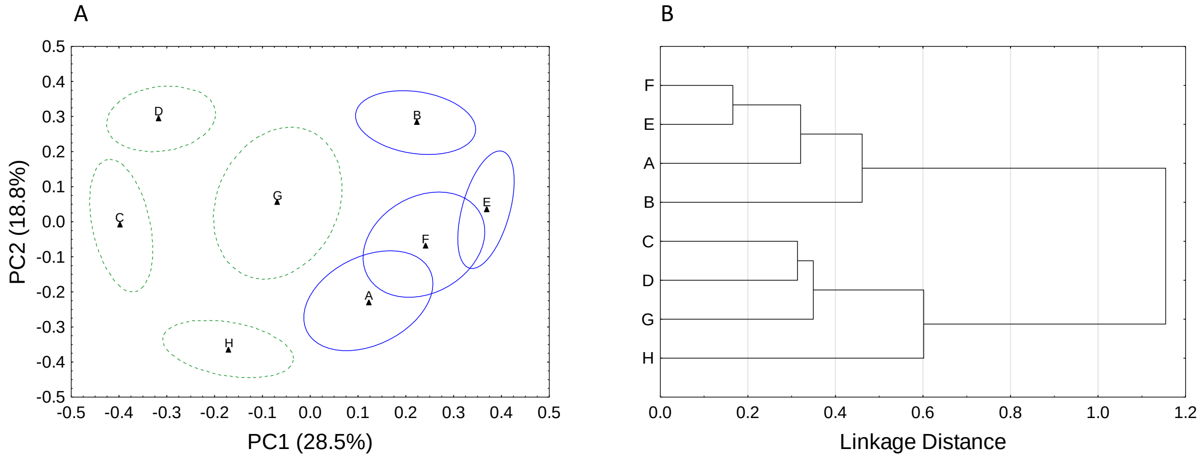


Fig. 4. DISTATIS compromise map (A) and AHC dendrogram constructed from AHC performed on the coordinates of the DISTATIS compromise map (B).

5.3.3 Differences and similarities between the wines investigating sorting groups

The sorting results visualised by agglomerative hierarchical cluster analysis (AHC) of the DISTATIS graph, consisting of the first two dimensions of the DISTATIS analysis, showed clear differences between some wines as well as similarities between others. It could be seen from the DISTATIS graph (Fig. 4A) and AHC dendrogram (Fig. 4B) that the wines could be divided into two groups along PC1 and PC2 with wines C, D, G and H forming one group and wines B, E, F and A another group. Furthermore, wines A and E, with the highest overall quality scores show overlapping confidence ellipses, also with wine F which has the third highest overall quality score, even if not different from wines B. Therefore, these wines were more similar to each other than wines B, C, D, G and H with lower overall quality scores. Quality seems to be either dependant on the sensory profiles of the wines or industry professionals intuitively sorted according to quality even though judges were not asked to sort wines according to quality. In fact, the judges did not know that they would score the quality of the wines until after the sorting task was completed. This is in-line with literature since Sáenz-Navajas et al. (2016) also reported that wine industry experts sorted according to quality. Therefore, the relationship between quality and the sensory attributes used to describe the sorting groups was investigated.

5.3.4 Aroma and flavour profile differences driving Sauvignon Blanc quality

The Pearson's correlation coefficients calculated between the CA deviates and quality scores (Table 2) indicated that the following attributes could be interpreted as drivers of quality: "passion fruit" associated with the quality of "the nose"; "asparagus" associated with taste quality and "peas" and "green pepper"' associated with the quality of "the nose", taste and overall quality. The correlation coefficients were > 0.7 and p-values < 0.05. Correlation coefficients for "tropical", "apple" and "grass" were less than -0.8 (Table 2), with a significant p-value, $p < 0.05$. This indicated that these attributes were less frequently associated with quality wines compared to when these attributes would be chosen randomly. In other words, these attributes were not associated with high quality wines.

In order to visualise all the information obtained and summarise the data analysis conducted the Pearson's correlation coefficients between the DISTATIS coordinates, PC1 and PC2, and the attribute citation were projected onto the DISTATIS compromise map (Fig. 5). To visualise the overall quality scores in an intuitive way, the size of the data markers representing the wines were adjusted. Large data markers represent high overall quality scores and smaller data markers lower scores. Attributes identified as drivers of quality were coloured blue and attributes negatively associated with quality were coloured red.

From the visualisation of the differences in the sensory characteristics, specifically odour, aroma and flavour (Fig. 5), it can be concluded that wines E and F were perceived similarly and attributes such as "peas", "green beans", "asparagus", "green", passion fruit" and "grapefruit" were cited frequently for these samples. Wine A was perceived as having a general green notes with similar attributes cited frequently as for wine E and F. The differences between wine A and, E and F were due to the fact that "green pepper" was used by all the assessors to describe wine A and "passion fruit" was cited frequently for wine F and E.

