Factors Affecting the Identification of Odours and Wines

THE UNIVERSITY OF SYDNEY

Alex Russell

School of Psychology

Faculty of Science

University of Sydney

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

2014

ABSTRACT

Identifying an odour can be difficult, particularly when odours are combined. However, wine experts are allegedly able to describe the odours and flavours of wines in ways that other experts can understand. These skills are learned, not innate. The aim of this thesis was to explore factors that potentially affect the identification of odours and wines in order to teach novices to identify wine samples.

In initial experiments, odour mixtures were employed to determine which labels novices could learn to use in an identification task. Participants were able to use appropriate (veridical) labels, as well as self-generated labels that were mostly non-veridical.

Similar experiments were conducted with wine samples. Participants could not use self-generated labels, but could use grape names (e.g. Shiraz) and short descriptors (e.g. pepper). However, performance was not as high as some groups in the odour experiments. This may indicate difficulty in detecting the elements in the wines. Thus, wine samples were adulterated to enhance these elements. Participants could use appropriate labels when the wines were adulterated, but performed poorly when the wines were no longer adulterated, indicating that this training method may not be effective.

Participants in later experiments were allowed to taste the wines and were able to use grape names and descriptors to identify the wines, depending on the samples used. One final experiment involved training using conceptual information (e.g. information about wine regions), which was no more effective than sensory training.

The major results of the thesis are that novices can discriminate between wines using olfaction alone and that novices can learn to apply labels to wine with small amounts of training, but this learning may not transfer to other wines. There do not appear to be any shortcuts to becoming a wine expert, but novices can use labels to identify wines within a few hours of intense training.

Keywords: Olfaction, olfactory training, gustation, gustatory perception, wine, wine perception.

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge my supervisor, Professor Bob Boakes. I am very grateful to have had the opportunity to work with you during my fourth year and throughout this PhD. Thank you for your support and patience. It has been an honour and a priviledge.

I would also like to thank Professor Justin Harris for being my associate supervisor, along with Dr Angus Hughson (my unofficial associate supervisor), Dr Ian Johnston and Dr Alex Holcombe, who all provided very useful thoughts, advice and support during the process.

I have been very lucky to gather a great group of friends amongst my fellow postgraduate students during my time here, including Dr Lauren Mahlis, Dr Joanna Fardell, Ben Smith, Kate Inglis, Dr Dan Costa, Dr Elian Fink, Dr Kristina Fritz and many others, all of whom have been there through various parts of the thesis.

I would also like to thank the various staff members who have employed me to teach at the university, including Dr Margaret Charles, Dr Sabina Kleitman and Dr Caleb Owens. I have also had support from other staff behind the scenes with my teaching, including Professor Sally Andrews and Sadhana Raju. Thank you all.

I have been fortunate to be involved in research projects with various research teams, from which I have learned things that have contributed directly and indirectly to this thesis. I would like to thank the following for the opportunity to be involved in these collaborations: Professor Alex Blaszczynski, Professor Nerilee Hing, Dr Sally Gainsbury, Dr Helen Breen, Professor Dick Stevenson, Associate Professor Anina Rich and many others.

Through all of these collaborations, I have met some of the smartest people I know, which has been incredibly inspiring.

Special thanks to Tastemaster, Quest and Perfume Manufacturers for providing the odour samples that were used during the odour experiments. Thank you also to the Australian Wine Research Institute for providing me with an odour sample. My research was supported by grants from the Postgraduate Research Support Scheme, Postgraduate Research Grants and the Student Travel Allowance, all from the School of Psychology and Faculty of Science at the University of Sydney. I am very grateful for this support.

Finally, but most importantly, I would like to thank my family. My grandparents, my late dad Stephen, my mum Pamela for absolutely everything and my siblings Anthony and Emily for their love and support, along with the extended family of Russells, Baldrys, Brydons and Morans. My family now includes my lovely fiancée Steph (thank you so much for all of the support you've given me throughout all of this!) and my in-laws-to-be: Paul, Anna, Al (along with Andrew, Duncan and Leah), Bren and the extended family of Lees and Murphys. Thank you all for allowing me to drag you through this whole thing. It has been a long journey and I couldn't have done it without all of you standing beside me.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

All of the work contained within this thesis is original, with the exception of quotations of previous research by others, which is attributed throughout. One paper was published from this research, jointly authored by myself and my supervisor. This work is noted in the body of the thesis. None of the work contained in this thesis has been submitted for another degree.

This research was approved by the University of Sydney Human Research Ethics Committee, project number 10235.

CHAPTER ONE – HUMAN CHEMOSENSORY PERCEPTION

Introduction

The study of human chemosensory perception provides unique challenges compared to the other senses. With visual perception, it is possible to describe the features of the stimulus using a standard vocabulary that may help another person find that stimulus within a scene (Clarke, Elsner, & Rohde, 2013). The same is not the case with chemosensory perception. As described below, odours and flavours can be particularly difficult to identify, and yet odour and flavour experts are able to describe chemosensory stimuli so that another person with sufficient training can understand their description. This effect is particularly true for wine experts, who use a standardised language to describe their perception of a wine for a variety of purposes.

The aim of this first chapter is to describe the nature of chemosensory perception and the abilities of humans to detect, discriminate, recognize and identify chemosensory stimuli.

Smell, taste and flavour

The term *olfaction* refers to the sense of smell, whereby volatile, soluble gaseous molecules (odorants) enter the nasal cavity and react with the receptor cells in the olfactory epithelium. These chemical odorants can reach the olfactory epithelium either via the nostrils (*orthonasal olfaction*) or via the mouth and nasopharynx (*retronasal olfaction*) (see, for example, Engen, 1982). Information about the odorant then travels to the olfactory bulb via the olfactory nerve. The olfactory bulb has neural connections with the piriform cortex (the olfactory cortex), which, in turn, has neural connections with the entorhinal cortex (responsible for familiarity of input signals, or memory), which has neural connections with the amygdala, which is primarily responsible for the processing of fear (e.g. Fernando, Murray, & Milton, 2013). Chemosensory input is unique in that it bypasses the thalamus, which processes other sensory signals including those related to auditory, somatosensory and visual stimuli (Ward, 2013). For a more detailed overview of olfactory circuit anatomy, see Wilson and Stevenson (2006). The *odorant* is the chemical that is detected by the receptors, while an *odour* is our brain's interpretation of that odorant.

Gustation refers to the sense of taste, where molecules called *tastants* react with the receptor cells located on the tongue and, to a lesser extent, other parts of the mouth. The five tastes that are currently recognised in humans are: sweet, salty, sour, bitterness and umami (the savoury or meaty taste characteristic of substances such as monosodium glutamate).

Flavour refers to the combination of gustation and retronasal olfaction, along with other sensations, such as irritation, which is detected by the trigeminal nerve (Small & Prescott, 2005). The terms *flavour* and *taste* are often used by laypeople as synonyms (Rozin, 1982), such as in the term *winetasting*, although a winetaster is generally interested in much more than just the gustatory element of a wine. In this chapter, olfaction and flavour studies are discussed together because olfaction is a fundamental component of flavour (Small & Prescott, 2005).

Olfactory stimuli

Different odorants tend to be perceived as different odours, but this is not always the case. For example, both amyl hexanoate and ethyl butyrate have a pineapple quality (Sigma-Aldrich, 2012). Furthermore, the perceived odour of the odorant can change depending on its concentration. For example the chemical 4-mercapto 4-methylpentan-2-one smells like box tree or blackcurrant in low concentrations (0.1 ng/L alc), whereas at higher concentrations (3 ng/L alc), the same chemical smells like cat's urine (Howell et al., 2004). The key chemicals that are responsible for the particular smell of an odour are termed *character impact odorants*.

There is no predictive model of what an odorant will smell like based on any properties of the chemical. Machines cannot determine the odour or flavour of a stimulus (Piggott, Simpson, & Williams, 1998). However, some predictions can be made based on chemical structure, such as esters tending to smell fruity, although there is no way to predict which fruit any ester will smell like (Gilbert, 2008). Various classification schemes have been proposed in order to identify dimensions of olfaction (e.g. Henning, 1916), with little success.

Furthermore, most odours are the product of more than one odorant (Gilbert, 2008; Gottfried, 2010; Stevenson & Wilson, 2007a). These are referred to as "complex" stimuli, compared to monomolecular ("simple") stimuli (see Gottfried, 2010). For example, the emergent odour properties of a tomato consists of approximately 400 different volatile odorants, yet the odour of tomato can be reproduced using a combination of only 16 of these odorants, although the final product may smell somewhat artificial (Gilbert, 2008). Once perceived, interactions between these odours can occur in the olfactory bulb or antennal lobe (Wilson & Stevenson, 2006). In these multi-odorant mixtures, the combination of scores or hundreds of odorants produce a unified perceptual experience called an *odour object* (Gottfried, 2010; Stevenson & Wilson, 2007a, 2007b).

While machines such as gas chromatographs can help determine which elements are contained within an olfactory stimulus, they cannot tell us what the stimulus smells like. Instead, trained humans are used to evaluate, classify and identify odours and elements within multi-odour mixtures. The remainder of this chapter discusses human abilities in this regard.

Human olfactory abilities

Detection/sensitivity

Olfactory sensitivity tasks can refer to the ability to detect a stimulus as a whole, or the ability to detect elements within a compound. Sensitivity to whole stimuli in humans and other primates has been described as *microsmatic*, in that it is greatly reduced compared to other mammals such as dogs or rodents (Rouquier, Blancher, & Giorgi, 2000). However, performance in this regard is not necessarily as bad as might be expected. Cain (1977) showed that humans can detect as little as 5% intensity changes in n-butyl alcohol, which was much better than previously thought (for more information, see Lawless & Heymann, 2010; Stone, Bleibaum, & Thomas, 2012).

Odour sensitivity is not fixed. The repeated presentation of an odour stimulus can result in the lowering of these thresholds, increasing the ability to detect the stimulus, even in untrained (or non-expert) participants (Dalton, Doolittle, & Breslin, 2002; Doty, Huggins, Synder, & Lowry, 1981; Rabin & Cain, 1986), suggesting that training can lead to better olfactory sensitivity.

The detection of elements within a compound is also of interest. Many of the approximately 400 volatile odorants involved in the odour of tomato are present at subthreshold levels and thus cannot be detected, which is why identifying the key elements and mixing them at the correct concentrations can lead to a relatively faithful reproduction of the original odour (Gilbert, 2008).

Discrimination

The main question when testing odour discrimination is whether participants can tell the difference between two odours. Discrimination is generally tested using either the same-different procedure, where the participant smells two odours (or tastes two tastes) in succession and is asked to state whether the two stimuli are the same or different (e.g. Dwyer, Hodder, & Honey, 2004), or the triangle test, where two of three presented stimuli are the same and the participant is asked to pick the odd one out (Amerine, Pangborn, & Roessler, 1965).

While stimuli in some other modalities can be presented simultaneously (e.g. in vision, where three items could appear on the screen simultaneously), this task is necessarily sequential in olfactory and gustatory research, which thus requires some level of short-term memory (Dacremont & Valentin, 2004).

This task does not require the participant to make judgements about whether they have smelled the odours before, nor are they required to determine what the odours are, but some cognitive factors are involved in this level of olfactory processing. Hedner, Larsson, Arnold, Zucco, and Hummel (2010) conducted an experiment where 170 men and women were tested on olfactory sensitivity (detection), discrimination and identification, along with numerous cognitive tests covering factors such as executive functioning, semantic memory and episodic memory. Higher performance in executive functioning and semantic memory were related to better performance in the discrimination and identification tasks, but not the detection task, indicating that discrimination may indeed benefit from certain types of cognitive profiles.

In olfactory terms, discrimination can also refer to the discrimination of elements within a stimulus. Laing and Francis (1989) found that untrained participants can distinguish up to four elements within an odour mixture and that this appears to be the case even after extensive training (Livermore & Laing, 1996), suggesting that this may be a physiological limit. This is discussed in further detail below.

Recognition and odour memory

In odour recognition studies, participants are asked to determine whether they have smelled a particular odour on a previous occasion, with either a short or long interval between initial stimulus presentation and the recognition task to test short-term (Engen, Kuisma, & Eimas, 1973) or long-term (Engen & Ross, 1973; Walk & Johns, 1984) odour memory.

Engen and Ross (1973) tested the long-term odour memory of 37 participants by asking them to smell 48 odours and then asking them to smell pairs of odours to determine which of the pair were part of the initial 48 odours. Their participants were randomly allocated to one of four interval conditions, where the groups were tested either immediately, after one day, after one week or after one month. While the initial recognition rate was deemed to be relatively poor (around 70% correct) compared to recognition of stimuli in the other senses, there was virtually no change in recognition memory over time.

There is conflicting evidence about whether describing or verbalizing an odour helps with later recognition. Engen and Ross (1973) found that the presence of either a veridical label (e.g. wintergreen) or “association” label (e.g. Life Savers) did not increase recognition performance above the approximately 70% recognition rate in the previous experiments in the paper and concluded that the presence of labels during encoding does not help later recognition. Stevenson, Case, and Boakes (2005) found that odour memory appears to be relatively resistant to interference and forgetting when verbal interference tasks are used. However, in contrast to these results, Lyman and McDaniel (1986) found that participants who either attempted to name an odour and provide a dictionary-like definition, or who describe a life episode of which the odour reminded during encoding, performed better in recognition compared to a no-strategy control, suggesting that verbalization can aid odour memory. Finally, identification can, under certain circumstances, have negative effects on odour recognition, an effect described as verbal overshadowing (Melcher & Schooler, 1996), particularly when the odour is as complex and difficult to describe as a wine sample.

Identification

Identification tasks require participants to smell an odour (or perceive a flavour) and attempt to identify it by name. The response options can either be free response (“What is the odour?”) or multiple choice (e.g. “Which of these four words best describes the odour?”). Compared to the other senses, olfactory stimuli are difficult to identify (Bitnes, Ueland, Moller, & Martens, 2008; Hughson & Boakes, 2002a; Richardson & Zucco, 1989; Royet, Plailly, Saive, Veyrac, & Delon-Martin, 2013).

In general, performance in a free response identification task is relatively poor, such as in the study by Cain (1979), where participants initially correctly identified only half of the 80 everyday odours presented to them. This led Cain to state that the three key ingredients to successful odour identification are a) commonly encountered substances, b) a long-standing connection between an odour and its name and c) aid in recalling the name. Cain also stated that “the absence of any one ingredient impairs performance dramatically, but the presence of all three permits ready identification of scores of substances, with performance seemingly limited only by the inherent confusability of the stimuli.”

Participants may also experience the “tip of the nose” phenomenon, where they realise that they have smelled the odour before and feel that with the right cue they could identify the odour, but cannot do so (Lawless & Engen, 1977). When the name of the odour is revealed, the participant recognizes it immediately and wonders why they could not think of it in the first place.

When cues are given, such as in multiple-choice options, identification of the source of an odour is much higher. Normal performance in the University of Pennsylvania Smell Identification Test (UPSIT: Doty, Shaman, & Dann, 1984) is generally considered to be around 90% correct identification. However, in multiple-choice tests, identification of the stimulus may not be required in order to be correct. Instead, the participant may simply eliminate the foils (Sulmont-Rosse, Issanchou, & Koster, 2005).

Odour identification is even more difficult when more than one odour is present in a mixture. Laing and Francis (1989) described an experiment whereby participants were given odour mixtures containing between one and five separate odours and were asked to identify them. When only one odour was presented, it was identified correctly in 81.8% of trials. When more than one odour was present, at least 91% of the participants were able to correctly identify at least one of the elements in the mixture, but only 35.2% could identify both in the two-odour mixtures, 14.0% could identify all three in the three-odour mixtures, 3.7% could identify all four in the four-odour mixtures and 0% could identify all five odours in the five-odour mixtures, leading Laing and Francis to believe that humans cannot identify more than four elements in a mixture. Subsequent work has confirmed that four appears to be the limit to how many odours can be identified within a mixture (Jinks & Laing, 1999a, 1999b, 2001; Laing, Eddy, & Best, 1994; Laing & Glemarec, 1992; Livermore & Laing, 1998).

As Laing and Francis (1989) noted, a possible alternate explanation for poor performance tasks that require participants to identify elements in odour mixtures could also be because the odours in the mixture “blend to form a new odour with few of the characteristics of the constituent odours". While some rules have emerged in terms of the probability of odour identification and relative intensity of elements in mixtures (see Olsson, 1994), there is still no theoretical model of how odours mix to form new emergent qualities. In order to create a new mixture, odour/flavour chemists must embark upon extensive pilot testing until an appropriate combination is found. These combinations depend on the purpose of creating the mixture (e.g. disguising unpleasant odours, creating perfumes, etc) and are often created with "notes" in mind. The position of notes refers to temporal processing, where "top notes" are those perceived first, such as bergamot, followed by "middle notes", like eugenol, and "basic notes", such as angelica root (Poucher, 1993). Poucher (1993) describes the importance of understanding previously successful mixtures along with the importance of imagination on the part of the flavourist or perfumier.

Identification of the source of an odour is not the only way to communicate the olfactory experience to others. Baccino et al. (2010) asked "senders" to describe odours using any terms they wanted and "receivers" to attempt to match the description to the stimulus. They found that the descriptions contained terms that could be classified into five categories via qualitative analysis: descriptions of the source of the odour, hedonics, intensity, intrinsic properties of the odour (e.g. "a heavy scent") and the effect that the odour has on the person smelling it (e.g. "it grabs your nose"). One important finding from this study was that when the sender used more terms, recognition rate was significantly lower, suggesting that, while there are numerous ways to identify or describe an odour, limiting such information may be useful in describing an odour experience to others. However, such descriptions (intensity, hedonics, etc) can apply to a range of stimuli and may not be very useful for communicating information about the nature of the odour.

Effects of context – associations and expectations

There are various factors that have an effect on our perception of odours, such as stimulus intensity, repeated exposure, sex and hormonal state, age, emotional status, experimental instructions and previous semantic knowledge about odours (Rouby, Pouliot, & Bensafi, 2009).

Odour identification is not a purely bottom-up process. It is clear that expectations can influence the subsequent perception of an odorant, such as the colour of the stimulus in which it is presented (Koza, Cilmi, Dolese, & Zellner, 2005; Shankar, Levitan, Prescott, & Spence, 2009; Zellner, McGarry, Mattern-McClory, & Abreu, 2008). Similarly, the colour of a wine can have an effect on how it is perceived (Morrot, Brochet, & Dubourdieu, 2001; Pangborn, Berg, & Hansen, 1963). One particularly strong effect is odour-colour congruence, where a lemon solution is more likely to be correctly identified if it is coloured yellow rather than purple. This effect can be present even when the colour is not a part of the odour stimulus. Davis (1981) found that the presence of colour cues (either the actual colour or even just the name of the colour) can increase correct identification when the colour is congruent with the odour or increase incorrect identification when the colour is incongruent with the odour.