Wines that were generally perceived as having a green character with attributes such as "green pepper", "peas", "green beans" and "asparagus" were scored high for quality. These wines, specifically wine A, F and E were, in addition to the green notes mentioned above, positively correlated with attributes such as "passion fruit", "grapefruit", "cat pee" and "tomato leaf" indicating that many of the judges cited these aromas for those wines. More specifically "cat pee" was cited frequently for wine A, "tomato leaf" for wines A and F and "passion fruit" for wine F and E and "grapefruit" for wine A, F and E. It is interesting to note that these wines were negatively correlated with notes such as "pineapple" for wine B and E, "fruit salad" for wine A and F and "tropical" for wine B and E.

It appeared if green notes in general are associated with high quality South African Sauvignon Blanc wines, with "grass" as the exception for this sample set. The fact that "grass" was not associated with high-quality could be due to the overpowering effect of the "green pepper" notes in the high-quality wines masking the "grass" in those wines and might not have been perceived by the judges. Another possibility could be that judges choose to use the term "green" by considering that the "grass" note was included in the more general "green" description. The possibility is not ruled out that some bias may have been introduced in the judges' evaluations due to the specific sensory methodology used, or during the attribute combination step in the descriptor clean-up. It could be considered that different criteria for combining attributes could be used; for example, instead of combining attributes based on the citation frequencies attributes belonging to the same attribute family could be combined. This is the first report where "grass" was negatively correlated with notes such as "green pepper" and "asparagus". Generally grass notes for New World Sauvignon Blanc wines such as Australian and New Zealand wines are associated with high quality (Parr et al., 2006). "Grass" is frequently combined with "green pepper" and other green notes. In order to draw conclusions on the association of the "grass" note with the quality of Sauvignon Blanc wines further investigation, where a larger number of South African Sauvignon Blanc wines are included and different sensory methods for profiling are used, is needed. Wines with general tropical fruit characteristics, "fruit salad" and "apple" were scored lower for quality, while wines with "passion fruit" notes were scored higher.

Figure 5 provides researchers and industry professionals with a sensory map or graph to identify the drivers of wine quality by visualising sensory attributes and quality, with a single graph. This plot can be used as a final visualisation tool, but the necessary quality control steps, to ensure that the wines differ significantly in quality should still be conducted on the data. The following steps are recommended as quality control steps: (i) evaluate judge consensus by interpreting the confidence ellipses on the DISTATIS graph; (ii) conduct ANOVA to ensure that quality differences between wines are significant; and, (iii) identify drivers of quality by computing correlation coefficients instead of only inspecting the final graph containing all the information. Clustering can be used to identify groups of samples, but inspection of the DISTATIS graph only to identify groups, might in many cases be sufficient.

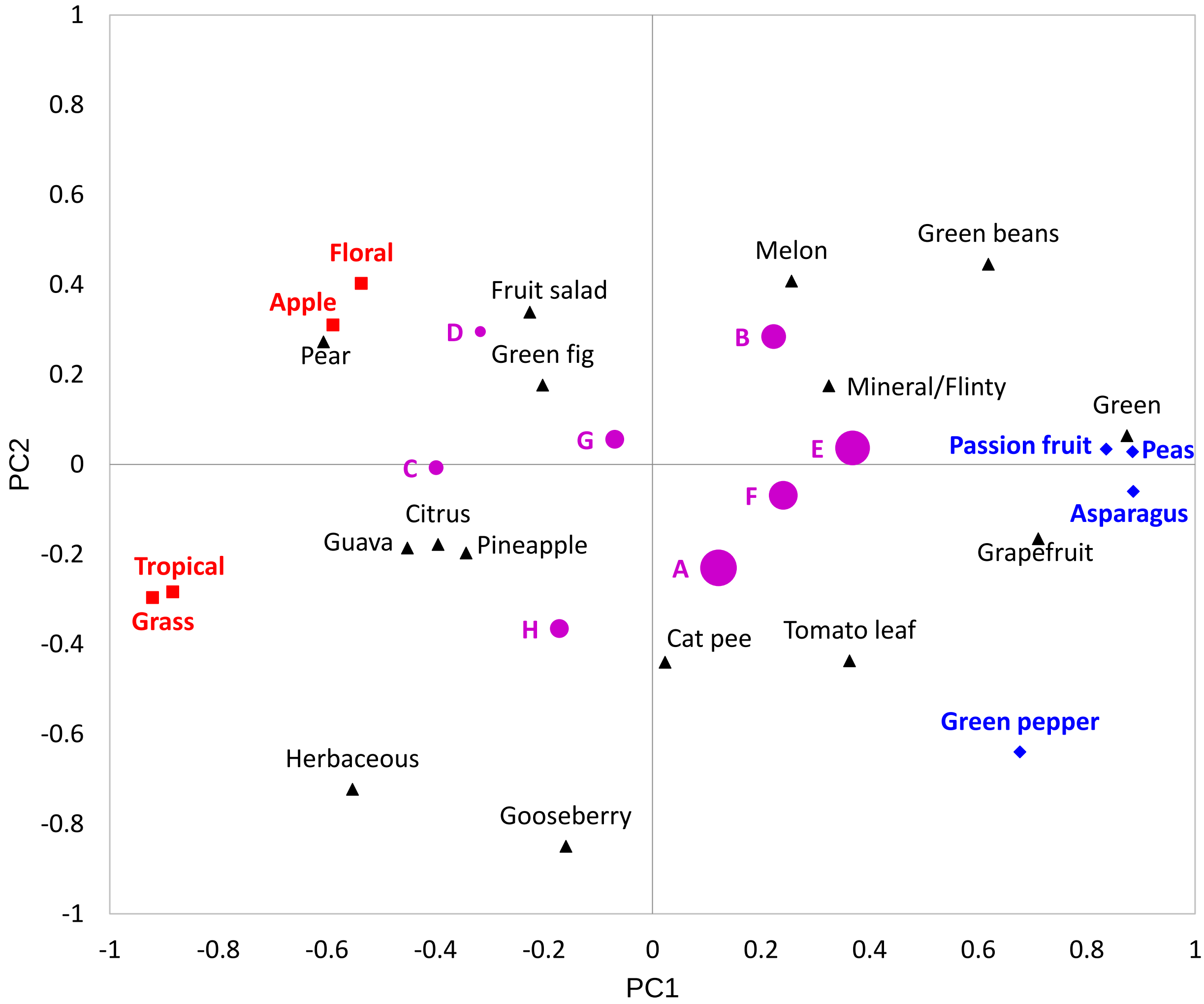


Fig. 5. Multivariate sensory map used to illustrate drivers of quality. The map includes a number of elements. 1) A two-dimensional DISTATIS compromise map with products using ● in purple as markers. These markers are sized according to the overall quality scores. 2) The projection of the Pearson's correlation coefficients between sensory attributes used and the DISTATIS product coordinates indicated with ▲, ♦ and ■ as markers. Attributes and ♦ markers in blue indicate attributes that are positive drivers of quality being positively correlated with high quality wines. Attributes and ■ markers in red represent attributes negatively correlated with quality and can be noted as negative drivers of quality.

TABLE 2
Pearson's correlations coefficients and *p*-values calculated between correspondence analysis (CA) deviates representing sensory attributes and quality scores.

	Correlation coefficients and correlation coefficient p-values					
	Quality of "the nose"		Quality of "the palate"		"Overall quality"	
Sensory Attributes	Correlation coefficient	*p*-value	Correlation coefficient	*p*-value	Correlation coefficient	*p*-value
Guava	-0.22	0.59	-0.19	0.66	-0.21	0.61
Passion fruit[a]	0.70	0.05*	0.68	0.06	0.68	0.06
Grapefruit	0.45	0.26	0.51	0.19	0.48	0.22
Citrus	-0.21	0.62	-0.21	0.61	-0.22	0.61
Asparagus[a]	0.69	0.06	0.70	0.05*	0.69	0.06
Peas[a]	0.79	0.02*	0.77	0.02*	0.78	0.02*
Green beans	0.19	0.66	0.10	0.81	0.14	0.74
Green pepper[a]	0.77	0.03*	0.87	<0.01**	0.82	0.01**
Mineral / Flinty	-0.26	0.54	-0.27	0.52	-0.27	0.52
Tropical[a]	-0.82	0.01**	-0.78	0.02*	-0.80	0.02*
Melon	-0.18	0.67	-0.27	0.51	-0.23	0.58
Green fig	-0.42	0.30	-0.51	0.20	-0.47	0.24
Floral[a]	-0.77	0.03	-0.84	0.01	-0.80	0.02
Pear	-0.58	0.13	-0.67	0.07	-0.63	0.01
Apple[a]	-0.80	0.02*	-0.82	0.01**	-0.82	0.01**
Tomato leaf	0.62	0.1	0.62	0.09	0.62	0.1
Green	0.60	0.12	0.51	0.20	0.55	0.16
Herbaceous	-0.50	0.21	-0.38	0.35	-0.44	0.28
Gooseberry	-0.13	0.76	0.00	0.99	-0.06	0.90
Grass[a]	-0.86	<0.01**	-0.84	<0.01**	-0.85	<0.01**
Fruit salad	-0.52	0.19	-0.59	0.12	-0.54	0.16
Cat pee	0.13	0.76	0.08	0.84	0.11	0.80
Pineapple	0.22	0.60	0.09	0.84	0.17	0.69

[a]Sensory attributes that can be interpreted as drivers of quality with Pearson's correlation coefficients between CA deviates corresponding to sensory attributes and quality scores ≥ 0.7. Attributes in blue are correlated with high quality and hence positive drivers of quality. Attributes in red are negatively correlated with quality and therefore, negative drivers of quality.

Pearson's correlation coefficient p-values: * ≤ 0.05, ** ≤ 0.01.

5.3.5 Relationship between Sauvignon Blanc "palate" quality and taste, mouthfeel and chemical analysis

Taste and mouthfeel attributes, anticipated to be drivers of the "palate" quality score, were only cited by a few judges. The frequencies at which these attributes were cited were not high enough to include these attributes in the statistical data analysis, since less than 20% of the panel cited similar taste and mouthfeel attributes. Taking the technical information and chemical analysis of the wines (Table 1) into consideration, it could be seen through inspection that the quality of the wines was not dependant on, or correlated with the chemical values of pH, titratable acidity (TA), residual sugar (RS) or alcohol.