Top-down influences of odour perception and identification are not limited to components of the actual stimulus. Descriptions of food have been found to bias sensory perceptions in restaurants (Wansink, van Ittersum, & Painter, 2005) and hedonic perception in both adults and children (Bensafi, Rinck, Schaal, & Rouby, 2007; Lange, Martin, Chabanet, Combris, & Issanchou, 2002). In one experiment, half of the participants were told that pink ice-cream was "ice-cream", while the other half were told it was "frozen savoury mousse". The stimulus was smoked salmon ice-cream and those in the latter group, while still marginally disliking the stimulus, liked it much more than those who were expecting ice-cream (Yeomans, Chambers, Blumenthal, & Blake, 2008), indicating the role of expectation on olfactory perception and cognitions.

Summary

Olfaction and gustation appear to operate differently to the other senses. We understand that chemicals are responsible for odours, but we do not understand why odorants have the odour qualities that they do, which means that we cannot predict which odour will be related to a chemical based on a chemical analysis. Instead, we must sniff the odour and attempt to identify it.

While humans are able to detect and discriminate between olfactory stimuli quite well, identification of a stimulus in the absence of cues is a difficult task. However, while there are other ways to describe an olfactory stimulus (e.g. intensity, hedonics), they may only be useful when they are the unique and salient feature of that particular stimulus.

CHAPTER TWO - WINE PERCEPTION AND EXPERTISE

Wine expertise has long been an area of interest to psychologists. James (1890) identified wine expertise as an example of perceptual learning, stating that “one man will distinguish by taste between the upper and lower half of a bottle of old Madeira”. This chapter discusses the sensory experience involved in tasting wine, followed by a discussion of the nature and current theories of wine expertise.

The chemical basis of wine sensory qualities

Wine is generally considered to be a very complex beverage in terms of the number of aroma substances involved. One study estimates the number of volatile odorants in wine at over 800 (Marti, Mestres, Sala, Busto, & Guasch, 2003). Thus wine is much more complex than the mixtures used by Laing and colleagues (see page 9).

The odours involved in any wine can come from three different sources: the varietal character of the grape(s) involved (which also depends on the region in which the grape was grown and various vineyard variables), the winemaking technique (such as primary and secondary fermentation, oak treatment and skin contact) and the aging process (Jiang & Zhang, 2010). The specific combination of these factors results in differences between wines, such that a wine produced from one particular vineyard at a certain winery will taste quite different to a wine made from the same type of grape from another vineyard.

Chemicals in wine

Just as certain specific chemicals have been identified as character impact odorants for various odours (see page 3), specific chemicals have been described as key odorants in wine samples. For example, Shiraz is generally described as a “peppery” or “spicy” wine. Researchers at the Australian Wine Research Institute isolated the key odorant behind the peppery and spicy nature of wines made from Shiraz and identified it as rotundone, which is also the key impact odorant in white and black pepper (Siebert, Wood, Elsey, & Pollnitz, 2008; Wood et al., 2008).

Various character impact odorants have been identified in wines made from different grapes, such as 4-mercapto-4-methylpentan-2-one (4MMP), responsible for blackcurrant flavours in Cabernet Sauvignon and box tree or “cat’s urine” flavours in some Sauvignon Blancs, while the preferred Sauvignon Blanc flavours of passionfruit, gooseberry, grapefruit and guava are due to 3-mercapto-hexan-1-ol (3MH) and 3-mercaptohexyl acetate (3MHA) (Tominaga, Masneuf, & Dubourdieu, 1995; Tominaga, Murat, & Dubourdieu, 1998; Tominaga, Peyrot des Gachon, & Dubourdieu, 1998).

The flavour of a wine is based on more than just the chemicals in the grape. For example, Gawel, Royal, and Leske (2002) studied the effects of various different types of oak barrels on the sensory profile of Chardonnay, as judged by trained assessors. Amongst other things, they found that oaks with a stronger perceived oak influence have a lower perceived intensity of fruit. Finally, the chemicals responsible for specific aging flavours have also been studied, such as those found in aged Rioja (Aznar, Lopez, Cacho, & Ferreira, 2001).

While the chemicals described above primarily relate to odour sensations, chemicals such as malic and lactic acids and sugars are also present in wine and are the primary chemicals that are detected via taste receptors, along with bitter elements in the case of certain types of wine spoilage.

Furthermore, the sulphur-based preservatives, alcohol level and tannin (in red wine) stimulate the trigeminal nerve and the sense of touch, to round out the flavour of the wine.

How components of wine stimulate the senses

The first sensory contact with a wine usually involves vision. The visual properties of a wine (or beer) can change the perceived odour and flavour of the stimulus (Lelievre, Chollet, Abdi, & Valentin, 2009; Parr, White, & Heatherbell, 2003; Zampini, Sanabria, Phillips, & Spence, 2007). One famous example is the experiment by Pangborn et al. (1963), where white wine was coloured pink to resemble a rosé wine, a style that is often sweeter than most white wines. The rosé version of the wine was rated as a sweeter wine than was the white version of the wine, despite the colouring having no taste.

Visual cues in the wine can also be used to determine characteristics of the wine such as age (older wines tend to have a browner colour, for both white and red wine), the alcohol of the wine (such as observing the viscosity of the wine by examining the “legs”) and also to give cues about the winemaking technique (such as cloudy wine, where fining and filtration has not been used).

The next sensation will usually be orthonasal olfaction (sniffing). The volatile odorants react with olfactory receptor cells embedded in the olfactory epithelium. Whether or not these odorants are then consciously perceived depends on the complexity of the wine, the salience of the odorant (or the combination of odorants), the skill of the person and other factors, such as whether their attention has been directed towards or away from the odour with visual cues or experimental instructions.

Following orthonasal olfaction, the next step is usually “tasting” the wine, where the purpose is not just to detect the specific tastes (sourness, sweetness and perhaps bitterness), but also to determine specific flavours through the combination of these taste elements with the odours that are detected via retronasal olfaction, the trigeminal elements of sulfur and alcohol and the touch sensations of astringency, temperature and effervescence.

Why sensory evaluation is necessary

As discussed in Chapter 1, there is no way of predicting what odour will be perceived when an odorant is sniffed. We know the odours associated with certain odorants because people have previously sniffed and described their perception. Furthermore, the flavour of a wine is based on hundreds of different chemicals. These odorants can combine to form a new odour, or odour object, in unpredictable ways (Laing & Francis, 1989).

Wine evaluation can be used for various purposes. One such purpose is to detect elements within the wine, such as determining whether a wine contains “off” characteristics that indicate spoilage. One example of this is the detection of 2,4,6-trichloroanisole, known as TCA, which is the main chemical involved in cork taint (Prescott, Norris, Kunst, & Kim, 2005). Another is to communicate to the consumer what the wine tastes like (Gawel, 1997), which serves as an important marketing tool (Edwards, 1986). Thus, in order to communicate the properties of the wine, they must be described in enough detail to give the consumer an idea of what to expect to help them determine if it is one that they would like to buy.

What can wine experts do?

What is a wine expert?

Gawel (1997) described the difference between those who have practical experience in wine but no formal training (e.g. wholesalers, retailers and importers) and winemakers, who are not only highly experienced in tasting wine, but have also undergone formal training. Gawel described experience as familiarity with a class of products due to long-term exposure, where that exposure has occurred in conjunction with “considered thought as to the product’s sensory characteristics”, whereas training is “a uniform and directed program of instruction”. Melcher and Schooler (1996) outlined another categorization scheme for wine experts. They used three groups for their experiment. The first were non-red wine drinkers, who had virtually no perceptual or descriptive experience with the stimulus. Participants in the second group were regular wine drinkers, who had developed a palate for red wine, that is, had perceptual experience, and yet did not know how to describe wines with much precision. Finally, the wine expert group had developed extensive vocabulary dedicated to the chemosensory properties of wines.

Many studies into wine expertise use the experience criterion, rather than formal training. For example, the experts in the study by Hughson and Boakes (2002a) were required to have at least 10 years tasting experience in the wine industry, with no requirement of formal training. Parr and colleagues used an even broader category, allowing not just established winemakers, but also wine-science researchers, wine professionals (e.g. wine judges), graduate students in Viticulture and Oenology and people with more than 10 years of wine involvement (Parr, Heatherbell, & White, 2002; Parr, White, & Heatherbell, 2004).

While there is no accepted definition of what constitutes a wine expert in the literature, all agree that an expert requires not just a large amount of perceptual experience, but also some sort of non-perceptual training. In particular, the main aim of this training appears to be to learn how to use the extensive and precise vocabulary that other wine experts use. In the remainder of this chapter, the meaning of the term “expert” will be made clear by the context in which it is used.

Roles within the wine industry

Wine experts are employed in a number of different ways and thus need to use their skills for different purposes (Peynaud, 1987). A winemaker will taste grapes as they ripen to determine the optimal time for picking, while in the winery the same winemaker will taste the fermenting mixture to determine whether any adjustments need to be made (e.g. fault correction). Following this, the winemaker must work out which batches to blend together (including batches made from the same grape but from different vineyards, or batches from different grapes) and, if necessary, which batches to discard. Here, their sensory evaluation of the grapes and wine informs vineyard and winery decisions, with a particular emphasis on quality control.

The wine judge or wine writer is involved in a more independent evaluation of a wine in order to determine its quality. The wine judge must be able to verbalise why they gave a wine a particular score, while a wine writer must be able to describe the sensations of the wine to an audience in order to help them determine whether it is a wine that they would like to buy.

The wine retailer must then determine whether this is a wine that they would like to sell, which is usually determined by whether it is the style of wine their clientele is likely to buy. They must also be able to describe the flavours of the wine to the consumer (Peynaud, 1987).

The majority of these roles require the expert to describe the sensory profile of a wine. The following sections discuss the chemosensory abilities of wine experts.

Detection and discrimination performance by experts

A relatively old study examined detection abilities of wine experts and novices using wine-related stimuli (e.g. grape seed tannins) and found no significant differences (Berg, Filipello, Hinreiner, & Webb, 1955). Similarly, when testing detection thresholds, Bende and Nordin (1997) found no significant difference between wine experts and novices using 1-butanol, nor did Parr et al. (2002). However, 1-butanol is not a wine related odour, although it is widely used for detection threshold measures (Albrecht et al., 2008). A more recent study by Hayes and Pickering (2012) reported the testing of 331 participants who were classified as novices or experts through the use of a questionnaire. A relationship was found between the perceived bitterness of 6-*n*-propylthiouracil, a substance known as PROP that is commonly used to determine sensitivity to bitterness and taste in general, and wine expertise, such that mean rated PROP bitterness was significantly higher amongst wine experts compared to novices.

While experts do not appear to have the ability to detect odours that non-experts cannot, there is some evidence that wine experts may be more sensitive to particular sensations. Furthermore, there is some evidence to suggest that experts are better than novices at discriminating between wine samples. Perceptual training using wine can increase performance in a task that requires a same/different judgement of two wines, despite an initial and persisting bias towards a “different” judgement (Owen & Machamer, 1979; Walk, 1966). This has also been tested using a “triangle test”, where two of three samples are identical and the task is to determine which one is different. While experts are better than novices in some experiments (e.g. Solomon, 1990), this difference is not always observed (e.g. Solomon, 1997). Thus, there appears to be limited evidence that experts have superior detection or discrimination ability compared to novices, even for wines.

Recognition performance by experts

Recognition tasks generally require the participant to taste a wine, or a number of wines, and then later identify it or them against a series of foils. Wine experts generally perform better than novices at wine recognition tasks (Lawless, 1984; Solomon, 1990). In these studies, experts and novices described a set of wines and later selected the wine that matched either their own descriptions or descriptions generated by other novices or experts – a task known as a matching task. Experts can match their own descriptions (Lawless, 1984), as well as the descriptions of other experts (Solomon, 1990), to the wines, but novices perform poorly at matching their own descriptions, or a composite of descriptions written by other novices, to the wines (Lawless, 1984). Valentin, Chollet, and Abdi (2003) performed a similar experiment, where they asked novices and experts to describe wines and then match the descriptions of others to the same wines. They found that both experts and novices were able to match the descriptions that were written by experts more easily than those written by novices. Solomon (1988) attributed the superior ability of experts in terms of matching tasks to the more consistent use of verbal descriptors.

Similar findings have been reported from matching experiments using beer experts. Chollet, Valentin, and Abdi (2005) reported a slightly different matching procedure, where untrained novices and trained beer assessors each described 18 beers. During the matching task, each beer was accompanied with two descriptions written by other participants. One of the descriptions was written about that particular beer, while the other was drawn from one of the other 17 beers at random. In general, the trained assessors could match the descriptions written by other trained assessors at a rate significantly higher than chance, but were less successful with descriptions written by novices. As expected, novices were not particularly successful in matching descriptions even when written by experts. However, one interesting finding was that when the beers were supplemented (i.e. contained additives to enhance specific flavours), the novices could match the descriptions written by trained assessors. This suggests that novices can gain some meaning from expert descriptions if the task is made easier by enhancing the most salient aspect of the beer.

Gawel (1997) found that wine experts with formal training produced better descriptions than those who had experience with wine, but no formal training. Furthermore, Gawel asked his participants to underline the words in the descriptions that helped them the most during the matching task and found that, for those with no training, tactile and palate intensity terms actually impaired performance. When the group that had not had formal training could match their descriptions, they underlined concrete aroma terms and some tactile terms. Gawel took this to mean that experts with and without formal training rely on the perception of different aspects of the wine. More specifically, untrained experts appear to rely mostly on aroma, while trained experts appear to rely on more than just aroma. Taken together with the finding by Valentin et al. (2003) that the experts' descriptions were more precise (discussed in more detail below), it appears that differences between experts and novices in matching tasks are likely to be due to the communicative value of the descriptions (Gawel, 1997). That is, the participant needs to be able to make a connection between the description and the sensation in the wine.

Parr et al. (2002) used a different procedure to test wine recognition memory. They presented experts and novices with 12 wine-related stimuli (such as vanilla/oak and buttery) and asked the participants to smell them and then attempt to identify and remember each stimulus. After 10 minutes, they were presented the same 12 wine-related stimuli amongst 12 distractor stimuli. For each odour sample, the participants were again asked to state whether they were old or new stimuli. As previously noted, no significant differences were found between experts and novices in terms of general sensitivity, but in the same study they found that experts do perform better in olfactory recognition tasks when using wine related stimuli, including chemicals with the odour properties of vanilla/oak, butter and melon.

Zucco, Carassai, Baroni, and Stevenson (2011) tested novices and experts in a wine recognition task by asking the participants to smell (but not taste) a wine, followed by four wines a few seconds later. One of these four wines was the original target. The experts performed significantly better than the novices at this task, suggesting that the superior performance in wine recognition tasks may not require explicit verbalization.

Identification performance by experts

Wine experts have been tested on their ability to identify odours that are not related to wine samples. Zucco et al. (2011) found that sommeliers that are currently undergoing or have completed formal sommelier training, do not perform better than untrained wine drinkers in terms of identifying everyday odours, whether they were wine-related (e.g. leather, rose, tobacco) or everyday items (e.g. oregano, garlic, fish). Similarly, Parr et al. (2002) found no difference between experts and novices in terms of identification, even in the case of wine-related odorants. However, Bende and Nordin (1997) found some evidence of a difference in identification ability between novices and experts, although they acknowledged that this finding was possibly due to a few of the odours that they used (e.g. cloves, lemon) because wine experts may have had more professional exposure to them, suggesting that this finding may have been specific to that set of odorants.

The results above refer to identification of single-odour stimuli. When tested on multi-odour solutions, there appears to be some evidence that experts are better than identifying elements within multi-odour mixtures. Livermore and Laing (1996) tested experts and novices using mixtures of up to five odours, asking them to identify as many odours in the mixture as possible. The experts did not perform better than novices on single odour stimuli, but did perform better on stimuli containing two or three odours. When the mixtures contained four or five odours, once again there was no significant difference between groups, although this is presumably due to a floor effect as performance by both groups was quite poor.

Taken together, these findings suggest that expert performance does not depend on a superior ability to identify a single odour. However, experts do appear to perform better than novices at identifying elements within multi-odour solutions. All of these studies refer to identification of odours that were presented in isolation from other odours and not in a wine context.

Identification of a stimulus, or important components of a stimulus, appears to be a key factor in that it has an effect on processes such as detection thresholds (Tempere et al., 2011), hedonics (Hersleth, Mevik, Naes, & Guinard, 2003) and the subsequent acceptance of a wine (Blackman, Saliba, & Schmidtke, 2010), along with odour memory, as seen in matching tasks (Melcher & Schooler, 1996). Thus, understanding the factors that aid identification of wine is important.

Description performance by experts

When smelling a wine, it appears that experts are generally able to identify major attributes of the wine and to create more detailed and accurate descriptions of wines compared to novices (Lehrer, 1975). The poor matching performance of novices described above (see Recognition performance by experts) has been attributed to their poor descriptions (Lehrer, 1983). Furthermore, the descriptions by experts tend to contain more specific features of wines compared to novices and to those with a lower level of expertise, termed “intermediates” (Solomon, 1997) and appear to contain more concrete terms compared to non-experts (Lawless, 1984). Thus the matching ability described above may be because, within a wine context, experts are able to not just extract the most relevant features of a wine, but also to describe the features in a way that other experts can understand, despite not displaying superior performance in identifying wine-related odours when they are isolated from a wine sample.

Similarly, Valentin et al. (2003) asked experts and novices to describe wines and then match descriptions generated by others to the same wine samples. They found that the experts tended to use more precise terms to describe the wines. Furthermore, they found that the descriptions of experts differed the most from novices when they used specific (e.g. strawberry) as opposed to general (e.g. fruity) terms.

Experts also describe more than just the specific flavours of a wine. In a wine evaluation, an expert can also describe configural properties, such as the “length” or the “balance” of a wine (Hughson & Boakes, 2001), or the aging potential of a wine (Langlois, Dacremont, Peyron, Valentin, & Dubois, 2011).

Novices can rate wines in terms of hedonics, but they are usually quite poor at stating why they like or dislike a particular wine. In contrast, experts are better at describing the reasons why they like or dislike a particular wine (Hopfer & Heymann, 2014). However, the same study indicated that experts do not always agree with each other in terms of ratings that may be more subjective, such as the overall quality of the wine.

However, experts are not infallible in their descriptions and appear to be particularly susceptible to visual mismatches. Morrot et al. (2001) found that when their trained panel tasted white wines that had been artificially coloured red with an odourless dye, they tended to use descriptors that were suitable for red wines. Similarly, as described above, Pangborn et al. (1963) found that white wines coloured as rosés were judged as being sweeter than the same white wine samples when they were not coloured pink.

Categorisation tasks

Experts' descriptions tend to covary by grape type compared to non-experts and, when asked to cluster wines together, experts are more likely to do so according to grape type (Solomon, 1997). Ballester, Dacremont, Le Fur, and Etievant (2005) asked a panel of staff and students from the Université de Bourgogne and their friends, all of whom were considered to be novices, to sort 18 wines into as many categories as they liked. These wines were made from numerous different white wine grape varieties. The results suggested that the sensory boundaries between grape types are not clear-cut to novices.