Wines A and E were perceived as the highest quality wines, with no significant difference between the two. Wine A had an alcohol content of 12.5% v/v and wine E 14.0% v/v, spanning the minimum and maximum range for the set of samples. These two wines were also different with regards to TA content with wine A having 8.5 g/L, the highest of all the wines, and wine E 6.7 g/L, the second lowest of all the wines. The wine with the lowest residual sugar content (1.3 g/L), wine G, and the wine with the highest (3.3 g/L), wine B, did not differ significantly with regards to quality either.

From these observations it seemed as if retronasal perception related to the flavour of the wines played the most important role in the perception of the "palate" quality rather than taste perception such as sweet and sour and the perception of alcohol burn. It was noted by Sáenz-Navajas et al. (2016) that the concepts of wine taste and mouthfeel quality are build only in context with wine odour and aroma quality. This means that the quality of a wine as perceived on the palate is mostly based on flavour and balance of the wine rather than the perception of the individual basic tastes (sweet, sour and bitter) and moutfeel sensations (astringency and alcohol burn).

5.4 Conclusions

Both the quality assessment and description of the sensory characteristics at smaller cellars are conducted as a tasting accompanied by a discussion of opinions. In cases where formal sensory analysis is used in combination with quality scoring, methods based on quantitative descriptive analysis (Stone & Sidel, 2004; Lawless & Heymann, 2010) such as descriptive analysis (DA), that involves extensive panel training is used.

The procedure presented in this article can be useful for the South African wine industry to obtain fast, objective scientific sensory data for relating sensory characteristics to quality. This procedure can be used in-house by cellars to relate quality parameters to intrinsic sensory properties like odour, aroma, flavour, taste and mouthfeel characteristics. However, taste and mouthfeel attributes were not identified as drivers of quality during this study. Colour was not a major role player in overall quality. It might, therefore, be sufficient to only score overall quality instead of the appearance, the "nose" and the "palate" quality separately. When quality as perceived specifically on the "palate" is required a different strategy should be investigated.

The most time-consuming part of this method was the reduction of the number of attributes through combination of similar attributes. Further research is needed to identify and address possible biases introduced with this step. We propose the combination of attributes belonging to the same descriptor family as an alternative when specific detail is not required.

The method could be used in a similar way to preference mapping (McFie & Thomson, 1988; Van Kleef et al., 2006; Lawless & Heymann, 2010) constructing sensory maps where liking data, obtained from consumers, are combined with profiling data. As opposed to classical preference mapping where quantitative descriptive analysis (QDA™) data are used for profiling, sorting data could be used as profiling step to determine drivers of liking. It has been shown that both consumers and wine industry professionals have the ability to profile wines using sorting (Bester 2011).

However, it should be kept in mind that the quality perceptions of wine by industry professionals do not necessarily correlate with consumer preference, liking and acceptance. It is, therefore, recommended that complementary consumer studies are used for marketing studies and testing consumer perception. The procedure suggested here to determine quality drivers should be used during product development and to acquire analytical sensory data on quality and profiling of wines. This type of data will correspond better to wine competition data than consumer perceptions since wine experts are used as judges during competition tastings.

Another application could be to relate both quality data, from professionals and liking data, from consumers to intrinsic sensory properties using a similar procedure to the one presented in this article. The results obtained could be useful for benchmarking, product development and marketing where it is often crucial to relate quality perception of a product to the intrinsic, sensory characteristics like odour, aroma and flavour perception.

This procedure is particularly suited for industry applications for a number of reasons. Tasting groups consisting of industry professionals are common. Sorting is a relatively easy task that does not require sensory training and can be performed by experts as well as novices. A single graph can provide information regarding the quality and sensory characteristics of the products. It is a fast, low cost, objective scientific method and the results are easy to interpret providing key information useful for product development and marketing.

This method could be equally useful in research, as a rapid sensory tool, where the differences between wines in terms of quality or liking are needed to supplement research in oenology and viticulture.

5.5 References

Abdi, H., Valentin, D., Chollet, S., & Chrea, C. (2007). Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. *Food Quality and Preference, 18*, 627–640.

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.

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

Bester, I., (2011). Classifying South African Chenin blanc wine styles. MSc thesis, Stellenbosch University. http://hdl.handle.net/10019.1/18048

Botonaki, A., & Tsakiridou, E. (2004). Consumer response evaluation of a Greek quality wine. *Acta Agricola Scandinavia, Section C, Food Economics, 1*, 91–98.

Bredahl, L., Grunert, K. G., & Fertin, C. (1998). Relating consumer perceptions of pork quality to physical product characteristics. *Food Quality and Preference, 9,* 273–281.

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.

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.

Charters, S., & Pettigrew, S. (2007). The dimensions of wine quality. *Food Quality and Preference. 18*, 997–1007.

Chollet, S., & Valentin, D. (2000). Le degré d'expertise a-t-il une influence sur la perception olfactive? Quelques éléments de response dans le domaine du vin. *L'Année Psychologique, 100*, 11-36.

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.