Ballester, Patris, Symoneaux, and Valentin (2008) used a similar procedure to compare wines made from only two grapes: Chardonnay and Melon de Bourgogne. They found that experts could categorise the wines based on grape type, whereas novices could not.

However, experts are not always superior to novices in categorisation performance. Ballester, Abdi, Langlois, Peyron, and Valentin (2009) asked experts and novices to smell red, white and rosé wines in dark glasses, so that the colour was obscured. They were then asked to categorise the wines into red, white and rosé categories based on smell alone. Contrary to their expectations, experts and novices were both able to categorise red and white (but not rosé) wines, with no significant difference in performance. While this task is not a descriptive task, it does require some knowledge about what these categories are.

How do wine experts perform these tasks?

Before discussing proposed theories of wine expertise, it is useful to look at studies of expertise in other domains. Early theories dating back to Galton (1869) were based around the idea that experts possess superior innate capacities to novices and that these capacities cannot be improved with practice. As noted earlier, recent studies suggest that this is unlikely to be the case. de Groot (1978) and Chase and Simon (1973) examined chess expertise and concluded that experts do not survey every piece and every possible move on an entire chess board. Instead, it appears that experts detect patterns (such as the relative positions of two or three chess pieces) and that these patterns are detected very quickly, rather than requiring an extensive search. Chase and Simon (1973) proposed that chess expertise was not innate, but instead was determined by acquiring knowledge in their domain over many years.

The studies comparing general olfactory sensitivity in wine experts (described above), in which it has generally been found that experts do not have superior olfactory abilities to novices, provide evidence against Galton’s theories that wine expertise is innate. Instead, wine expertise is acquired through years of training. This training, along with training in other domains of expertise, is specific to the stimuli of that field. That is, wine experts are trained on wine-relevant stimuli, while perfumiers are trained on perfume relevant stimuli, and so on. Because of the specific nature of this training, it is not surprising that expertise in one domain does not transfer to another, whether the domain is chemistry or wine (Bende & Nordin, 1997; Lawless, 1984).

One theory of wine expertise is that experts can identify more flavours in wines than novices, due to knowledge gained through training (Gawel, 1997; Lawless, 1984; Solomon, 1997). Hughson (2003) used the example of a wine expert tasting a wine, detecting a lime note and then inferring that the wine is likely to be a Riesling, as lime is a distinct feature of Riesling, but not other white wines, such as Chardonnay. Based on this initial assessment, the expert then searches for corroborating features that have been found in other examples of Rieslings that the expert has previously experienced. As the search is directed based on knowledge, the search is more efficient than that of novices. That is, the experts know how to filter and use the information available to them in an initial sensory evaluation in order to maximize their chances of successful identification based on rules that are learned from previous experience and training. This is similar to the aforementioned studies of chess experts, where the expert filters which pieces are likely to be useful for the next move by looking at groups of pieces, rather than each individual piece, and then using that information to guide further decision-making.

For wine experts, these knowledge units may be based around grape type, or varietal, as suggested by the findings (described above) that experts categorise wines based on grape type, whereas novices do not. There is further evidence for this hypothesis in the studies reported by Hughson and Boakes (2002a). The authors tested novices and experts on verbal short-term memory of wine descriptions, without requiring the participants to describe the wines. The descriptions consisted of three terms. When these three terms all described the characteristics of a wine made from one grape type, experts displayed superior recognition performance than when the three terms were shuffled, so that the three words together did not describe the characteristics of wine made from one grape type. This finding corresponds with findings in other areas, such as chess expertise, where experts can recall the positions of chess pieces on a board when they could occur during a game of chess, but not when the placement of pieces was not possible during a game (Chase & Simon, 1973).

Summary of the nature of wine expertise

When the layperson thinks of wine experts, they often think their abilities either equate to “mystical powers” (Stone et al., 2012) or are fraudulent (Sage, 2002). While experts do not appear to have more sensitive chemosensory systems compared to novices (Marino-Sanchez et al., 2010; Parr et al., 2002) they do appear to be better at describing wines in a way that other experts can understand (Gawel, 1997), as evidenced by matching experiments (Lawless, 1984; Solomon, 1990). Their ability to describe odours appears to be specific to wines and does not appear to translate to wine-related odours when smelled in isolation of wine (Zucco et al., 2011). Thus, it appears that the abilities of wine experts are due to how they process the information they receive, including being able to categorise wines (Ballester et al., 2008) and describe wines (Gawel, 1997; Lawless, 1984; Solomon, 1990; Valentin et al., 2003). Finally, these abilities do not appear to be innate. Instead, they are learned via expertise and stimulus-specific training (Gawel, 1997) that does not appear to generalise to other stimuli (Chollet et al., 2005).

CHAPTER THREE – TRAINING OF THE CHEMICAL SENSES

Different types of chemosensory training are available for different purposes. This chapter first discusses wine training books and courses that are typically available to the general public, followed by the more intense training that is used in chemosensory industries, such as the training of sensory panels and perfumiers. The final section of the chapter discusses training in the academic literature, including sections on perceptual learning and various studies that have involved olfactory training.

Commercial wine training courses

Books

Wine training books are usually aimed at novices, although they can include information relevant to all levels of expertise. They usually include some theoretical training (e.g. the steps involved in winemaking and the influence they have on the perceptual attributes of the final product) as well as training in terms of the process of tasting (e.g. swirling the wine) and what to pay attention to, starting with the visual aspects of the wine, the orthonasal smell of the wine and then the flavour along with any other sensations detected in the mouth. Such books cover this material in various levels of detail (Betts, 2013; Goodall & Eyres, 2013; Jackson & Jackson, 2009; Kreider, 2011; Robinson, 2008; Schuster, 2009), sometimes including which odours and flavours are typically found in wines made from certain grapes, but do not describe any actual training methods. One of the more advanced books is "How to Taste: A Guide to Enjoying Wine" by Robinson (2008). It includes some practical exercises to help the participant learn more about their senses, such as wearing a blindfold and asking an accomplice to pass the participant a food that they like, along with another that is similar but slightly different (e.g. smoked salmon and smoked mackerel) in order to compare and contrast the stimuli. She then suggests doing the same with a red and white wine, while blindfolded. Robinson (2008) also suggests using commonly available products to learn about aspects of wines, such as using apple juice as an example of malic acid, while using milk as an example of lactic acid (p.21). However, these types of exercises are rare in the wine training books.

Most books recognise that the difficult aspect of wine training is learning to verbalise more than just hedonics, such as the specific odours and flavours of a wine. In order to make this particular task easier for novices, some books indicate the particular odours that are usually associated with each type of grape, such as blackcurrant (or cassis) typically being associated with wines made from Cabernet Sauvignon (Goodall & Eyres, 2013). At least one includes a “scratch and sniff” component, where 16 wine-related odours are released when scratched (Betts, 2013). The assumption behind these guides is that the reader can recognise that particular odour in a wine context.

While most of these books discuss the typical flavour profiles of wines made from particular grapes, they usually do not offer practical exercises, with the exception of the books described above. Thus, they are generally aimed at improving the conceptual, rather than the perceptual, knowledge of the reader.

Internet-based courses

Winetasting courses are now offered online, bringing the added benefit of interaction. One example ("The Complete Wine Course," 2008) includes both the type of theoretical knowledge found in books, as well as an interactive way to enter tasting notes using drop-down menus (e.g. Depth of Colour can be Pale, Medium Deep, Deep or Opaque). In terms of entering the specific odours or flavours associated with the wine, an interactive, clickable version of the Wine Aroma Wheel (Noble et al., 1987; Noble et al., 1984) appears, giving the user a structured list of odours to choose from (e.g. Cherry, Mint, Eucalypt). The training associated with this particular program also includes interactive quizzes that test theoretical knowledge, once again based around conceptual knowledge rather than perceptual knowledge.

Practical wine training courses

One limitation to books and online courses is that, with the exception of innovations like the scratch and sniff book written by Betts (2013) and examples from Robinson (2008), it is left to the novices to match the descriptions of the odours and flavours with the actual stimuli within the wine, based on whatever wines are available to the participant.

While there is little published literature on the nature of these courses, the following information comes from personal experience and correspondence with instructors in these courses. Unlike books and online courses, practical wine training courses involve a class of participants tasting wine together, under the lead of an instructor. As well as this practical tasting element, these courses also include the type of theoretical material described in wine books. They can vary from single, hour-long sessions to courses that run for months, depending on the purpose of the course and the desired level of training. The instructors usually select the wines on the basis that they are faithful representations of their style. They may be single-varietal wines or wines blended from numerous varietals. A course may cover just one style of wine (e.g. just Australian Shiraz) or many styles, such as courses where the participants taste an example from each of the major styles that are common in the market. Furthermore, the wines may come from the same region, different regions and even different countries, allowing the participants to taste a wide range of wines. Classes that focus on one particular theme are sometimes described as “masterclasses”. The number and type of wines tasted depends on the length and cost of the course.

In general, the instructor will describe the particular flavours of each wine as the participants taste them and participants are generally not required to produce their own descriptions of the wines, although they may be encouraged to do so. The advantage of these courses is that the participants not only get to taste the wines, but they get to talk about them with the instructor and fellow participants if they wish to do so. This discussion can serve as corrective feedback, which is an important training mechanism. Unlike most books, these practical courses aim to build on the conceptual *and* perceptual knowledge of the participants.

A different approach is that of Wine Awakenings (Wine Awakenings Inc, 2014), a North American company that manufactures and sells varietal-specific aroma kits containing 12 of the most common aromas found in wines such as Chardonnay, Cabernet Sauvignon and Pinot Noir, as well as a wine fault detection kit. Similar kits are available from Le Nez du Vin (Le Nez du Vin), which is French for “the nose of the wine”. Customers smell the aromas within these kits and learn to identify the aromas by name, although these are isolated odours and are not presented in wine samples. The kits also include wine-related literature, such as “How to Taste Wine” (Robinson, 2008).

Formal tertiary winemaker training

Those who study undergraduate qualifications to become winemakers undergo formal training and assessment in all areas of winemaking. For example, the 2015 syllabus for the Bachelor of Viticulture and Oenology at the University of Adelaide (University of Adelaide, 2014) includes study units in biology, chemistry, physics, statistics, agriculture, viticulture and sensory analysis. According to the sensory analysis course outline, the students are taught the theory behind chemosensory perception. The course also includes a practical program, during which the students are taught basic skills in the sensory assessment of wine using model solutions to depict basic tastes and their interactions.

Intensive training

Sensory panels

As described in Chapter 1, instruments and machines can detect the presence of chemicals, but cannot describe the odour or flavour of a stimulus (Piggott et al., 1998). Instead, relatively small panels undergo extensive training in order to be able to provide a sensory evaluation of a stimulus, usually for commercial purposes. Panelists have been used in a variety of industries, such as food and beverage, agriculture, flavour, consumer products and perfume.

Uses of sensory panels. There are two types of panels: consumer panels and sensory panels. Consumer panels rate products on factors that are of interest to marketers, such as hedonics and purchase intent (Lawless & Heymann, 2010). As consumer panels do not usually undergo any type of training, they will not be discussed further here. Sensory panels are used for a variety of tasks, such as determining whether a new product on the market is sufficiently discriminable from an existing product, whether particular odours or flavours are detectable in a product (i.e. detection and identification of components) and the relative strength of each of these odours or flavours within the mixture (Meilgaard, Carr, & Civille, 2006; Stone et al., 2012). These ratings from panels, which are considered to be more objective than the ratings from consumer panels (Meilgaard et al., 2006), can then be compared to consumer hedonics to determine which components drive liking and, presumably, sales. While some panels only rate stimuli on a number of attributes, others may also provide a verbal description of the stimulus.

Sensory panels are considered to be analytical instruments in that the responses from panelists should be objective, although there may be some question as to just how objective such a rating by a human can be. While there are variations of the procedure that leads to these ratings, such as The Flavour Profile Method and Quantitative Descriptive Analysis, all fall under the umbrella term of *descriptive analysis.*

Panels are efficient in terms of sample size, generally consisting of between six and 20 trained assessors (Stone et al., 2012), with 10 to 15 being a typical number (Næs, Brockhoff, & Tomic, 2011). As the purpose of these panels is not to represent the subjective responses of a population, the demographic composition of the panel is not important.

As different panels work in different industries, with different stimuli for different purposes, there are a myriad of considerations in terms of training procedures (Næs et al., 2011). The following discussion on panels is an overview of these considerations.

Training of sensory panels. Meilgaard et al. (2006) described a typical training method for sensory panels. Initially, panelists are instructed to “precondition” their senses by not doing anything that could impede their sensory acuity, such as wearing perfumed cosmetics or eating food with a strong flavour. The subsequent step is to demonstrate the correct technique when using equipment, while stressing the importance of using the same routine at all times. Furthermore, panelists are informed that they should ignore their personal preferences and focus on the task at hand, such as detecting differences between and amongst samples.

Panelists should initially be presented to samples with “large, easily perceived sensory differences”, which are subsequently replaced with sequentially smaller sensory differences until panelists are confident they can detect the difference. Sensory panels rate numerous products in one sitting, which raises the problem of adaptation or fatigue. To counter this problem, Meilgaard et al. (2006) suggest that panelists take shallow sniffs and leave "tens of seconds" between each sample.

In some cases, panels need to be trained on particular sensory attributes that are to be identified within samples. In these cases, panelists are required to be able to not just detect the attributes, but also to use the appropriate terminology to describe them (see below). Furthermore, they are trained to use a line scale to indicate how much of the attribute is present (Stone et al., 2012). The ratings on these scales are not intended to be relative ratings, that is, it is not simply a matter of whether one product has more or less of an attribute than another product. Instead, they are supposed to be absolute ratings that indicate how much of the attribute is present and should be consistent between samples with the same attributes and testing sessions, although there is necessarily a degree of context in terms of order of stimuli. These scales are calibrated both within and between panelists (Meilgaard et al., 2006; Næs et al., 2011).

Similar information about selecting, training and monitoring panels is detailed in International Organization for Standardization (ISO) and American Society for Testing and Materials (ASTM) standards (American Society for Testing and Materials, 1981; International Organization for Standardization, 2012). However, these standards are designed for panels trained in any of the senses and are thus somewhat vague, with statements such as "Multiple samples may be analyzed within a panel session; the actual number of samples to be analyzed is fatigue-dependent" (American Society for Testing and Materials, 1981), without any indication of how many samples may induce fatigue for any of the senses.

In practice, panelists are usually trained on between 10 and 20 different attributes that are relevant to a product (Næs et al., 2011). Training and calibration continue through regular training sessions in order to maintain performance, so the exact amount of training varies widely depending on a number of factors. However, Lawless (1999) urged a note of caution that the descriptive analysis method of dividing the perception of an odour into “simple and apparently independent scales may produce the illusion that the odour experience is a collection of independent analyzable ‘notes’ when it is not.” Wolters and Allchurch (1994) also state that, while panelists are trained on numerous attributes for any one product, it is likely that only a small number of these attributes are “pre-determined, more or less objective and known dimensions in which products differ from each other”. Thus, while most descriptive panel methods require training on these attributes and panels are generally deemed to be as objective as possible in terms of their measurement of these attributes, there is some question as to whether their ratings are as objective as they are often thought to be. Examples of possible biases are the halo effect or proximity error, where the judgement of one attribute may influence the ratings of subsequent attributes (Kemp, Hollowood, & Hort, 2011).

Furthermore, panels have been used for wine perception studies (e.g. Heymann, Hopfer, & Bershaw, 2014; Pagès, 2005; Perrin et al., 2008; Ross, Weller, & Alldredge, 2012) they have been found to be more sensitive to changes in stimuli than regular consumers (e.g. King & Heymann, 2014). Carlucci and Monteleone (2008) described a study where panel training occurred in two parts. The first part involved training the panelists in aroma identification, while the second involved intensity evaluation of the aroma descriptors.

Amount of training required. Panels are specific, in that they need to be trained on the particular product that they are employed to test. The length of training varies depending on the complexity of the product, but can range from 25 hours to 100 hours for training on multiple attributes (Meilgaard et al., 2006). This amount of training is required in order to be able to detect, identify and quantify numerous attributes within a sample.

However, lengthy training is not always necessary. Stone et al. (2012) described training protocols that last between 8 and 10 hours. Similarly, Kreutzmann, Thybo, and Bredie (2007) described the training of a 10-member sensory panel that consisted of seven sessions over four days. This panel was trained on 13 attributes over five different carrot genotypes and was presented with each genotype once per training session. Furthermore, feedback was only given during the first four sessions. Thus, the amount of training required for these panels appears to differ dramatically depending on the purpose, but it is clear that effective training can occur in a relatively short period of time.

Wolters and Allchurch (1994) compared four different panels that had undergone different amounts of training. They compared a conventionally trained descriptive panel (60h of training), a panel that received reduced amounts of the same type of training (30h), a panel that performed a profiling task rather than a descriptive task (15h) and an untrained panel (0h). They found that those who received more training performed better in some measures, such as the absolute number of discriminating attributes used, but also found that all of the groups that received training were still able to rate the attributes of the stimuli in a similar way. The result suggests that while some of the more precise tasks performed by panels may require extensive training, a relatively small amount of training may still allow a panel to produce accurate and useful data.

For example, Etaio et al. (2010) described the training of a wine sensory panel. Their assessors were trained over 15 sessions, each one between 90 and 120 minutes on eight sensory parameters of the wines: odour intensity, odour complexity, aroma intensity, aroma complexity, balance and body, global aroma persistence, colour hue and colour intensity. Most of these assessors had previous experience in sensory descriptive analysis in cheese and/or wine evaluation. As with other sensory panels, this task still requires the panelists to be able to identify the relevant attributes in order to be able to rate them on that particular scale. However, the participants in this particular study were not required to provide a description of the particular flavour of the wine. Despite this relatively short amount of training, the panel was able to perform to an acceptable standard.

Training a panel to identify and quantify attributes of a sample requires a large outlay in terms of time and resources. As such, methods have been proposed to reduce the amount of training time, such as being allowed to reevaluate the sample after feedback (Findlay, Castura, Schlich, & Lesschaeve, 2006) or to remove the requirement of precise quantification of attributes by ranking products along dimensions (Cartier et al., 2006).

Perfumiers

Calkin and Jellinek (1994) described perfumier training as a matter of learning the thought processes that connect the perception of the odour to its name, rather than perceptual training, where the aim is to train the nose or receptors. Their students were initially introduced to no more than 50 of the most important materials. Over the course of several weeks of training, this number reached 162 materials. Their task was to learn to classify odours according to a system by comparing and contrasting different odours. Some comparisons were between two similar odours (e.g. two different woody notes), while others were between two more distinct odours (e.g. a rose note and a woody note). At each training session, they would smell and then attempt to identify each material, receiving feedback from the teacher and other students when they were incorrect.

Gilbert (2008) described another similar training method, called the Givaudan method. Here, odours are presented in a grid of rows and columns, where each row is a family of fragrances (e.g. floral, woody, citrus) and each column is a training session. In the first session, students sniff each sample in a column in order to compare and contrast different families. This process continues for as many sessions as there are columns. Once this training is complete, the same process occurs across rows, where different odours from within a family are compared. Thus, the method uses direct comparisons in order to help the novice perfumer to learn to identify the important elements of perfume making.

These training procedures both rely on contrast between stimuli in order to highlight the differences between them. This concept is discussed further in the perceptual learning section below. A second important point to note about these methods, as well as the methods for training sensory panels, is that the trainees are required to use terminology that is standard across the industry. Thus these studies can inform a better understanding of wine sensory analysis.

Standard terminology

Different industries have different standard terminologies to describe attributes of the stimulus. In order to improve training, terminology prompts have been developed, often in the case of wheels, such as the Beer Flavour Wheel (Meilgaard, Dalgliesh, & Clapperton, 1979). This wheel served as inspiration for the well-known Wine Aroma Wheel, initially published by Noble et al. in 1984 and updated in 1987. The latter is arranged in three tiers, with the inside tier describing broad categories (e.g. fruity, nutty), the middle tier describing subcategories (e.g. tropical fruits, berry fruits) and the outside tier describing specific odours and flavours (e.g. pineapple, raspberry – see Figure 1). The Beer Aroma Wheel also includes attributes other than the specific odours, such as mouthfeel (e.g. fizzy, flat, creamy) and bitterness. In contrast, Noble and colleagues focused mostly on olfactory attributes, although some trigeminal attributes (hot, cool) were included.

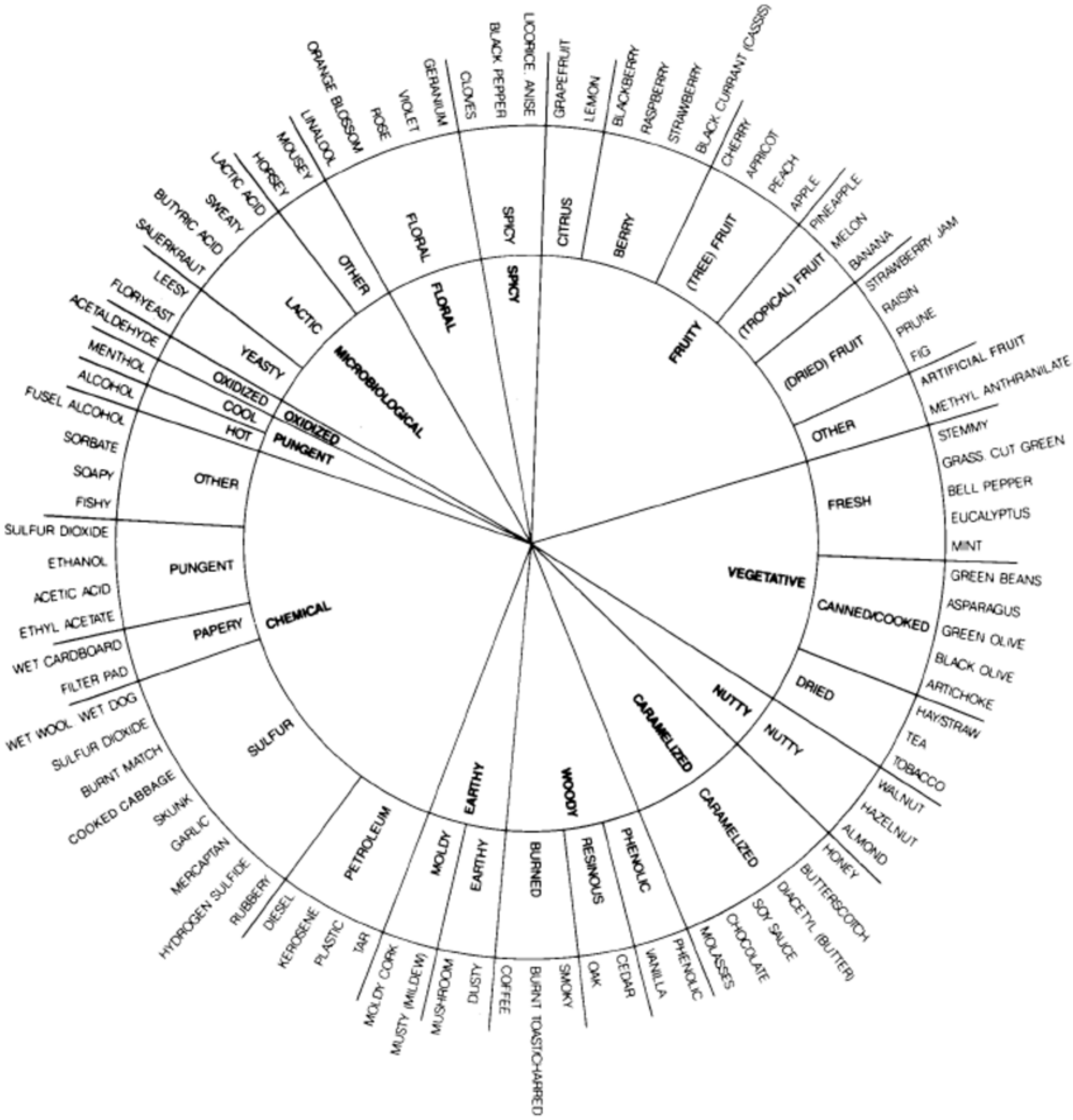

Figure 1. The Wine Aroma Wheel (Noble et al., 1987).

The purpose of the Wine Aroma Wheel is to serve as a prompt to help reduce the difficulties of odour identification that were described in Chapter 1. Noble et al. (1987) also included reference standards that could be created by taking either red, white or rosé wines (depending on the standard to be created) and adding specified amounts of additives. For example, the standard for apricot is created by taking 25mL of white base wine and adding 15-20mL of apricot nectar, while the standard for mint can be created by placing 1 crushed mint leaf or 1 drop of mint extract into 25mL of either red or white wine. These standards were provided to increase the utility of the aroma wheel and were based on standards that were created for the Beer Flavor Wheel (Meilgaard, Reid, & Wyborski, 1982).

The Wine Aroma Wheel is a useful guide in terms of identification of many specific odour and flavour attributes within wines. Descriptive analyses of wines involve more than just the specific odours and flavours. In general, they include notes about the colour, specific odours, relative levels of acidity and sweetness, “balance” (a term used to describe the relative harmony of various attributes of the wine), length (persistence of aftertaste) and various other attributes of the wine (Peynaud, 1987). As the purpose of this thesis is to determine the factors that help novices to identify specific odours within wines, the following discussion focuses on this particular aspect of wine expertise.

Olfactory training in academic research

Perceptual Learning

Perceptual learning refers to “relatively long-lasting changes to an organism’s perceptual system that improve its ability to respond to its environment and are caused by this environment” (Goldstone, 1998). These changes can occur due to intentional learning, such as when a rat receives food following a certain response (e.g. a bar press during a conditioning trial) or when a human is told whether they were correct in their choice, or via incidental learning, where changes are driven by mere exposure (Mitchell & Hall, 2014).

The abilities of wine experts are often cited as examples of perceptual learning. As mentioned in Chapter 2, James (1890) cited the example of being able to distinguish wine by taste depending on whether it came from the top half or the bottom half of a bottle of Madeira. Similarly, the descriptive and matching abilities of experts described in Chapter 2 derive from extensive training with various forms of feedback.

In order to describe the predominant theories of perceptual learning with feedback, it is first necessary to explain the notation used. Stimuli are generally considered as compounds, consisting of numerous features (Mitchell & Hall, 2014). Two stimuli may be described as AX and BX, where A and B are the features that are unique to each compound and X refers to the features that are common to the compounds. To extend this notation to wines, one could consider two red wines: a Shiraz and a Pinot Noir. In this example, X refers to the features that are common to these red wines, such as alcohol, tannin, etc. A, then, refers to the features that are unique to the Shiraz, such as a spicy or peppery element, while B would refer to the unique elements of the Pinot Noir, such as cherry flavours or gameyness. Given that wines can differ so much depending on the vintage, the location of the vineyard, winemaking techniques and numerous other factors, the components that fall into A, B and X depend on the set in use.

An important finding in the perceptual learning literature is that intermixed trials (AX, BX, AX, BX, etc) result in more learning than blocked trials (AX, AX, BX, BX) (Dwyer et al., 2004; Symonds & Hall, 1995). While there may be different explanations for this (see Mitchell and Hall, 2014, for a summary), this particular finding is a useful consideration in terms of olfactory training. The same principle is used in the training of perfumiers, as described above.

Perceptual learning does not require explicit feedback. It appears that mere exposure to a stimulus can result in perceptual learning (e.g. Hall, 1980). However, the purpose of this thesis is to discuss active training procedures, so the accounts of perceptual learning without feedback will not be further discussed here.

Furthermore, most perceptual learning research has been conducted using non-human animals. While these animals can learn to identify which compound is associated with a reward, they obviously cannot name the unique element. The earlier sections of this chapter described the typical training that is involved in commercial wine courses or sensory panels. The next section discusses academic studies in which the participants have been actively trained to identify either stimuli as a whole (e.g. single-odour solutions), or components within stimuli (in the case of multiple-odour solutions).

Olfactory experiments

As discussed in Chapter 1, Cain (1979) described the difficulties of odour identification. Cain’s twelve participants also underwent training on the 80 odours that they smelled. After smelling each odour, they rated them in terms of familiarity and then attempted to identify the odour. The stimuli were then presented a second time and participants were asked to name the stimuli, in order to assess how consistent they were in terms of odour identification. Cain gave his participants feedback during this second block.

This feedback was based on the labels that the participant had used during the first assessment of each odour, whether the labels were correct or not. After this second block, the participants returned for four subsequent test sessions separated by 2-3 days, where the procedure in the second block was repeated. In the initial block, the average number of items correct among the 80 odours was 36, with a range of 25 to 43. By the final session, this number increased to an average of 61.5 out of 80 correct.

Cain also compared the relative performance on correct, or veridical, labels, to near misses (e.g. describing cloves as nutmeg) and to far misses (e.g. describing machine oil as cheese). He described near misses as “serviceable”, in that they may have still been useful as they applied to an odour that was similar to the one to be identified, while far misses do not. Where participants had initially used a veridical label, then they rarely failed to apply it during the subsequent sessions. However, performance using near misses was frequently incorrect and far misses even more so. Furthermore, some participants were allowed to change their labels during training resulting in better performance, as more than half of these changes improved the label (e.g. far miss to veridical). As in some perceptual learning studies, feedback was used as a training mechanism, here resulting in an improvement in performance over a large set of odours in a relatively small amount of training compared to that used for sensory panels.

Desor and Beauchamp (1974) reported a similar study, where three participants were trained to identify each of 32 everyday odours three times a day until they reached perfect performance, using feedback as a training mechanism, which took 5, 9 and 11 days for the three participants. Five days after the last participant reached perfect performance, all three participants were tested on the set of 32 odours, each presented three times, for a total of 96 trials. The three participants were correct in 94, 95 and 96 trials respectively in this test session.

Taken together, these two studies show that training to identify a large number of odours can occur within a relatively short period of time and that feedback is a very useful training mechanism. However, in the studies above, the participants were all trained on single-odour stimuli. As described in Chapter 1, when odours are mixed together, identification is a much more difficult task.

Livermore and Laing (1996) used a trained panel of ten women and an expert panel of eight professional perfumers and flavourists to investigate the influence of training and experience on identifying odours within multicomponent mixtures. Their participants smelled stimuli consisting of between one and five of the seven odours used in the experiment and indicated which odours they believed were present in the stimulus using a graphics tablet. At the conclusion of each trial, they received immediate feedback, informing them of the correct answer or answers. In this experiment, both groups were trained on both the individual odorants and, during two weeks of testing, on random mixtures of the odours. The participants reached perfect identification of the individual odorants within five days of relatively light training. However, training did not improve performance in terms of the number of odorants in a mixture that could be identified. However, the expert panel did perform better than the trained panel at identifying 2- and 3-odour solutions, indicating that experience can have an effect on performance.

The distinction between the studies by Cain (1979) and by Livermore and Laing (1996) is that the participants in the latter experiment were given multiple choice prompts, whereas Cain’s participants were not. Despite these differences, performance in both groups improved with training. That Cain’s participants could increase their performance on such a large number of stimuli so quickly is particularly remarkable.

Beer and wine experiments

The advantage of studying beer and wine expertise is that there is a body of pre-existing wine experts who generally do not require further training. For this reason, most studies have used pre-existing experts to test hypotheses, rather than spending time and resources on training (Hopfer & Heymann, 2014; Hughson & Boakes, 2002a; LaTour, LaTour, & Feinstein, 2011; Melcher & Schooler, 1996).

Chollet et al. (2005) described the training of a beer sensory panel over a two-year period using beers with added flavours, such as caramel, banana, butter and lilac. The participants were trained for one hour per week to detect and identify the added flavours, as well as to assess the beers for characteristics such as bitterness and levels of alcohol, hops and malt flavours. However, this study did not compare different approaches.

LaTour et al. (2011) did compare perceptual and conceptual learning of wine. Like Chollet et al. (2005), LaTour et al. (2011) added elements to their stimuli to enhance certain elements. Those who received perceptual training smelled wines that included additives, similar to those described by Noble et al. (1987), and were asked to identify the strongest aroma in the wines. Conceptual training consisted of a PowerPoint presentation, containing the kind of information described in wine books, such as the steps involved in winemaking, information about wine regions and grapes and specific information about the type of wine used in the experiment (Zinfandel). The participants completed a recognition task whereby they initially tasted a wine and then were asked to identify it amongst four other samples of the same wine with various levels of added sweetener. The participants also described these wines in terms of "taste, smell, feel or related associations". The participants did not receive any feedback in terms of the use of descriptors and were thus not trained in this particular aspect of the experiment. Those who received perceptual training focused more on the sensory aspects of the wines in their descriptions, while those who received conceptual training were more likely to use terms from the conceptual training. The aim of the experiment was to study the verbal overshadowing effect (Melcher & Schooler, 1996), so while the authors provided the result of the statistical test comparing these two training groups and a no-training control group, they did not report a post-hoc test comparing the overall performance of the perceptual and conceptual groups. However, from examination of the means, those in the perceptual training group performed better than those in the conceptual training group. The authors concluded that perceptual training is more important that conceptual training, but conceptual training is still an important aspect of wine training.

Summary

Commercial wine training courses usually involve a large amount of conceptual training (e.g. how wines are made, wine regions, grapes, etc), while directed perceptual training usually only occurs in the types of courses that involve participants attending winetasting evenings with an instructor. In contrast, sensory panels undergo intensive perceptual training in order to be able to detect, identify and quantify all of the important attributes of a product. This training involves regular training sessions, whereby the panelists are presented with a small number of stimuli. They will usually attempt to identify them and rate their concentration as objectively as possible using line scales, although it is acknowledged that there is some degree of subjectivity to these ratings. They will then receive corrective feedback from the instructor, which is an important training mechanism (e.g. Findlay, Castura, & Lesschaeve, 2007; Findlay et al., 2006; Frank, Rybalsky, Brearton, & Mannea, 2011).

Most studies in academic literature have taken advantage of pre-existing experts in order to eliminate the requirement of training. However, those that have actively trained participants have generally involved intermixed trials and feedback. Some studies have included adulterated samples in order to highlight the most important aspects of a stimulus. In general, most of these studies use training procedures that make sense intuitively, but very few are evidence-based (e.g. LaTour et al., 2011).

While some studies have looked at the effects of training on discrimination or recognition, the purpose of this thesis is to examine which aspects of training help novices to overcome the problem of learning to identify elements in complex mixtures.

CHAPTER FOUR - GENERAL METHOD

Most of the experiments reported in subsequent chapters fall into two broad categories: experiments designed to test whether the participants could discriminate between samples (“discrimination experiments”) and experiments designed to train and test people in terms of identification of samples (“identification experiments”). The experiments within each category generally followed a similar formula. This chapter describes the general methods used for each type of experiment in order to avoid repetition throughout the thesis. Any differences between the methods described below and those used in an individual experiment are noted where that experiment is reported. This research was approved by the University of Sydney Human Research Ethics Committeee, reference 10235.

Participants

All participants in all experiments were first-year Psychology students and all participated for course credit. Almost all of the participants were between 18 and 21 years of age, although a small number of mature age students took part in some experiments. The age range of the participants, along with the number of participants and the number of females (from which the number of males can be determined), is reported for almost all experiments, except in the first experiments, where no mature age students took part and thus age was not recorded as it was essentially constant.

Advertisements for participants were placed through an online sign-up system available only to first-year Psychology students from the University of Sydney (SONA). In the advertisement, it was stated that all participants must be non-smokers. Furthermore, it was stated that English must be their first language, as language is an important component of the experiments. No participants took part in more than one experiment.

It was required that participants were wine novices and so all participants completed a slightly modified Australian Wine Knowledge Questionnaire (AWKQ: Hughson & Boakes, 2001). Almost all of the participants were considered to be novices on the basis of their answers to this questionnaire (scores of less than 4 out of a possible 8, with most participants scoring approximately 1 out of 8), but in some experiments there were at most two mature-age students who had had some exposure to wine during their life, but had not studied it and would thus be classified as “intermediate” (Melcher & Schooler, 1996). However, the removal of their data made no difference to the pattern of results for any experiment and they were thus included in the subsequent results.

Prior to attending any sessions, an e-mail was sent to all participants stating that they must not eat or drink anything in the immediate lead-up (defined as two hours) to each session that was likely to affect their senses of taste or smell, such as onions, coffee or chocolate.

The number of participants available from a subject pool and thus each experiment involved groups of up to 20 participants. The choice of this number was based on similar studies (e.g. Dwyer et al., 2004). In some cases, significant results were detected before 20 participants in each group were reached and the experiments were terminated early.

Materials

Olfaction experiments – odour mixtures

All olfactory samples were presented using the same equipment for all experiments. The olfactory stimuli were presented on make-up removal pads in airtight specimen containers with screw-on lids to prevent evaporation. Initial pilot experiments (see Appendix D) were conducted using squeezable sniff bottles (Décor plastic sauce bottles), but the participants did not follow instructions to only squeeze the bottles three times in an attempt to better smell the samples. Thus this alternative presentation method was used in order to have greater control over the participants' behaviours.

All participants in all experiments were first given a practice block of odours. These odours were flowers (specifically, rose) and pear (see Table 1). These odours were presented as single-odour solutions and were chosen because they are easily discriminated by most people with normal olfactory perception according to pilot testing. Thus, these trials also served as a simple screening task for those with olfactory impairment. Where a participant was incorrect for one or more trial, their data were considered for exclusion. No exclusions made any difference to the pattern of significant and non-significant results in any experiment and thus their data were included for all analyses.