Faye, P., Brémaud, D., Durand-Daubin, D., Courcoux, P., Giboreau, A., & Nicod, A. (2004). Perceptive free sorting and verbalization tasks with naive subjects: an alternative to descriptive mappings. *Food Quality and Preference, 15,* 781–791.

Honoré-chedozeau, C., Lelièvre-desmas, M., Ballester, J., Chollet, S., & Valentin, D. (2017). Knowledge representation among assessors through free hierarchical sorting and a semi-directed interview : Exploring Beaujolais wines, *Food Quality and Preference, 57*, 17–31.

Lattey, K. A., Bramley, B. R., & Francis, I. L. (2009). Consumer acceptability, sensory properties and expert quality judgements of Australian Cabernet Sauvignon and Shiraz wines. *Australian Journal of Grape and Wine Science. 16,* 189–202.

Lawless, H.T., & Heymann, H. (2010). *Sensory evaluation of food. Principles and Practices* (2nd ed.). Springer, New York, USA.

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.

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.

Lelièvre-Desmas, M., Valentin, D., & Chollet, S. (2017). Pivot profile method: What is the influence of the pivot and product space? *Food Quality and Preference, 61*, 6–14.

McCloskey, L.P., Sylvan, M., & Arrhenius, S.P. (1996). Descriptive analysis for wine quality experts determining appellations by Chardonnay wine aroma. *Journal of Sensory Studies, 11,* 49-67.

McFie H. J. H., & Thomson D. M. H. (1988). Preference mapping and multidimensional scaling. In: J. R. Piggott (ed.), *Sensory Analysis of Foods* (pp. 381-409). New York: Elsevier Applied Science.

Neisser, U. (1987). *Concepts and conceptual developments: ecological factors and intellectual factors in catergorization*. Cambridge: Cambridge University Press.

Pagès, J. (2003). Recueil direct de distances sensorielles: application à l'évaluation de dix vins blancs de Val de Loire. *Science des Aliments, 23,* 679-888.

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,* 642-664.

Parr, W. V, Green, J. A., & White, K. G. (2006). Wine judging, context and New Zealand Sauvignon Blanc Évaluation de vins, contexte et Sauvignon Blanc Néo-Zélandais. *Revue européenne de psychologie appliquée*, *56*, 231–238.

Parr, W. V., Green, J. A., White, K. G., & Sherlock, R. R. (2007). The distinctive flavour of New Zealand Sauvignon blanc: Sensory characterisation by wine professionals. *Food Quality and Preference, 18*, 849–861.

Parr W. V., Ballester J., Peyron D., Grose C., & Valentin D. (2015). Perceived minerality in Sauvignon wines: Influence of culture and perception mode. *Food Quality and Preference, 41*, 121-132.

Perrin, L., & Pagès, J. (2009). Construction of a product space for from the Ultra-flash profiling method: application to ten red wines from the Loire Valley. *Journal of Sensory Studies, 24*, 372-395.

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

Piombino, P., Nicklaus, S., Fur, Y. Le, Moio, L., & Le Quéré, J. (2004). Selection of Products Presenting Given Flavor Characteristics: An Application to Wine. *American Journal of Enology and Viticulture, 55*, 27–34.

Rosch, E. (1973). Natural categories. *Cognitive Psychology, 7*, 328– 350.

Sáenz-Navajas, M.-P., Avizcuri, J.-M., Ballester, J., Fernandez-Zurbano, P., Ferreira, V., Peyron, D., & Valentin, D. (2015). Sensory-active compounds influencing wine experts' and consumers' perception of red wine intrinsic quality. *LWT – Food Science and Technology, 60,* 400–411.

Sáenz-Navajas, M., Miguel, J., Federico, J., Ferreira, V., Fernández-Zurbano, P., & Valentin, D. (2016). Understanding quality judgements of red wines by experts: Effect of evaluation condition. *Food Quality and Preference, 48,* 216–227.

Sáenz-Navajas, M.-P., Ballester, J., Pecher, C., Peyron, D., & Valentin, D. (2013). Sensory drivers of intrinsic quality of red wines: Effect of cultures and level of expertise. *Food Research International, 54*, 1506–1518.

Stone, H., & Sidel, J. (2004). *Sensory Evaluation Practices* (3rd ed.). San Diego: Academic Press Inc.

Thuillier, B., Valentin, D., Marchal, R., & Dacremont, C. (2015). Pivot© profile: A new descriptive method based on free description. *Food Quality and Preference, 42*, 66–77.

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.

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*, 1563–1578.

Valentin, D., Parr, W. V, Peyron, D., Grose, C. & Ballester, J. (2016). Colour as a driver of Pinot Noir wine quality judgments: An investigation involving French and New Zealand wine professionals. *Food Quality and Preference, 48*, 251–261.

Van Kleef, E, Van Trijp, H. C., & Luning P. (2006). Internal versus external preference analysis: An exploratory study on end-user evaluation. *Food Quality and Preference, 17*, 387–399.

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.

Verdú Jover, A. J., Lloréns Montes, F. J., & Fuentes Fuentes, M. del M. (2004). Measuring perceptions of quality in food products: the case of red wine. *Food Quality and Preference, 15*, 453–469.