In some experiments three odour mixtures were used for the training and test blocks: AX (Vanilla + Citral), BX (Melon + Citral) and CX (Banana + Citral). The choice of odorants, as well as the concentrations of each of the unique odorants, were chosen on the basis of pilot testing so that the unique odorants would be perceived as similar in intensity and would be difficult, but not impossible, to detect when mixed with a subjectively stronger lemon-like odour of Citral (Table 1). The concentration of Citral was the same for all of the mixtures. In creating these experiments, the relationship between odour quality (probability of identification) and intensity was considered (Atanasova et al., 2005; Frank, Goyert, & Hettinger, 2010; Olsson, 1994).

As in all the experiments reported here, the mixtures were presented at room temperature of approximately 22 degrees and a fan was used to exhaust odours from the laboratory. All mixtures samples were dissolved in water. Water was used as it would have been possible to use the same mixture for potential retronasal experiments which were considered, but ultimately never tested.

Table 1. Concentration and supplier of odours.

Odour name	Dilution Ratio	Supplier	Catalogue/ Order number
Rose (Flowers)	1:600	Quest	AP07749
Pear	1:600	Quest	AP06882
Citral	1:500	Perfume Manufacturers	N/A
Vanilla	1:600	Tastemaster	080820
Melon	1:725	Quest	AP05403
Banana	1:725	Quest	CD53114

Note: All odorants dissolved in filtered water.

Olfaction experiments – wine samples

In the olfactory experiments where wines were used as stimuli, the wines were soaked into the make-up removal pads and presented in the same type of specimen containers as the odour samples. Where visual differences were present between the wines, odourless dark-blue or black food colouring was added to the wine in order to mask these visual differences. The odourless nature of the food colouring was tested in pilot tests.

Flavour experiments

In some experiments participants were required to taste the wines, rather than just smelling them. In these experiments approximately 10mL of wine (measured according to levels in the cups) was presented in opaque, black, plastic cups to eliminate visual cues. This amount of wine was chosen to eliminate resampling, so that the tasting procedure was constant amongst participants. Standard ISO wine glasses were considered, but as many glasses were required for each testing session, and the fact that visual differences were still apparent in the glasses despite the use of food-colouring, a cheaper alternative was required. Prior to presentation each cup was sealed with aluminium foil to reduce evaporation and to reduce the possibility of excess odours in the testing laboratory. All wines were tasted within two hours of opening the bottles in order to reduce changes in sample quality, which was much more stringent than previous studies that have used bottles opened for up to seven hours (e.g. Hughson & Boakes, 2002a).

In all experiments, water was made available throughout and participants were encouraged to smell and drink the water between trials, which is a common technique used to refresh the palate (e.g. Labbe, Damevin, Vaccher, Morgenegg, & Martin, 2006).

Procedure

Discrimination experiments

The discrimination experiments reported in Chapters 5-8 and Appendices E to H were conducted using triangle tests. Each trial consisted of three samples, of which two were identical and one was different. Each stimulus served as a target and a foil the same number of times and the position of the target was counterbalanced across the experiment. The participant’s task was to smell all three samples – one at a time in a specified order – and respond by indicating which sample they thought was the unique sample in the set of three. The participants did this by circling A, B or C on a sheet of paper, corresponding with the first, second and third sample in each trial (which were labeled 1A, 1B and 1C for the first trial, 2A, 2B and 2C for the second trial, and so on). The samples were arranged into blocks consisting of six trials. The discrimination experiments were only run as olfactory experiments. The blocks in these experiments were not split over separate sessions.

Participants were required to wait for 5s between samples and for 1-min between trials. No feedback was given in any of the discrimination experiments. Once a participant had indicated their answer, they were instructed to put the samples aside and were not allowed to smell that particular set again.

A different discrimination procedure (the same-different task) was utilised for some experiments reported in the appendices, where the procedure for these experiments is also reported.

Identification experiments

In most of the identification experiments, participants were allocated to one of two groups. Participants were tested in groups of up to three at a time and all participants in a testing session were always allocated to the same group. Those in consecutive sessions were placed in different groups.

Some identification experiments were olfactory experiments, while later identification experiments were flavour experiments. All identification experiments required the participants to learn to identify three different stimuli, except for Experiment 17 and Experiment A1 in Appendix D, where two stimuli were used. The response options for the identification task were varied between experiments.

The stimuli were presented in blocks that usually consisted of 18 trials (six of each of the three stimuli), where each trial consisted of a single stimulus. A 45-s inter-trial interval was used within blocks and participants had up to 45s to respond for each stimulus. An experiment usually consisted of two training blocks and one test block, with either 5 min or more than 24 h between blocks, depending on the experiment.

In training blocks, participants were required to smell or taste each sample and then respond by stating which of the three stimuli they thought it was by using the given labels. Participants then immediately received feedback that told them whether their response was correct or incorrect and, if incorrect, what the correct response was. Participants were not allowed to smell or taste a stimulus again after they had received feedback. In the flavour experiments, participants could possibly gain visual cues from the wines as they spat them into the spittoon. To avoid this possibility, participants were required to spit the wine out after responding to each sample, which meant that they did not see the wine until they had completed the task related to that sample.

Essentially the same procedure was used in the test blocks as in the training blocks, except that participants no longer received feedback after each response.

The timing of stimuli, collection of responses and delivery of feedback were generally performed using a custom Inquisit script (Millisecond Software LLC, 2011). The order of stimuli was also randomised for each participant using the Inquisit script, with the constraint that no stimulus could appear more than three times in a row. Participants were directed to each sample in turn using unique, random, four-digit codes that were written on the containers for each odour or wine sample.

Analyses

Discrimination experiments

In most discrimination experiments, two different sets of stimuli were used for all participants: the odour mixtures described on page 55 and a set of three wines. For each set of stimuli, the total number of correct trials (out of

18) were compared to a chance performance of six items correct, in order to determine whether the participants could discriminate between the stimuli at a rate above chance. Each possible target and foil combination appeared three times in the experiment for each participant and was compared to chance, to determine whether any particular combination was likely to be problematic.

Where two sets of stimuli were used, the overall discrimination performance (out of 18) was compared using a paired-samples t-test, to determine whether one set of stimuli was significantly more discriminable than the other.

Identification experiments.

Where identification experiments involved two groups, the groups were compared in terms of overall performance (out of 18 trials) at each block, including both training and test blocks. Trend analyses were run across training blocks to determine whether each group was improving in the task as training continued. Assumptions were tested for all analyses and all data met all assumptions. Other analyses are outlined with each experiment.

CHAPTER FIVE – USING LABELS TO IDENTIFY ELEMENTS IN ODOUR MIXTURES

As described in previous chapters, appropriate cues are generally considered to be an essential ingredient for successful odour identification (Cain, 1979) and relatively short amounts of training are effective in improving odour identification performance (Desor & Beauchamp, 1974). Furthermore, as more odours are added to a mixture, identification performance on each of the odours decreases, despite training (Livermore & Laing, 1996). However, in a wine sample, there are hundreds of volatile odorants (Peynaud, 1987) and the specific odorants involved in certain flavours are therefore much more dilute than simple binary odorant mixtures. An initial pilot experiment is described in Appendix D, where the common element was not strong enough and the identification task was too easy.

The present experiments were concerned with the general question of how the effectiveness of training people to discriminate between similar odours might depend on the type of label used in the task when the target odour is relatively weak. This question is of interest in the context of experts' ability to identify odours that are highly confusable (difficult to discriminate) to non-experts. Thus, this study bears on the question of how the expertise displayed by, for example, flavourists or wine judges, can best be trained. Although our primary concern is with the discrimination of wines, in the first three experiments the task was to identify the unique element in each of three confusable odorant solutions over multiple exposures and to learn which label applies to which sample on the basis of feedback given after each trial. The samples were all mixed with a strong lemon (Citral) odour in order to make the solutions confusable. These experiments served as a preliminary study for the wine experiments described in subsequent chapters. The experiments described in this chapter were published (Russell & Boakes, 2011).

Experiment 1 – The Effect of Appropriate Labels vs Self-Generated Descriptions on Odour Identification

The main aim of Experiment 1 was to test the prediction that supplying an appropriate label, e.g. ‘Vanilla’ for a Vanilla + Citral compound, would enable participants to perform with greater accuracy in the identification task than participants required to generate their own labels or descriptions for the three compounds. In the present context we use the phrase ‘appropriate label’ to indicate the name of the added element provided by the supplier of the odorant/flavouring. The idiosyncratic label produced by a participant in the Description group in the present experiment could well be more ‘appropriate’ for that individual. The present procedure provides an objective measure of the degree to which a label is appropriate, namely, in terms of whether a label given to an odorant increases the average participant’s ability to identify that odorant. Thus, the increased level of odour identification obtained when Cain (1979) or Desor and Beauchamp (1974) provided names for the familiar odours in their studies can be taken as indicating that these names were indeed appropriate for their odours.

The main aim of Experiment 1 was to test the prediction that supplying an appropriate label, e.g. ‘Vanilla’ for a Vanilla + Citral compound, would enable participants to perform with greater accuracy in the identification task

than participants required to generate their own labels or descriptions for the three compounds.

Method

Participants. Sixteen first-year Psychology students (13 female) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. The materials used in this experiment were the odour mixtures described on page 55.

Procedure. Participants attended two sessions. An initial one hour session involved consent paperwork, screening, initial exposure to each odour and two training blocks. The next day participants returned for a 30-min session involving the test block, a follow up questionnaire and debriefing. Participants were given water to smell and drink throughout. Participants were given strict instructions to smell the samples in the order given by the experimenter.

On arrival for the first session all participants were given the following instructions:

The aim of this study is to determine which method of training provides the best results in terms of identifying the samples based on smell alone. All of the test odours are mixed with a strong lemon solution. Your goal is to ignore this lemon odour and try to identify the odours based on the unique smell in each.

The Description group (*n* = 8) were required to describe each odour however they liked and were then required to choose the appropriate description at each trial. These participants received the following additional instructions:

There are three different unique odours. You may describe each odour in any way you like. You are to attempt to identify the unique element in each sample according to your description and you will receive feedback after each trial. Your goal is to learn from this feedback.

The Appropriate group (*n* = 8) received the following additional instructions:

There are three different unique odours: Vanilla, Melon and Banana. You are to attempt to identify the unique element in each sample and you will receive feedback after each trial. Your goal is to learn from this feedback.

Screening. These trials served as a practice task to introduce participants to the procedure. Two trials of each screening odour were presented in random order. The task was to identify each trial as either ‘Flowers’ or ‘Pear’. All participants correctly identified these odours on all four trials, so no further results from these trials will be reported.

Initial exposure to the three main odours. All participants were given an initial exposure to each odour for 3s with a 50-s interval between exposures. These exposures allowed the Description group to create an initial description of the three odour samples, as well as allowing the Appropriate group to begin to identify each unique odour. Each group was told after each exposure which sample they had smelled: the Appropriate group was informed that the unique odours were ‘Vanilla’, ‘Melon’ and ‘Banana’, while for the Description group the odour samples were labelled as A, B and C, these letters corresponding to their descriptions. The Description group had access to their descriptions throughout the entire experiment.

Training. In each of the two training blocks, participants were presented with 18 trials comprising six trials of each odour. Three semi-random sequences were created with the constraints that no stimulus appeared more than twice in a row and the distribution of stimuli was approximately the same in the first and second half of each block. In a given testing session each participant was allocated a different sequence from the other participants in the session. For each trial participants were asked to identify the unique odour and, after writing down which of the three they believed it to be, they discovered the correct identification by examining the bottom of the container, where the identity of the added odour (e.g. ‘Vanilla’ or ‘A’) was written. As in the initial screening test, there was a 50-s interval between trials. After the first block of 18 training trials participants were given a 5-min break, during which they could read a magazine.

Test. Eighteen test trials were presented the next day, using the same procedure and sequence randomisation as in training, except that no feedback was given.

Follow-up questionnaire. After the final test trial, participants were given a questionnaire that included two open-ended questions; "Did you use a strategy to discriminate the odours and, if so, what was it?" and "Do you have any other comments to make about the experiment?"

Analysis. A 2x(2) ANOVA was employed to test for the main effects of group and training block and an interaction effect. Independent samples t-tests were used to test for differences between the groups in each block. Both groups were tested against chance (33%) at each block with t-tests for a single mean. All analyses were conducted using raw scores (the number of correct responses), but are reported below as percentages.

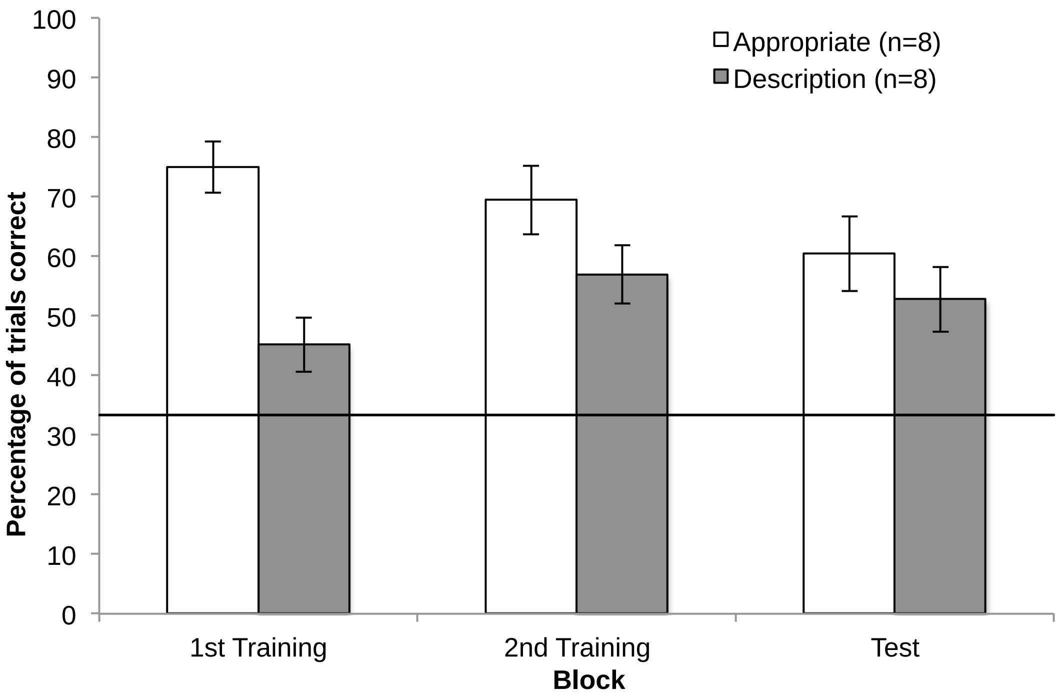

Figure 2. Mean (+/-SE) percentage of trials correct in each training and test block by group in Experiment 1.

Note: Feedback was given in the training blocks, but not in the test block. The horizontal line indicates chance level of responding (33%). Participants were given either appropriate labels (Appropriate group) or they produced their own descriptions of the odour compounds (Description group).

Results

Number correct. Figure 2 shows that, over the two training blocks, the Appropriate group correctly identified more odours than did the Description group, $F(1,14) = 10.93$, $p < 0.01$. There was no overall increase in performance from Block 1 to Block 2, $p > 0.10$, but there was a significant group by training block interaction, $F(1,14) = 10.75$, $p < 0.01$. From inspection of Figure 2, it is clear that this interaction reflects greater improvement in the Description group from first to second training block than in the Appropriate group. While the Description group improved with training, $t(7) = 2.13$, $p = 0.034$, this result was not significant using the Bonferroni correction to control familywise error rate. No change was detected in the Appropriate group, $p>0.05$. In the first training block the Appropriate group correctly identified more items than the Description group, $t(14) = 4.78$, $p < 0.001$, but the groups no longer differed in the second training block or the test block, both *p*s > 0.1. Mutually orthogonal contrasts revealed that both groups scored significantly more items correct than chance (33%) in each of the three blocks, with the smallest *t*-value of $t(7) = 2.62$, $p = 0.017$.

Descriptions. Participants in the Description group used a mixture of absolute terms (e.g. “fairy floss”, “cleaning fluid”, “sharp”), relative terms (e.g.

“more sweet than A”, “the weirdest smell”) and personal terms (e.g. “Aunt Lily’s clothes”, “my old diary”) to describe the odour compounds. A full list of descriptions is given in Appendix A. These self-generated descriptions were idiosyncratic. In particular, participants who used relative terms to describe the odours focused on a single dimension (such as sweetness) to describe the odours, trying to order them along this single dimension.

There was little agreement between participants on absolute terms and the majority of these could be classified as “far-misses”. Only three of the eight participants in the Description group identified one of the unique elements in their descriptions (either melon or banana but not vanilla) and none identified more than one.

Questionnaire. Reported strategies included: using the relative sweetness of the samples, attempting to differentiate the samples on perceived intensity of the lemon odour, learning one or two of the three samples and using a process of elimination to identify the third, or by remembering the previously presented smell to determine if the current odour was similar or different. Some participants also reported that the unique element was more apparent at different times during sniffing, most likely due to temporal coding (see Whalley, 2013). For example, one participant reported that the Melon sample was detected earlier in the sniff while another reported that Vanilla was an “aftertaste”.

Five participants reported that two of the odours were very similar, while the other was easier to discriminate. There was no agreement amongst participants about which of the odours were similar.

Discussion

The main finding from Experiment 1 was that participants given appropriate labels to choose from (Appropriate group) were better able to identify the highly confusable compounds than participants who generated their own labels, consistent with the finding by Parr et al. (2002) that verbal interference is associated with self-generated descriptors. While not a statistically significant increase, with training the latter (Description group) improved their performance up to a level similar to that of the Appropriate group. These results indicate that initially in such a task either being given labels is helpful, producing one's own description is harmful, or both. The experiments that follow addressed these questions.

Experiment 2 – The Effect of Appropriate vs Inappropriate Labels on Odour Identification

The performance of the Appropriate group in the previous experiment raised the question as to whether, as intended, the labels facilitated identification because each was appropriate for the specific compound to which it was attached or whether providing the labels facilitated performance because they were generally appropriate for this set of odours. Experiment 2 addressed this question by comparing an Appropriate group, for which conditions were essentially identical to those for the Appropriate group in Experiment 1, with an Inappropriate group, in which participants were also given the names for the unique elements in each odour (Vanilla, Melon or Banana), but these were attached to an inappropriate compound. For example, a participant in the Inappropriate group might be informed that the label for the Vanilla + Citral sample was Melon.

Method

Participants. Twenty-seven first-year Psychology students (18 female) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. As for Experiment 1.

Procedure. Screening, training and follow-up questionnaire procedures were the same as in Experiment 1, except for the following aspects. First, the test block was reduced to twelve trials and was conducted 5 min after the conclusion of the second training task; thus, participants attended a single one hour session. Second, participants were prompted to smell each sample in turn by a signal from a computer that also presented the 3-label choice on each trial; participants used a computer keyboard to select one of the labels and received immediate feedback as to which was the correct label on the computer screen.