Ziethalm, V. A., Berry, L. L., & Parasuraman, A. (1988). Communication and control processes in the delivery of service quality. *Journal of Marketing, 52*, 35–48.

Chapter 6

General discussion and conclusions

6. General discussion and conclusions

6.1 General discussion

The development, optimisation and comparison of rapid sensory analysis methods received a substantial amount of attention lately in the field of sensory science research (Valentin et al., 2012; Varela & Ares, 2012). These methods were proposed as alternatives for descriptive analysis (DA, Stone & Sidel., 1974), to address some of the disadvantages of DA such as time-consuming training. Although these methods have been introduced and some validated, a need for optimisation, testing of certain aspects of the methodology as well as testing the suitability for specific applications and product matrices emerged mainly due to industry demand. The demand for tailor-made sensory analysis methods to analyse wine fast and in a cost-effective manner has become increasingly important. Rapid sensory methods have been used to profile wine, but no study to date compared more than three of these methods to each other and to a trusted reference method such as DA.

The main objective of this research project was to identify rapid sensory methods suitable for wine profiling using wine industry professionals and trained panels. To achieve this a number of research studies were conducted.

During the first study (described in Chapter 3) different rapid sensory methods were compared to DA to identify the most suitable methods for profiling of wine using industry professionals as judges. The rapid methods compared were sorting (Lawless et al., 1995; Chollet et al., 2011), Napping (Risvik et al.,1994; Pagés, 2003), check-all-that-apply (CATA, Adams et al., 2007) and rate-all-that-apply (RATA, Ares et al., 2014). In order to standardise the procedure for all the methods and simplify data analysis a previously determined list, set-up by consulting industry aroma wheels and previous studies were used as the descriptive step. The following criteria were taken into account to assess the suitability of the rapid methods for wine profiling:

a) Multivariate sensory map quality obtained, judged by the explained variance, overlap of confidence ellipses and the distance between samples on the map.
b) Similarity between the DA and rapid method sensory maps measured by RV coefficients (Robert & Escofier, 1976).
c) The difficulty experienced by the panel to execute the task, rated by sensory judges on a 9-point easiness scale after the sensory evaluation session.
d) Sensory evaluation time needed to obtain the results.

All rapid methods provided good quality results. The first two factors of the multivariate sensory maps explained more than 50% of the variance within the data set, where single-block analyses were conducted on panel averages, and 40%, where multi-block analyses on individual data were performed. DA produced one-dimensional results with 91% explained along principal component (PC) 1 mainly characterising wines in terms of woody notes such as “oaky”, “vanilla”, “caramel” and “toasted bread” or fruity notes e.g. “peach”, “passion fruit” and “pineapple”. A similar observation can be made for RATA even though not as extreme as for DA. The second dimension of the sensory maps obtained from CATA, sorting and Napping indicated differences between samples that could be observed along the second dimension originating from e.g. green, sweet aromatic and mineral notes that could not be observed on the DA and RATA maps.

The confidence ellipses around the samples on the sorting and CATA maps showed the least overlap and the Napping map the most. It is clear that sorting highlights differences and similarities between samples since similar samples’ confidence ellipses overlap almost completely, where CATA map confidence ellipses show a continuum and samples are spread further apart. This can be attributed to the inherent nature of the sensory method. When sorting is conducted, judges are instructed to group the samples and describe the group as a whole, where when CATA is performed each sample is described individually.

By inspection, all the maps look similar which is statistically confirmed with RV coefficients (Robert & Escoufier, 1976) calculated between the multivariate map configurations ranging from 0.69 to 0.83. It is interesting to note the good similarity between the DA and RATA sensory spaces indicated by an RV coefficient of 0.82. This was unexpected since an unstructured line scale was used without calibration of the panel or training on how to use the scale.

From these observations, it is clear that a similar broad descriptions of the sensory space could be obtained using all of the methods, but sorting highlighted similarities and differences where CATA provided more detailed profiling information.

Napping was rated by sensory judges as the most difficult rapid method to perform, it was the only rapid method rated as more difficult to perform than DA. In addition it took the longest time to complete the task when compared to the other rapid methods Therefore, a “practice session” might be useful to familiarise the judges with the task prior to performing the Napping experiment itself. CATA attributes was rated as the easiest method to perform followed by Sorting as the second easiest. From the first study it was clear that CATA attributes and sorting are suitable methods for rapid profiling of wine even without any familiarisation with the technique.

CATA attributes are particularly useful in the wine industry since sensory lexicon in the form of aroma wheels and mouth feel wheels are available and frequently used by both wine industry professional panels and as ballot training (Lawless & Heymann, 2010) material for trained panels. Since monadic sample presentation is used during sensory evaluation when using these types of methods judges cannot relate the samples within a sample set to each other directly. It is, therefore, difficult to describe small differences between different samples in a sample set. It has been proposed by Thuillier et al. (2015) and Lelièvre-Desmas et al. (2017)

that pivot profile (PP) is particularly suited for discriminating between similar products, in other words, products in a sample set where the within-set variation is small. However, one of the main considerations and difficulties performing reference-based sensory methods is choosing the reference sample.

In the second study, PP was validated for wine sensory evaluation using three different sets with respectively low, medium and large within-set variability between samples by:

(1) Testing the stability of PP when the pivot sample is changed.
(2) Comparing results obtained from PP and FC when analysed with correspondence analysis (CA). Where FC is a CATA variant that involves training of judges using a reference standard for each term.

The specific criteria used to determine if the PP configuration is stable and PP results compared well to FC results were:

(1) The variance explained, overlap of confidence ellipses and distances between samples as a measures of the quality of the sensory map.
(2) Similarity between the sensory maps measured by RV coefficients.
(3) The difficulty of the task as experienced by the panel, rated on a 9-point easiness scale.
(4) Panel performance where consensus was measured with Cohen's kappa coefficients and repeatability with the reproducibility index (Ri) value.

It was clear from the RV coefficients calculated, between CA plots constructed for both the PP and FC attributes' data, that the differences between the different PP data sets, when changing the pivot, were larger than the differences between the PP and FC data.

Good results were obtained with both methods when a sample set with medium complexity was analysed. It was surprising to note that the low complexity sample sets with small within-set variation between samples were better profiled using FC due to larger confidence ellipses around samples on the PP CA plots. This was the first time to the knowledge of the authors that confidence ellipses were constructed for PP data. Lelièvre-Desmas et al. (2017) concluded that PP is more suited for sample sets with small within-set variability, but confidence ellipses were not calculated during that study and the study was conducted on beer and not wine. The sample sets evaluated during this study was selected to include extreme cases, the differences between samples in the set with the small within-set variability was probably smaller than the differences between the sample in the set tested by Lelièvre-Desmas et al. (2017). The results obtained in this study for the sample set with the medium within-product variability corresponded well to the results found by Lelièvre-Desmas et al. (2017) for small within-set variability. It is, therefore, important to have some knowledge about the sample set complexity if PP is considered. It this type of information is not available FC might be a better choice.

The objective of the last study, included in this dissertation, was to apply one of the rapid sensory methods identified in the first or second study as sensory profiling method to determine drivers of quality as perceived by industry professionals.

Since CATA and sorting with a verbalisation step were identified as the easiest rapid methods to perform and the results corresponded well with the results obtained during DA one of these methods was used. Although sorting was used, the workflow proposed in this study can be applied when CATA is used as descriptive step by performing multiple factor analysis (MFA) or CA instead of DISTATIS to produce a sensory map.

The specific aim of this experiment, was to combined sorting with quality scoring to obtain a sensory map, similar to a preference map, that represents: (1) analytical profiling information, (2) sensory attributes, and (3) quality scores, out of 20. Instead of mapping preference, the quality perception of industry professionals was mapped. This procedure was tested in order to provide industry professionals with a procedure to use during the production process to relate quality to the sensory profiles of products and ultimately identify drivers of quality using in-house panels.

DISTATIS (Abdi et al., 2007) was applied to the sorting data to produce a multivariate map and visualise the consensus between the sensory judges. The sensory attributes generated were projected onto the DISTATIS compromise map using Pearson’s correlation coefficients. Quality scores were analysed by means of one-way analysis of variance (ANOVA) and indicated on the same plot using the size of the marker, higher quality wines were represented by larger data markers and low quality wines by smaller markers. In order to evaluate this procedure the following criteria were considered:

(1) The quality of the data was assessed with a focus on the agreement between sensory judges when performing sorting.
(2) Discrimination of samples on the compromise map.
(3) The significance of the differences between quality scores.

A good agreement among the sensory judges was observed, meaning that they grouped the wines similarly, only one of the 24 judges could be regarded as an outlier. Discrimination of wines on the compromise map was good showing that differences between wines could be visualised. Significant differences between the quality scores of the wines were observed when one-way ANOVA was performed on the quality scores. It could, therefore, be concluded that the proposed procedure can successfully be used to determine drivers of quality using wine industry professionals as judges. However, the statistical analysis of the data could be tedious to perform if the analysts are not familiar with the techniques. In order to overcome this drawback an “all-in-one” software application could be developed since all the statistical methods used are well-known and validated.

6.2 Summary of research findings

The main research outcomes and finding of this project are:

- Knowledge on the suitability, limitations, pros and cons of different rapid sensory methods for wine profiling that can be used to make informed decisions to choose fit-for-purpose methods for research applications were acquired.
- Rapid sensory methods suitable for industry applications such as wine style identification were identified using Chenin Blanc as an example.
- The suitability for PP to describe wines with different within-set variability was shown.
- Guidelines on choosing a pivot sample, when PP is conducted, taking the within-set variability into account was established.
- A statistical procedure to calculate confidence ellipses for PP data to improve the visualisation of the sensory map constructed by means of CA was proposed.
- A procedure to obtain drivers of quality visually displayed on a sensory map similar to a preference map was proposed.