The third, and most important, difference in procedure was that for participants in the Inappropriate group (n = 14), an item was considered "correct" if they identified the item according to the training allocation. As in the example above, if the Vanilla + Citral compound was presented, but previous feedback had informed the participant that the sample was 'Melon', then the correct response would be 'Melon'. The allocation of labels for this group was counterbalanced so that, for example, for approximately half these participants ‘Melon’ was the label allocated to Vanilla + Citral and for the other half ‘Banana’ was the allocated label. Apart from the minor differences reported above, conditions for the Appropriate group ($n = 13$) were exactly as in Experiment 1. See page 59 for general information about the procedure used in this experiment.

Results

As suggested by Figure 3, during training the percentage of correct trials for the Appropriate group was greater than that for the Inappropriate group in both blocks of trials and both groups showed slight improvement from the first to the second training block; thus, ANOVA found a main effect of group, $F(1,25) = 13.44$, $p < 0.01$, a main effect of training, $F(1,25) = 7.53$, $p < 0.01$, but no interaction, $p > 0.1$. The groups differed in the first training block, $t(25) = 4.96$, $p < 0.001$, the second training block, $t(25) = 4.43$, $p < 0.001$, and in the test block, $t(25) = 2.73$, $p = 0.012$.

The Appropriate group scored significantly higher than chance (33%) in all blocks, with the smallest *t*-value of $t(12) = 6.35$, $p < 0.001$. The Inappropriate group scored significantly higher than chance in the second training block, $t(13) = 3.69$, $p = 0.003$, and in the test block $t(13) = 2.42$, $p = 0.031$, but not in the first training block, $p > 0.10$.

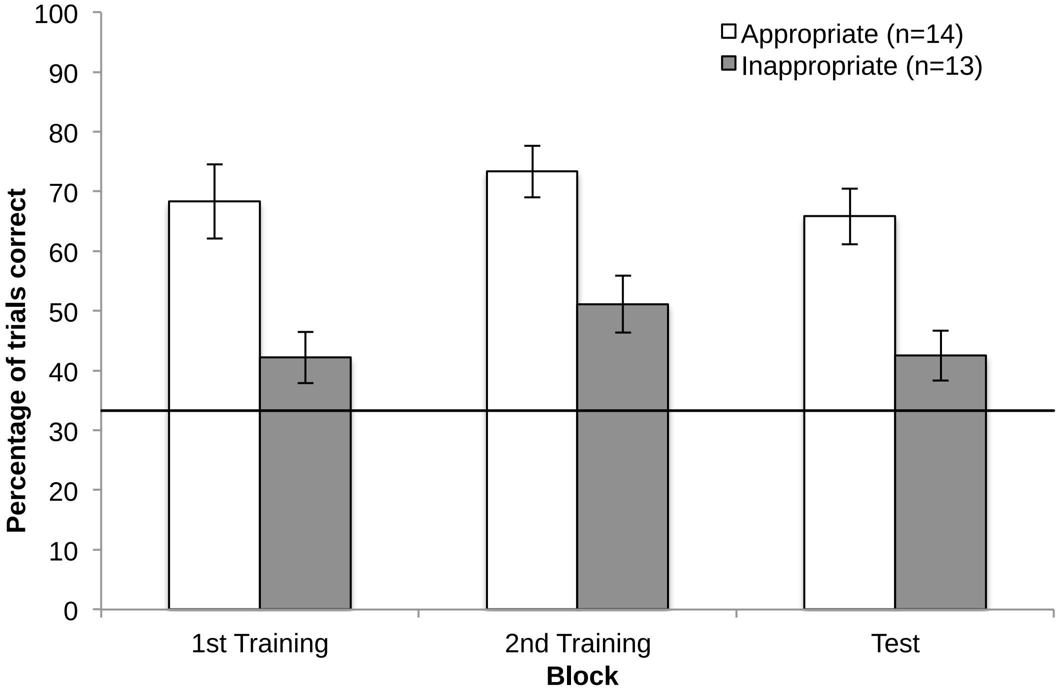

Figure 3. Mean (+/-SE) percentage of trials correct in each training and test block by group in Experiment 2.

Note: Feedback was given in the training blocks, but not in the test block. The horizontal line indicates chance level of responding (33%). Both groups were given the same set of labels; in the Appropriate group, these were applied to the appropriate odour compound (as in Experiment 1), but they were allocated to an inappropriate compound in the Inappropriate group.

Discussion

As expected, the Appropriate group performed at a higher level than the Inappropriate group. Nonetheless, participants in the latter group learned to apply the inappropriate labels, performing above chance, although not to the same level as the Appropriate group. It is possible that the presence of the correct labels in the list of choices may have made the task easier for the

Inappropriate group, in that these participants could have learned to answer (for example) Melon when they smelt Vanilla and this was indeed the case for two participants who identified this relationship in the follow-up questionnaire. However, if this were also true for many more participants in the Inappropriate group, this group's performance should have been similar to that of the Appropriate group.

Experiment 3 – The Effect of Appropriate vs Irrelevant Labels on Odour Identification

Given that participants in Experiment 1 learned to apply their own labels to odours and that participants in Experiment 2 learned to give inappropriate labels to the odour compounds, the question then arises as to whether people can learn under the present conditions to apply labels that are unrelated to any of the odours in the set. If training on inappropriate labels is detrimental to odour identification performance, then training participants using labels that are not appropriate to them, such as those used amongst experts to describe wines, is also likely to be detrimental.

Method

Participants. Twenty-nine first-year Psychology students (18 female) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. As for Experiments 1 and 2.

Procedure. The Appropriate group (*n* = 13) received exactly the same conditions as the Appropriate group in Experiment 2. The Irrelevant group (*n* = 16) were given the same samples to smell as the Appropriate group, but were instead required to identify them as Pepper, Cherry or Blackcurrant, with the allocation of these labels to odour compounds counterbalanced across participants in this group. Apart from the different labels given to the Irrelevant group, the procedure was identical to that used in Experiment 2. See page 59 for a more information about the procedure used in this experiment.

Results

As seen in Figure 4, the Appropriate group identified more items correctly than the Irrelevant group in both training blocks and in the test block. An ANOVA applied to the training data found a main effect of group, $F(1,27) = 13.29$, $p < 0.01$, but neither a main effect of training nor a group by training interaction, both *F*s < 1. Tests for simple effects confirmed that scores were higher in the Appropriate than in the Irrelevant group in all three blocks, with the smallest *t*-value of $t(27) = 4.67$, $p < 0.01$.

Scores for the Appropriate group were higher than chance (33%) in all blocks, with the smallest *t*-value of $t(12) = 10.75$, $p < 0.001$. Those for the Irrelevant group were higher than chance in both training blocks, with the smallest *t*-value of $t(15) = 4.21$, $p = 0.001$, but not in the test block, $t(15) = 1.88$, $p = 0.08$.

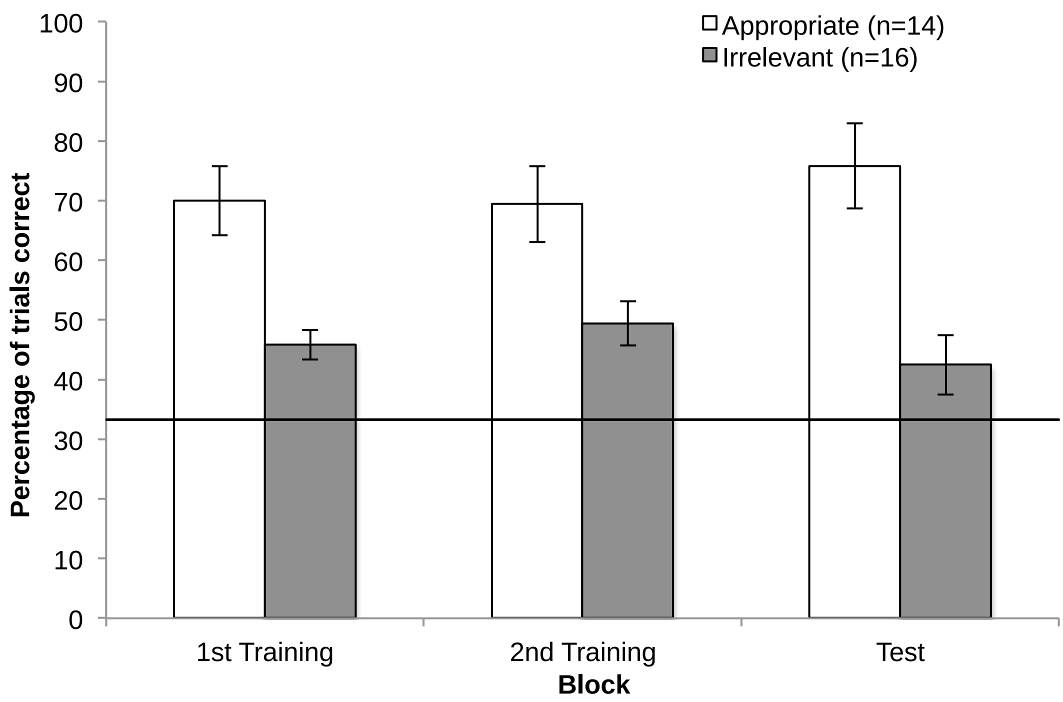

Figure 4. Mean (+/-SE) percentage of trials correct in each training and test block by group in Experiment 3.

Note: Feedback was given in the training blocks, but not in the test block. The horizontal line indicates chance level of responding (33%). Conditions for the Appropriate group were as in Experiments 1 and 2, while the Irrelevant group were given an unrelated set of labels for the odour compounds.

Discussion

These results indicate that participants can learn to apply arbitrary labels to the compounds under the present conditions as long as feedback is maintained, but not as well as when appropriate labels are given. However, when feedback was discontinued, as in the test block, performance using arbitrary labels was not sustained.

In all three of these initial experiments participants given appropriate labels to apply to the odour compounds performed more accurately than participants in the other conditions where labels were either self-generated, inappropriate, or unrelated, suggesting that the method employed here can serve as a test of the validity of the labels provided by the manufacturers of the odorants and flavours (O'Neill, Nicklaus, & Sauvageot, 2003). After all, these substances were not *real* vanilla, melon and banana, but substances judged by flavourists and others to have odours very closely resembling those of the real substances. Although the validity of these particular labels was never in doubt, in other contexts – such as descriptions of the complex odours of wines – the validity of labels supplied by an expert or manufacturer might be more open to question.

Summary

In addressing the question of why being given an appropriate label helped participants in the first three experiments, it should be noted that the task required two steps in the identification of the added odour on each trial: first, detection of this odour against the strong background odour of Citral and, second, choosing which of the three labels should be applied to the added odour. If the major difficulty lay in the first step, then providing a set of labels that was generally appropriate, although not paired with the appropriate particular odour – as in the Inappropriate condition of Experiment 2 – should support performance almost as good as in the Appropriate condition and better than when provided with a set of unrelated labels, as in the Irrelevant condition of Experiment 3. Although the present study did not include a within-experiment comparison of the Inappropriate and Irrelevant conditions, inspection of Figures 2 and 3 suggests that such an experiment would need considerable statistical power to detect any difference. Consequently, one may tentatively conclude that the main benefit of being given appropriate labels stems from improvement in identification rather than detection of the added odour. Thus, the top-down effect of the labels in the present experiment involving highly confusable odour compounds appears to be similar to that found in experiments where detection is not a problem (Cain, 1979).

Whereas we have examined the effects of labelling on odour identification, related effects have been found in recognition memory. Thus, performance in a recognition test can be enhanced when an odour has been labelled when first encountered in the experimental setting (Lehrner, 1993). Labelling effects have also been found in hedonic responses to ‘ambiguous’ odours where, for example, labelling a combination of isovaleric and butyric acids as ‘parmesan cheese’ evokes a very different reaction from that evoked when the same odour is described as ‘vomit’ (Herz & von Clef, 2001).

Top-down processes involved in identification have been studied in other sensory modalities, most notably using visual stimuli. Thus, a popular textbook example has been a picture of what at first seems a meaningless set of dark blobs on a light background; when told that it is a photograph of a Dalmatian dog in a park on sunny day, it becomes much easier to ‘see’ the dog in the centre of the picture (Lindsay & Norman, 1977). In the present sense this indicates that ‘Dalmatian dog’ is an appropriate label for the picture. A related example from visual perception is where a label, e.g. ‘young girl’ or ‘saxophone player’, biases perception towards one or other interpretation of an ambiguous figure. In the latter case it has been argued that the label influences attention to different parts of a complex visual pattern (Cavanagh, 1999). In contrast, top down effects in auditory perception have been explained as a result of pre-existing schemata that a label evokes (Bey & McAdams, 2002; Bregman, 1994). In relation to odour perception it is hard to see how a process analogous to visual attention could be operating, whereas the possibility is more plausible that a label evokes a schema, i.e. the representation of a particular pattern of olfactory stimulation, that can be matched against the incoming pattern of olfactory stimulation. Exactly how top-down processes operate in olfaction deserves far greater examination than has been given up to now. However, such questions go beyond the scope of the present study.

The other main finding from these experiments was that providing feedback in the kind of multiple choice tasks used here is not a very effective training method, as demonstrated by the results for the Inappropriate (Experiment 2) and Irrelevant groups (Experiment 3). It appears to be particularly difficult to train people to label an odour in a particular way when they do not already have a connection between the odour and the required label. Interestingly, the Description group in Experiment 1 was the only one that showed signs of improvement and this may indicate that, even though many of the descriptions could be classified as far-misses when compared to the veridical odour names, the participants could use a term that had a long-standing connection with the odour *for them*, even though the label or description may not have a connection to the odour for anyone else. This is in keeping with results reported by Lehrner, Gluck, and Laska (1999) and Parr et al. (2002) that showed that consistent use of a label was more important for odour-recognition accuracy than employment of the so-called veridical name of the odorant.

CHAPTER SIX – USING LABELS TO IDENTIFY ELEMENTS IN WINE SAMPLES

The experiments in Chapter 5 illustrated the importance of the labels used during an identification task. The next step was to determine which labels were appropriate for wines.

The labels used in the previous chapter described the unique element of each odour mixture. With wine samples, it is possible to describe a unique element of wine, such as the peppery element that is typical of wines made with the Shiraz/Syrah grape. It is also possible to describe the wine using a more holistic, non-sensory label, such as the word Shiraz.

Wine experts are able to identify major attributes of wines and to describe wines in more accurate detail than novices (Lehrer, 1975). Descriptions produced by wine experts are more likely to co-vary with grape type than those produced by novices (Solomon, 1997) and experts are better than novices at sorting wines into appropriate categories (Ballester et al., 2008). Furthermore, experts' descriptions are more likely to contain concrete (as opposed to abstract) terms for wines than those of non-experts (Lawless, 1984). Solomon (1990), for example, found that novices were unable to match wines to descriptions written by experts, suggesting that the descriptors may be meaningless to novices (see also Noble et al., 1987). Previous experiments of this kind have not involved multiple trials with feedback, as in the present study. Experts do not seem to have superior olfactory abilities to novices (Parr et al., 2004), suggesting that knowledge (Hughson & Boakes, 2002a, 2002b) and training (Gawel, 1997) are important factors. The question addressed in experiments reported within this chapter was whether the labels for wine samples that an expert has provided can help novices to identify a wine on the basis of its odour in the same way that providing appropriate labels improved performance in Experiments 1, 2 and 3. The first step towards answering this question was to find a set of three wines with similar discriminability between their odours to that of the odour samples used in Experiments 1–3. The first two experiments described in this chapter were published (Russell & Boakes, 2011).

Experiment 4 – Testing Discrimination of a Set of Red Wines

Experiment 4 used a triangle test procedure to compare the discriminability of the three odour mixtures from the previous experiments with that of the odours of three distinctive red wines.

Method

Participants. Fifteen first-year Psychology students (12 females, aged 18-27, *M* = 19.4, *SD* = 2.5) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. The odour samples were the same as those used in the experiments in Chapter 5. The red wines were a Yalumba ‘‘Galway Vintage’’ Shiraz (2007), a Yering Station ‘‘Mr. Frog’’ Pinot Noir (2007) and a Donelli Reggiano Lambrusco (non-vintage). See pages 55 and 57 for information about how the samples were prepared and presented.

Procedure. The participants' task was to identify the unique odour or wine in each trial (triangle test). The procedure used for this experiment is described on page 58. Each participant completed six blocks of six trials – three blocks of wine trials alternating with three of odour sample trials – resulting in a total of 18 trials of each kind. Seven participants started with a wine block and eight started with an odour sample block.

Results and discussion

No significant difference in performance was detected between the two types of stimuli, all *t*-values < 1. The average number of correctly identified wine stimuli was 10.0 ($SD = 1.96$), while the average number of correctly identified odour sample trials was 10.1 ($SD = 2.46$). Performance for both wine and odour samples was above the chance score of six out of 18 stimuli correct, $t(14) = 7.89$, $p < 0.001$, and $t(14) = 6.40$, $p < 0.001$, respectively. Finally, all samples within a stimulus type appeared to be equally discriminable from each other; that is, errors were approximately equally distributed amongst all target and foil combinations for both odour and wine samples (Table 2 and Table 3).

Table 2. Summary of errors by target and foil combination for odour samples in Experiment 4.

Target	Vanilla		Melon		Banana	
Foils	Melon	Banana	Vanilla	Banana	Vanilla	Melon
Mean errors (/6)	1.73	1.50	1.50	2.00	1.36	1.27

Table 3. Summary of errors by target and foil combination for wine samples in Experiment 4.

Target	Shiraz		Pinot Noir		Lambrusco	
Foils	Pinot Noir	Lambrusco	Shiraz	Lambrusco	Shiraz	Pinot Noir
Mean errors (/6)	1.67	1.85	1.29	1.54	1.50	1.67

These results indicate that, while participants can discriminate between the chosen wine samples at a level above chance, this is a difficult task and the level of discriminability is comparable to that of the odour samples. The question addressed by the next experiment is whether giving labels to these wine samples will allow participants to identify them, given the same training procedure as in Experiments 1–3.

Experiment 5 – The Effect of Grape Name vs Descriptor Labels on Identification of Wine Odours

This experiment compared the effectiveness of two types of labels: grape name (Shiraz, Pinot Noir and Lambrusco) and descriptors (spice and chocolate, black cherry and gamey, floral and raspberry, respectively). The descriptors were taken from the winemaker's tasting notes for each wine. The descriptors were chosen so that they were present only in the tasting note of that wine and were not present in the tasting notes of the other wines. It was predicted that these should aid identification more effectively than the grape names, since the latter have little meaning to people without any knowledge of wine and thus should be no more effective than the Irrelevant labels used in Experiment 3.

Method

Participants. Forty first-year Psychology students (24 females) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. The wine samples were the same as those used in Experiment 4, arranged into three 18-trial blocks, with stimulus sequences organized as in Experiments 1–3. See page 57 for information about how the samples were prepared and presented.

Procedure. The procedure was the same as that used in Experiment 3, whereby participants were instructed to smell each wine sample in order to identify it using either its grape name (Grape group, n = 21) or its description (Description group, n = 19), depending on the group to which they had been allocated. See page 59 for general information about the procedure used in the experiment.