6.3 Future recommendations

Although various methods belonging to the verbal-based and similarity-based categories were studied, for wine sensory applications during this project, only one method from the reference-based category was validated. Further work is needed to investigate the suitability, advantages and limitations of reference-based sensory methods such as polarised sensory positioning (PSP, Teillet et al., 2010) and polarised projective mapping (PPM, Ares et al., 2013) for wine sensory evaluation. One of the main aspects of PPM that for example requires further investigation, highlighted by Wilson (2018) when evaluating wine, is the choice and positioning of the poles. Comparing and testing reference-based methods for specific wine industry application such as benchmarking is needed. For example, PSP and PP can be compared as benchmarking tools since benchmarking involve direct comparison of products. In addition, proposing procedures which address specific needs such as profiling of large samples sets, where data obtained during multiple sensory evaluation session have to be aggregated, could be another possible application of reference-based methods.

In addition to sensory methodology studies, the development of procedures and techniques to analyse and visualise sensory method data is needed. One of the aspects that require further attention is the measuring of panel performance including repeatability and consensus. Even though procedures and methods were proposed for methods such as Napping and check-all-that-apply (CATA), further work is needed for reference-based methods such as PP. Visualising individual differences between judges is another field of study were there is scope for development of new or adapted statistical methods, again this is currently lacking for PP data for example.

Prior to multivariate statistical analysis performed, to obtain a sensory map, data, from verbal-based methods, are coded to extract themes or directly captured as attributes. These attributes are then reduced by either: (1) linguistic and semantic synonym combination or (2) statistical analysis used to identify attributes responsible for significant differences between the products. The efficiency, advantages and limitations of these strategies have not been tested and compared. Studies in this regard could shed light on which strategy would provide the best sensory map for specific objectives depending on whether for example a broad overview or more detailed profiles are required.

This study can provide the base for various experiments where rapid methods are applied to answer specific wine-related questions such as wine style description and identification, sensory aspects of tainted wines such as smoke taint, the effect of oenological and viticultural treatments on the final wine, investigating consumer and/or expert perceptions.

6.4 Concluding remarks

From this study, it became clear that rapid sensory methods are suitable for the analysis of complex matrices. It is, however, important to select a method that is fit-for-purpose even though most methods will provide similar multivariate sensory maps. The objective of a study should be considered carefully when the sensory experiment is planned. In addition, practical and logistical constraints should be considered.

DA provided one-dimensional multivariate maps, a possible reason for this might be the restriction of the number of attributes used during DA. Therefore, it is possible that unlike other food products, rapid methods might be better adapted for wine analysis than DA in many cases since the number of attributes is not restricted to a maximum of 15 to 25. It is, therefore, easier to accurately describe the smaller differences between samples in wine, a complex matrix. The fact that frequency counts are obtained and not intensities does not seem to be a big problem as the assumption that higher frequencies represent higher intensities holds for the most applications. This can be deduced from the fact that similar sensory maps are obtained when rapid method sensory maps, constructed form frequencies, are compared to DA sensory maps, constructed from intensity data.

To conclude, this study provides guidelines for the use of rapid sensory methods that can be used for both academic and industry focussed applications and sets the stage for further development and testing within this category of sensory tools.

6.5 References

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. *Poster presented at the 7th Pangborn Sensory Science Symposium.* 16 August 2007, Minneapolis, USA. Delegate Manual, pp. 12–418.

Abdi, H., Valentin, D., Chrea, C. & Chollet, S. (2007). Analyzing assessors and products in sorting tasks: DISTATIS, theory and application. *Food Quality and Preference*, 18, 627–640.

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*(2), 510–518.

Ares, G., Bruzzone, F., Vidal, L., Cadena, R. S., Giménez, A., Pineau, B., Hunter, D. C., Paisley, A.G. & Jaeger, S. R. (2014). Evaluation of a rating-based variant of check-all-that-apply questions: Rate-all-that-apply (RATA). *Food Quality and Preference*, *36*, 87–95.

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*(2), 104–115.

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*(6), 507–520.

Lawless, H. T., & Heymann, H. (2010). Sensory evaluation of food. Principles and practices. New York, NY: Springer.

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*(2), 91–98.

Lelièvre-desmas, M., Valentin, D., & Chollet, S. (2017). Pivot profile method : What is the influence of the pivot and product space ?, *61*(May), 6–14.

Pagès, J. (2003). Recueil direct de distances sensorielles: application a` l'e´ valuation de dix vins blancs de Val de Loire. Science des Aliments, 23, 679–888.

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

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

Stone, H., Sidel, J., Oliviers, S., Woosley, A., & Singleton, R. C. (1974). Sensory evaluation by quantitative descriptive analysis. Food Technology, 24–28.

Teillet, E., Schlich, P., Urbano, C., Cordelle, S., & Guichard, E. (2010). Sensory methodologies and the taste of water. *Food Quality and Preference*, *21*(8), 967–976.

Thuillier, B., Valentin, D., Marchal, R., & Dacremont, C. (2015). Pivot© profile: A new descriptive method based on free description. *Food Quality and Preference*, *42*, 66–77.

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*(2), 893–908.

Wilson, C., Brand, J., Toit, W., & Buica, A. (2018). LWT - Food Science and Technology Polarized projective mapping as a rapid sensory analysis method applied to South African Chenin Blanc wines. *LWT - Food Science and Technology*, *92*, 140–146.