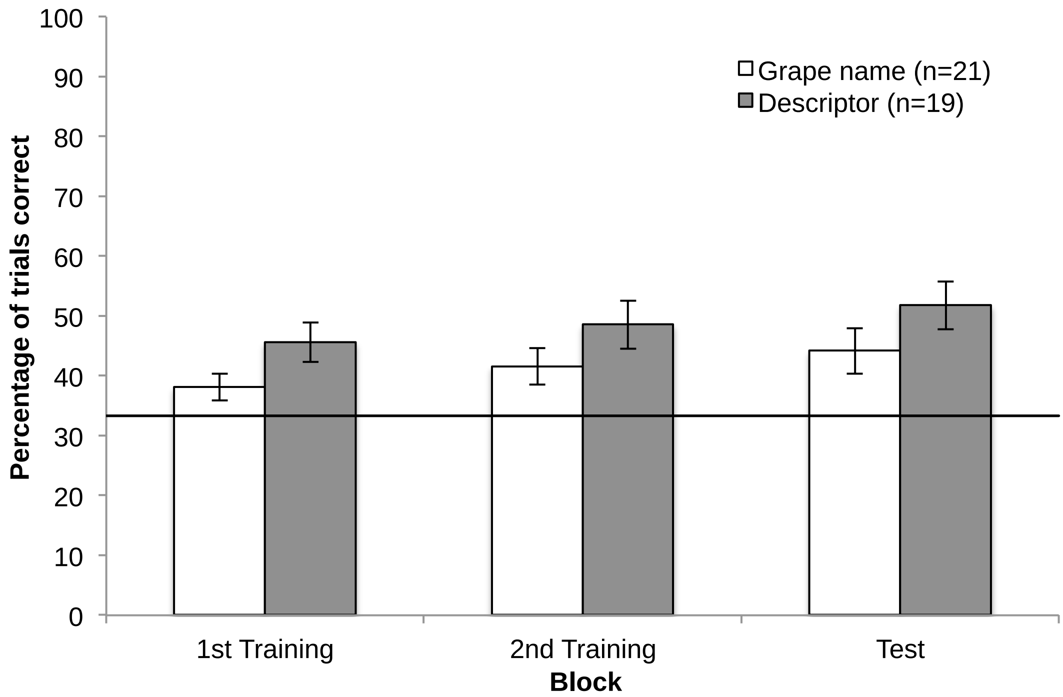

Figure 5. Mean (+/- SE) percentage of trials correct in each training and test block by group in Experiment 5.

Note: Feedback was given in the training blocks, but not the test block. The horizontal line indicates chance level of responding (33%). Participants were either given grape names (Grape name group) or two-word wine descriptions (Descriptor group) to match to the wine samples.

Results and discussion

Percent correct scores are shown in Figure 5. A 2x(2) ANOVA indicated a significant main effect for group ($F(1, 38) = 4.13$, $p < 0.05$) but not for training ($F(1, 38) = 1.38$, $p > 0.05$) nor for the group x training block interaction ($F(1, 38) = 0.01$, $p > 0.05$).

When the groups were compared at each of the training and test blocks, no significant differences between the two groups were detected, largest $t(38) = 1.93$, $p = 0.061$ for the first training block. Both groups scored significantly above chance (six trials correct) in all blocks (with the smallest *t*-value of $t(20) = 2.15$, $p = 0.044$ for the Grape group in the first training block). When correcting for multiple comparisons using the Bonferroni method, this particular block was the only one in which performance was approximately at chance level. The significant group main effect indicates that the descriptions were easier for the participants to use than the grape names, with the Description group scoring 1.3 more items correct during training compared to the Grape group. This is not entirely surprising given that grape names carry little information for novices compared to descriptors that also apply to non-wine samples. The nonsignificant differences between the groups at each block could be due to a lack of statistical power, particularly for the first training block. That both groups performed significantly above chance in the second training and test blocks indicates that these labels were not entirely useless for the participants. However, they did not facilitate identification performance to the levels seen in the Appropriate groups in Experiments 1–3.

This finding suggests that novices can learn to identify wines given a relatively short period of training using either labels that refer to unique elements in the wines or using more holistic labels. This does not necessarily contradict the findings of Solomon (1990) that novices cannot match descriptions written by experts back to the original wines because the present experiment involved training, whereas Solomon's participants did not receive multiple exposures to each sample and were not trained. However, the finding that novices can use both descriptors (e.g. pepper) and grape names is interesting. The finding suggests that novices are able to use information provided by others if it is appropriate to the wine (the descriptions), although not to the extent seen in Experiments 1-3 with the appropriate labels in the odour mixtures. This difference could possibly be attributed to the wines being much more complex stimuli.

Solomon (1990) also found that novices cannot match their own descriptions back to wine samples without training. This could be because the novices were unable to detect these elements in the wines or because the descriptions produced by novices were generally of poor quality (Lawless, 1984; Lehrer, 1975; Solomon, 1997). The results from Experiment 5 suggest that the former explanation may not be the case.

Thus, the following experiments were devised to test whether the participants could match their own descriptions to the wine samples given training. Experiments 6 and 7 were conducted after the experiments in later chapters and the set of wines used was different than those used for Experiments 4 and 5. Further details are reported in Chapter 7 and Appendices E and F. The wines were also changed in order to find a set of wines with fewer non-olfactory differences (such as the carbonation of the Lambrusco) that could be used in subsequent flavour experiments.

Experiment 6 – Testing Discrimination of an Alternate Set of Red Wines

As different wines were used to those in Experiments 4 and 5, the first step was to ensure that the new wine set was discriminable. A secondary aim was to search for a set of wines that was more discriminable than those used in Experiments 4 and 5. This was achieved using triangle tests, as was the case in Experiment 4.

Method

Participants. Sixteen first-year Psychology students (12 females, aged 18 to 22, *M* = 18.6, *SD* = 1.1) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. The red wines were a Lock and Key Shiraz (2009), a Yering Station ‘‘Mr. Frog’’ Pinot Noir (2009) and a McWilliam’s “Hanwood” Cabernet Sauvignon (2009). See page 57 for information on how the stimuli were prepared and presented. Unlike Experiment 4, no odour samples were used, as the aim of this experiment was to ensure that the wines were discriminable at a level above chance.

The experimenter (AR) timed the intertrial intervals and the participants recorded their responses by circling A, B or C (corresponding to the first-, second- and third-sniffed samples) on a sheet of paper. No feedback was given.

Procedure. General information about the procedure used in this experiment is outlined on page 58. To summarise, participants were presented with 18 trials (three blocks of six trials), with each trial consisting of three samples. Two of these samples were the same, while one was different. The participants were asked to smell each odour in order and determine, for each set of three, which wine was the unique sample. No feedback was given throughout.

Results and discussion

The average number of correct wine stimuli was 9.19 (SD = 2.83), which was significantly higher than chance performance of six out of 18 items correct $t(15) = 4.50$, $p < 0.001$. As each target and foil combination was presented three times, the chance error performance for each combination was two incorrect items. Results for each target and foil combination are shown in Table 4. Performance was not significantly different to chance (one trial correct from three trials of each target/foil combination) when the target was Shiraz and the foils were Pinot Noir or when the target was Cabernet Sauvignon and the foils were Pinot Noir ($t(15) = 1.57$, $p = 0.138$). However, all other combinations were statistically significantly higher than chance and the wines were thus considered to be discriminable by the participants.

Table 4. Summary of errors by target and foil combination in Experiment 6.

Target	Shiraz		Pinot Noir		Cabernet Sauvignon	
Foils	Pinot Noir	Cabernet Sauvignon	Shiraz	Cabernet Sauvignon	Shiraz	Pinot Noir
Mean errors (/6)	2.00	1.25	1.44	1.44	1.06	1.63

These results indicate that the participants can discriminate between these wine samples without the task being too easy. Thus these wines were considered acceptable for Experiment 7.

Experiment 7 – Can Novices Learn to Identify Wines Using Their Own Labels, Given Testing?

The aim of Experiment 7 was to test whether novices can use their own descriptions to identify wines above chance when given training, as seen with the odour samples in Experiment 1. Given that novices are generally suspected of producing poor descriptions for wines, a secondary question of interest was whether directly comparing and contrasting the wines during the description-generation phase of the experiment would help novices produce more useful descriptions. The rationale behind this secondary hypothesis was that, in perceptual learning experiments using visual stimuli, a masking screen is used between stimuli to make it more difficult to detect the differences between the stimuli. While the presentation of the odour stimuli in this experiment was necessarily serial, allowing the participants to smell two different wine samples with a short break between them may have allowed them to more easily identify the relevant unique elements in each wine and thus choose labels that are more useful for the task.

Thus, the main aim of the experiment were to determine whether novices could use their own descriptions of the wines in order to facilitate performance in a manner similar to that seen in Experiment 1. The secondary aim was to determine if direct comparisons during the description stage helped participants to determine the unique features of each wine and thus produce more useful descriptions compared to those who did not receive direct comparisons, which would be evident in better performance for those participants.

Method

Participants. Thirty-eight first-year Psychology students (32 females, aged 18-25, *M* = 18.7, *SD* = 1.5) participated in the experiment for course credit. See page 53 for general information about the participants.

Apparatus. The wines were the same as those used in Experiment 6: a Shiraz, a Cabernet Sauvignon and a Pinot Noir. Participants were given a sheet of paper on which to create their descriptions for each of the three wines. A copy of the Wine Aroma Wheel (Noble et al., 1987) was provided (see Chapter 3). See page 57 for information about how the samples were prepared and presented.

Procedure.

Description task. Participants attended a 1.5 hour session and were randomly allocated to the “comparison” or “non-comparison” groups based on their time of arrival. Unlike the general identification experiment procedure outlined on page 59, participants received four blocks of wines. The first block was the description phase, where participants were either given pairs of wines to describe, compare and contrast (comparison group, *n* = 19), or individual wines to describe (non-comparison group, *n* = 19). In both groups, participants were instructed to use words on the Wine Wheel (Noble et al., 1987) and were not allowed to use any other words. Participants in both groups were instructed to use words for each wine (referred to as Wine A, Wine B and Wine C) that did not apply to the other samples. Participants received two exposures to each wine or wine pair in order to facilitate these descriptions. The participants received only two exposures due to time constraints. Participants were also allowed to alter their descriptions during the training blocks. Thus, while the both groups were able to compare wines from one trial to the next, participants in the “comparison” group were able to directly compare the wines at the beginning of the experiment, whereas those in the “non-comparison” group compared them over a 45-s interval.

Identification task. The next two blocks were training blocks, during which participants smelled wine samples in a quasi-randomised order (with the limitation that each wine was not presented more than three times in a row), one at a time, and attempted to match their descriptions to each wine. All participants received feedback after their response, which indicated which response was correct (in the form of Wine A, Wine B and Wine C). Participants had their descriptions in front of them at all times and were allowed to alter their descriptions on the basis of this feedback during these two blocks.

The final block was a test block, during which participants received no feedback and were not allowed to alter their descriptions, but were still allowed to look at their descriptions.

Block 1 (the exposure block) consisted of six trials (two of each wine for the non-comparison group) or six pairs (each possible pair twice for the comparison group). Blocks 2-4 consisted of 18 trials – six of each wine. The response options in blocks 2-4 were “A”, “B” or “C”, which corresponded to their written descriptions.

Results

Accuracy of descriptions. The descriptions written by the participants for each wine are presented in Table 5. It is noted that, in some cases, the participants did not follow the instructions faithfully and used terms that are not on the Wine Aroma Wheel (e.g. yuck, sweet, etc). Their descriptions were compared to the winemakers' tasting notes and matching terms were identified (referred to below as "consistent" descriptions). The following were the unique terms in the winemakers' tasting notes: black pepper and spicy (Shiraz), blackcurrant (or cassis) and blueberry (Cabernet Sauvignon) and earthy and cherry (Pinot Noir). None of these terms appeared in the winemakers' tasting notes for the other wines.

In general, only some of the participants used the same unique terms as the winemakers for the wines. Upon smelling the Shiraz, seven of the 38 participants used the unique terms from the Shiraz winemaker's tasting notes. However, three of these seven participants also used terms from the Pinot Noir winemaker's tasting note, while another four only used words from that tasting note and five more participants only used words from the Cabernet Sauvignon tasting note.

When describing the Cabernet Sauvignon, eight participants used terms consistent with the winemaker's terms, four used words consistent with the winemaker's description of the Shiraz and three used words consistent for the Pinot Noir. No participants used words consistent for two of the wines in the same description.

For the Pinot Noir, eight participants used cherry or earthy (matching the winemakers' description), three used Shiraz terms and five used terms

Cabernet Sauvignon terms. Two participants also used terms that overlapped between two of the wines.

Table 5. Descriptions produced for the wine samples in Experiment 7.

Pp	Lock and Key Shiraz 2009	McWilliam's Hanwood Cabernet Sauvignon 2009	Yering Station Mr Frog Pinot Noir 2009
1	Preserved fruit, sherry, pinewood	Sweet, tropical fruit	Yeasty, bakers yeast, woody, cork
2	Spicy, soy sauce, yeasty, black olive	Fruity, raisin, tea, grape, slightly woody, cork	Floral, cherry, honey, plum, allspice
3	Dusty, earthy, raisin	Spicy, anise	Fruity, berry, sweet caramel, blueberry
4	Caramel, clove, mint	Cinnamon, ginger, blackberry	Butterscotch, vanilla, chocolate, maple syrup, mushroom, honey
5	Black pepper, earthy, soy sauce	Dried tobacco, woody, cork, vegemite, black olive, bell pepper, hemp	Fruity, fresh, caramel, sweet
6	Grape, cedar, oxidized, resinous, berry, dried	Citrus, pine, fresh, earthy, dried	Melon, mint, citrus
7	Woody, resinous, cork	Fruity, berry, grape, blackberry	Microbiological, yeasty, bakers yeast
8	Spicy, earthy, cheese, bread, blue cheese	Berry, cherry, plum	Raspberry, blackberry, dusty
9	Smokey, cork, strong, woody, earthy, spicy, black pepper	Smoky, plum	Blackcurrant, fruity, sweeter, jasmine, sherry, tree fruit
10	Yeasty, lactic, oxidised, preserved fruit, sweet	Yeasty, baker's yeast, tropical fruit	Berry, woody
11	Oak, cork, smokey	Plum, blackcurrant, pear	Prune, fig
12	Spicy, cola	Fruity, melon, cantaloupe, flowery	Dried hay, soil
13	Dried tea, berry, vanilla, butterscotch, woody, strawberry, green olive, jasmine	Apple, pecan, nutty, sherry, blueberry	Basil, banana, cork, mint, cedar, resinous
14	Melon, floral, raisin	Cherry, berry, earthy, resinous	Tea, mint, citrus, lime
15	Dense, earthy, strong	Woody, nutty, less strong (than A)	Fruity, berry, sweet, weak/light, floral?
16	Citrus, preserved fruit, nutty	Berry, strong, sweet	Sweet, fruity, sour, floral, weak
17	Woody, resinous, yuck	Sweet, molasses, strong	Floral, nice
18	Spicy, cinnamon, rose, dusty	Mint, butter, orange	Licorice, fruity

Table 5. Descriptions produced for the wine samples in Experiment 7 (cont).

Pp	Lock and Key Shiraz 2009	McWilliam's Hanwood Cabernet Sauvignon 2009	Yering Station Mr Frog Pinot Noir 2009
19	Woody, toasted, berry, sherry	Spicy, dried, hay, earthy	Floral, nutty, oxidised
20	Sharp blackberry scent, almost overpowering	Rounder, smooth pear scent	Distinct earthy, woody smell, similar to a. Damp forest.
21	Ginger, raisin	Berry, blackcurrant	Earthy, dusty
22	Oak, cherry	Cork, peach, plum	Violet, blackberry, cedar
23	Yogurt, lactic, microbiological, more acidic	Blackberry, sherry, lactic, fruity, more fruity, not as strong as A	Redwood, oak, more musky, not very strong
24	Fruity - preserved fig, honey, berry	Tree fruit - plum, cut green grass	Woody, resinous, pine, dusty, walnut
25	Fruit, dates, dried, sherry	Blackcurrant, plum, fresh, strong fruit	Sherry, oxidised
26	Berry, plum, citrus, fresh, basil	Prune, preserved fruit, fresh, sweet, lemon, burnt, spicy, anise	Smokey, earthy, pine, rosemary, apple, honey, sweet
27	Smokey, grape	Plum, sherry	Rosemary, dusty
28	Floral, cassis	Woody, sherry	Earthy, smokey, spicy
29	Appley, cranberry	Apricot, nectarine	Peachy, cherry, sweeter
30	Fruity, berry, cranberry	Fruity, raspberry	Fruity, tree fruit, plum
31	Cherry, woody, earthy, no perfume, bitter	Grapey, sweeter than A, juicy, slight perfume	Strong, spice, fruity, burns slightly, perfumed, deep
32	Pine, clove, rose, finger, fig, tang, floral	Date, hazelnut, rosemary, plum, aged cheese, spicey	Berry, black peppers, orange, blackcurrant, fainter
33	Cinnamon, berry, very fruity, honey, little bit spicy, floral, dried	Smokey, dried, floral, apple, cranberry, tree fruit	Butterscotch, raisin, honey, fresh, spicy, tropical fruit, apple
34	Berry, grape, apple, eucalyptus	Violet, Cranberry, blackcurrant, apple, sweet, grapefruit	Grape, apple, prune, watermelon
35	Woody, cork, pine	Herbs/leaves, dried, hay	Earthy, soy sauce, dusty
36	Berry (grape), blackberry	Black olive, grape	Coffee, anise
37	Blackcurrant, plum, redwood, cedar	Apple, oak, prune	Raisin, smokey
38	Grape, blackberry, blackcurrant, violet, apricot	Raisin, sherry, dates	Plum, cherry, anise, apple

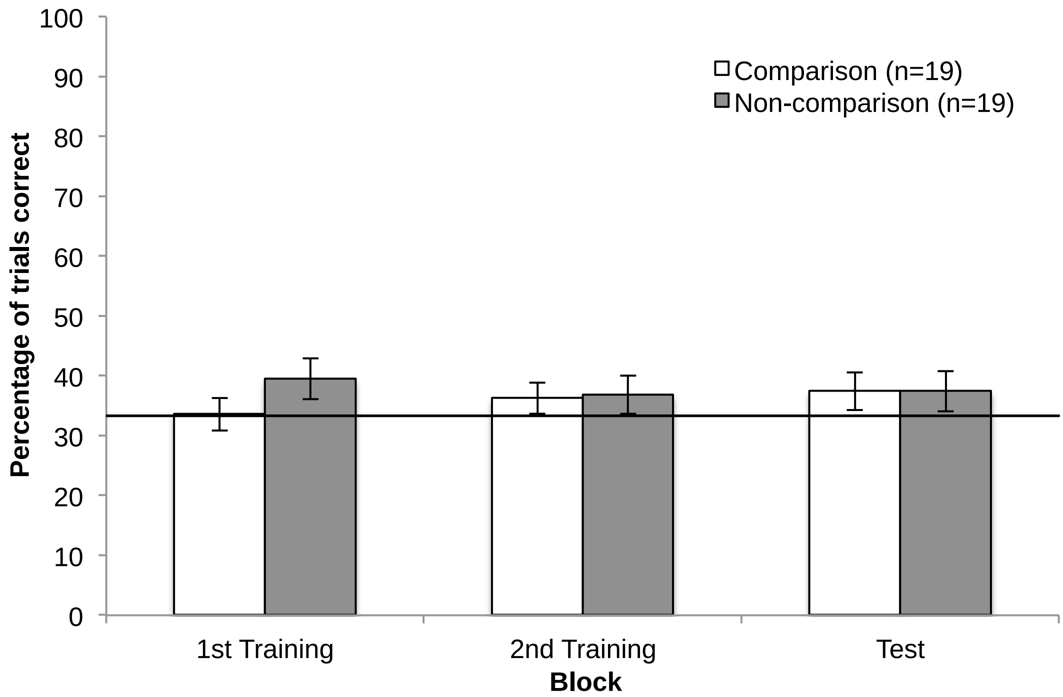

Figure 6. Mean (+/- SE) percentage of trials correct in each training and test block by group in Experiment 7.

Note: Feedback was given in the training blocks, but not the test block. The horizontal line indicates chance level of responding (33%). Participants were asked to make their own descriptions of the wine while either directly comparing and contrasting the wines (Comparison group) or by describing each wine individually (Non-comparison group).

Identification performance. Performance for each group is shown in Figure 6. The groups did not differ significantly in any block (largest $t(36) = 1.34$, $p = 0.188$ for the first training block). Furthermore, neither group performed significantly better than chance during any of the blocks (largest $t(18) = 1.81$, $p = 0.087$ for the non-comparison group in training block 1).

Furthermore, when each wine was compared to chance performance, the only wine in which the participants performed significantly better than chance was the Shiraz during the test block, $t(37) = 2.69$, $p = 0.011$.

There were no significant differences in performance between those who used accurate terms (e.g. spicy for Shiraz) compared to those who used inaccurate terms (largest $t(11.47) = 1.48$, $p = 0.166$) for any of the wines. Furthermore, those who gave consistent descriptions were compared to chance for each of the wines and performance was not significantly higher than chance (largest $t(7) = 1.13$, $p = 0.297$).

Discussion

The novice participants were unable to match their own descriptions back to the wine samples with this amount of training. This is in contrast with the result seen in Experiment 1, where participants were able to learn to use their own descriptions to identify the odour mixtures, despite many of the descriptions being unrelated to the odours.

An argument could be made that, as the words used by many of the participants were not consistent with the winemakers' tasting notes, the descriptions produced by the participants were not accurate. These terms were chosen not just because they appeared in the winemakers' tasting notes, but also because they are commonly used to describe distinguishing features of the three grapes used in the wines, although there is often some overlap between tasting notes of different wines. As seen in Table 5, there is little agreement amongst the participants in terms of which words to use for each wine, despite being given the same resource (the Wine Wheel). Despite this, the results from Experiment 1 suggest that participants can learn to use terms that are relevant *to them* when learning to identify odour samples.

These results for the present experiment demonstrate that the same is not true for wine samples, suggesting that the participants were unable to match relevant terms to the sensations derived from the wines and that the terms that they used were not useful for them in this particular context.

This measurement of description accuracy is not without flaws, as the descriptions written by these participants were effectively compared to one expert for each wine. It is possible that, while the term ‘spice’ was deemed as a correct term for the Shiraz, the other wines also contained a spicy element. Determining the accuracy of a description of something as complex as a wine is not easy and the measures described above should be interpreted with caution. However, given that the participants were able to use these terms in Experiment 5, they should have also been able to use them in the present experiment. As only a small number of the participants gave accurate descriptions for each wine, the non-significant finding for any of the wines could simply be due to lack of power.

One possible explanation for this is that the participants in this study were forced to use terms in the Wine Aroma Wheel. This restriction was introduced for two reasons. The first is that this was an attempt to force the participants use more concrete terms, rather than the abstract terms that Lawless (1984) observed in novice descriptions. Secondly, this experiment was performed after the experiments listed in subsequent chapters. During some of these later experiments, the participants were asked to use their own words to describe wines. These descriptions will be described in subsequent chapters, but were generally of very poor quality. In general, the terms that the participants used in their descriptions are mostly common terms, so it is unlikely that the terms meant little to them. It is more likely that the participants were unable to use these terms in a meaningful way *in a wine context*.

Summary

The results from Experiment 5 indicate that above chance performance was obtained when participants were given the grape name or terms from the winemaker’s description that are relevant to the unique elements in each wine sample, although performance is not as good as that seen with odour samples. Experiment 7 demonstrated that, unlike odour samples, the participants could not match their own wine descriptions back to the wines, even with training.

CHAPTER SEVEN – DISCRIMINATION BETWEEN ALTERNATE SETS OF WINES AND THE EFFECT OF LABELS ON DISCRIMINATION

The experiments in Chapter 5 demonstrated that participants can learn to identify confusable odour mixtures when given appropriate names for the unique elements. In Chapter 6, a set of red wines were tested and found to be of similar discriminability to the odour samples used in Chapter 5. The identification experiments in Chapter 6 indicated that the participants can also learn to identify wine samples, but not the same extent as seen with the odour samples in Chapter 5.

The relatively poor performance in identifying the wine samples could be due to a number of factors. For example, wine samples are more complex than the odour samples used in Chapter 5 and contain more overlapping volatile odorants from a perceptual perspective. The following experiments were designed to test different sets of wines to those used in Chapter 6 in order to determine whether we could find a set of wines that would be easier for the participants to discriminate. More discriminable wines may lead to better identification performance, which would facilitate further experiments. The experiments reported here were amongst a series of experiments designed to find a set of wines that would be easier to discriminate (see Appendices E and F).

Experiment 8 – Testing Discrimination of a Set of White Wines

Whereas the set of wines used in Experiments 4 and 5 were all red wines, Experiment 8 used a set of white wines to test whether these might be easier for the participants to discriminate and thus more useful for future identification experiments. This experiment used a triangle test procedure, similar to that used in Experiment 4 (Chapter 6). While previous studies (e.g. Gawel & Godden, 2008) have found that red wines are generally more discriminable than white wines, white wines can vary widely in terms of winemaking techniques. For example, it is relatively rare to find a red wine that has not received oak treatment, whereas many whites do not. Thus, while the better discrimination for red compared to white wines that has previously found is not in question, that does not necessarily mean that the particular sets of red wine used in previous experiments would necessarily be more discriminable than any set of white wines. Furthermore, there is some evidence that younger wine consumers, particularly female wine consumers, tend to prefer white wine (Hanni & Utermohlen, 2010). According to personal correspondence with wine retailers, this also appears to be the case with young female Australians. Thus it was possible that most of the participants may have had more exposure to white wine than red wine. Thus, discrimination experiments using white wines were conducted in order to attempt to find a set of stimuli that were more discriminable.

Method

Participants. Fifteen first-year Psychology students (nine females, aged 18 to 30, *M* = 20.1, *SD* = 3.0) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. The odour samples were the same as in the experiments in Chapter 5, namely vanilla + Citral, melon + Citral and banana + Citral. The odours were included as a control to serve as a check for motivation and perceptual ability amongst the participants. The white wines were a Yalumba ‘‘Mawson’s” Sauvignon Blanc (2008), a Yering Station ‘‘Mr. Frog’’ Chardonnay (2008) and a Leasingham “Exclusive Release” Riesling (2007). See pages 55 and 57 for general information about preparation of the samples.

Procedure. Trials consisted of three stimuli, two of which were identical and one different. The participants sniffed all three wines in a trial and were asked to identify the unique stimulus. No feedback was given. Each trial consisted only of odour stimuli or of wine stimuli. See page 58 for general information about the procedure that was used in this experiment.

Results

On average, the participants were correct in 6.67 (*SD* = 2.16) of the wine trials and 10.93 (*SD* = 2.84) of the odour trials. As seen in previous experiments, performance on the odour trials was significantly greater than chance ($t(14) = 6.73$, $p < 0.001$). However, performance on this set of wines was not significantly better than chance ($t(14) = 1.20$, $p = 0.252$) and also significantly lower than performance on the odour task ($t(14) = 5.45$, $p < 0.001$).

For the odour samples (Table 6), it seems that banana and melon were the most difficult stimuli to discriminate. For the wines (Table 7), it seems that almost all of the stimuli were difficult to discriminate, with a relatively high error rate for almost all pairings.

Table 6. Summary of errors by target and foil combination for odour samples in Experiment 8.

Target	Vanilla		Melon		Banana	
Foils	Melon	Banana	Vanilla	Banana	Vanilla	Melon
Mean errors (/6)	0.93	0.87	1.33	1.53	1.00	1.40

Table 7. Summary of errors by target and foil combination for wine samples in Experiment 8.

Target	Riesling		Chardonnay		Sauvignon Blanc	
Foils	Chard-onnay	Sauvignon Blanc	Riesling	Sauvignon Blanc	Riesling	Chard-onnay
Mean errors (/6)	2.40	1.40	2.00	1.67	1.60	2.27

Discussion

The participants were not able to discriminate between these wines at a level above chance, suggesting that this is a more difficult set of wines than the reds used in Experiments 4 and 5, which is consistent with the finding by Gawel and Godden (2008) that red wines are generally more discriminable than white wines. Furthermore, performance on the odour samples indicates that this result was not due to any perceptual impairment on the part of the participants. Two other possible combinations of white wines were tested and results were greater than chance, but still no better to those in Experiment 4 (see Appendix E and F), providing empirical evidence that the choice of red wines for subsequent experiments was appropriate.

Experiment 9 – Testing Discrimination of a Set of Wines Containing Both Red and White Wines

Given that discrimination performance amongst the white wines in Experiment 8 was not greater than chance and thus not a suitable replacement for the red wines used in Experiments 4 and 5, the aim of Experiment 9 was to determine whether a set of wines that included a mixture of white and red wines might be easier for the participants to discriminate.

Method

Participants. Sixteen first-year Psychology students (nine females, aged 18 to 52, *M* = 23.1, *SD* = 10.7) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. The odour samples were the same as in the experiments in Chapter 5 and Experiment 8 in this chapter. The white wines were a Yalumba ‘‘Y Series” Sauvignon Blanc (2008) and a Yering Station ‘‘Mr. Frog’’ Chardonnay (2008) and the red wine was a Yalumba “Galway” Shiraz (2008). See pages 55 and 57 for information about how the samples were prepared.

Procedure. The procedure was identical to that in Experiment 8.

Results

On average, the participants were correct in 10.41 (*SD* = 2.29) of the wine trials and 9.82 (*SD* = 3.09) of the odour trials. Performance on both stimuli was significantly higher than chance performance of 6 correct trials, $t(17) = 7.93$, $p < 0.001$ and $t(17) = 5.11$, $p < 0.001$ respectively. No significant difference was observed between performance on the wine and odour stimuli, $t(17) = 0.61$, $p = 0.553$.

For the odours (Table 8), once again it appears that most possible pairings were relatively easy to discriminate. However, for the wines (Table 9), it appears that some combinations were more difficult to discriminate than others. Notably, the Shiraz and Sauvignon Blanc were relatively easy to discriminate, as were the Chardonnay and Sauvignon Blanc. However, trials that required participants to discriminate between Shiraz and Chardonnay produced the most errors.

Table 8. Summary of errors by target and foil combination for odour samples in Experiment 9.

Target	Vanilla		Melon		Banana	
Foils	Melon	Banana	Vanilla	Banana	Vanilla	Melon
Mean errors (/6)	1.47	1.76	1.29	1.41	1.06	1.18

Table 9. Summary of errors by target and foil combination for wine samples in Experiment 9.

Target	Shiraz		Chardonnay		Sauvignon Blanc	
Foils	Chard-onnay	Sauvignon Blanc	Shiraz	Sauvignon Blanc	Shiraz	Chard-onnay
Mean errors (/6)	2.12	0.82	1.65	1.12	0.65	1.24

Discussion

Similar to the wines in Experiment 4, discrimination performance on this set of wines was not significantly different to the discrimination performance for the odour samples. This indicates that these wines would be a more useful set than those tested in Experiment 8. However, despite the finding that the participants could discriminate the wines at a level significantly higher than chance performance, this set of wines is still not particularly easy to discriminate. No other set of wines tested in Appendices E and F were more discriminable than the odour samples.

Furthermore, the results from Experiment 9 indicate that the main wine that stood out was the Sauvignon Blanc (a white wine), rather than the only red wine. This may be because the Shiraz and Chardonnay were similar in certain ways, such as both being exposed to oak during maturation. This finding also indicated that the food colouring added to the wines was sufficient in terms of minimising visual differences between the wines. Furthermore, the result indicates the difficulty of arranging a set of three wines that are all discriminable from each other.

Apart from the wines in Experiment 8, the vast majority of triangle tests conducted resulted in approximately 10 or fewer out of 18 correct trials for almost any stimulus (see also Appendices E and F). I was concerned that there may have been a problem with the triangle test procedure, in that performance never got any higher and that this test may have therefore been insensitive. This hypothesis was explored in Appendix G. The conclusion from that experiment was that the triangle test was a sensitive measure and was appropriate for determining the discriminability of the wine sets.

The triangle test procedure was also used for the experiment that follows.

Experiment 10 – Testing the Effect of the Presence of Appropriate Labels on Discrimination of Odour Mixtures

One possible interpretation of the results from the experiments that demonstrated the appropriate label effect in Chapter 5 could be that the labels primed the participants. The effects of priming on odour identification have been described before (Olsson, 1999). I am unaware of any previous papers examining the effect of priming on olfactory discrimination, although previous evidence for the effect of priming on discrimination has been found in visual perception (Verfaillie, 2000). Thus, a question stemming from the results from Chapter 5 is whether the priming effect aids with lower processing levels such as discrimination as well as higher processing levels such as identification.

Method

Participants. Twenty-five first-year Psychology students (18 females, aged 18 to 43, *M* = 20.6, *SD* = 5.3) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. The odour samples were as in the experiments in Chapter 5 and the previous experiments in this chapter. That is, the unique elements in each sample were vanilla, melon and banana. All were mixed with a relatively strong lemon (Citral) odour. See page 55 for information about how the samples were prepared.

Procedure. The experiment followed the same procedure as other discrimination experiments, outlined on page 58, with the following exceptions. Participants were randomly allocated to either the “label” group (*n* = 12) or the “no label” group (*n* = 13). The procedure was the same for both groups except that participants in the label group were made aware that there were three different types of samples; one with a vanilla odour, one with a melon odour and one with a banana odour. Participants in the no label group were not given this information.

Results

The mean number of trials correct for the label group was 10.00 out of 18 (*SD* = 2.24), which was not significantly different to the no label group (*M* = 10.00, *SD* = 2.99). Furthermore, both groups were significantly higher than chance performance of 6 out of 18 correct trials, $t(11) = 4.64$, $p = 0.001$ for the label group and $t(12) = 6.45$, $p < 0.001$ for the no label group. The errors for each target and foil combination for each group are shown in Table 10.

Table 10. Summary of errors by target and foil combination for odour samples by group in Experiment 10.

Target	Vanilla		Melon		Banana	
Foils	Melon	Banana	Vanilla	Banana	Vanilla	Melon
No label	1.31	1.23	1.85	1.38	1.08	1.15
Label	1.17	1.58	1.33	1.83	1.17	0.92

Discussion

That both groups performed significantly greater than chance was not unexpected. The lack of any difference between label and no label conditions leads to two possible conclusions. If the appropriate label effect described in Chapter 5 worked by guiding the participants to the appropriate perceptual information in the stimuli, then this effect should also work on discrimination performance. This was not the case in this experiment. Secondly, the appropriate label effect only appears to work on higher-level cognitive processes and any top-down influences on simple discrimination stemming from this effect are limited.

Summary

Experiments 8 and 9, along with the experiments presented in Appendices E and F, were designed to determine whether a set of wines could be found that was more discriminable than the set of red wines used in

Experiments 4 and 5 in Chapter 6, in an attempt to improve identification performance in the wine tasks. No such set of three wines was found, even when the set included a mixture of disguised white and red wines. Furthermore, the results from Experiment 10 suggested that the appropriate label effect observed in Chapter 5 does not appear to aid discrimination, indicating that the effect works on a higher level in the olfactory processing hierarchy.

CHAPTER EIGHT – ADDING ELEMENTS TO WINES AND THE EASY-TO-HARD EFFECT

The experiments in Chapter 6 indicated that participants were able to learn to identify wines using labels that were considered to be relevant to the wine (such as learning to use the labels “spice/pepper and chocolate” for Shiraz), but not to the extent seen in odour mixtures of similar discriminability in Chapter 5. This may be because the relevant element in the wines was more difficult to detect, as wines contain more odour elements than the binary odour mixtures used in Chapter 5. If that is the case, then the Shiraz would not be perceived as particularly peppery (compared to the other wines) by the participants. Similarly, the Pinot Noir would not be perceived as having cherry odours and the Cabernet Sauvignon would not be perceived as having blackcurrant or blueberry flavours compared to the other wines. However, these terms were the dominant descriptors in the winemakers’ descriptions for the wines. Furthermore, the participants were able to use them to some extent when learning to identify the wine samples in the experiments reported in Chapter 6.

It is possible that these elements were not very strong in the wines that were chosen for previous wine experiments. As such, the first experiment in this chapter aimed to determine whether wines could be found with odours that were typical of wines that have the qualities of pepper (for Shiraz), cherry (for Pinot Noir) or blackcurrant or blueberry (Cabernet Sauvignon) compared to wines made of other grapes. One reason for doing so was to determine which wines may be useful for subsequent experiments. A second, and more important, reason was to determine whether the difficulty experienced by novices in using these “appropriate” labels was that the labels did not actually refer to an element of the wines that they could detect. The second experiment was designed to test whether increasing the amount of the relevant odour in the wine would induce participants to rate the wines as having more of the relevant odour amongst novices and whether more than one concentration of additives could be found for the subsequent experiment, given that this technique is commonly used during olfactory training of panels (see the Intensive Training section of Chapter 3). The final experiment aimed to use the principle behind Pavlov’s “transfer along a continuum”, where learning gained from an easy version of a task transfers to more difficult versions of the task (Mackintosh, 1975; Pavlov, 1927). As described in Chapter 3, this method of training novices by adding elements to stimuli has been used before for beer (Chollet et al., 2005) and wine (LaTour et al., 2011).

Experiment 11 – Comparisons of the Perceived Intensity of Prototypical Wine Notes by Grape

This experiment was designed to answer two questions: a) whether the participants tend to consider wines made from Shiraz as peppery, wines made from Cabernet Sauvignon as having either blackcurrant and/or blueberry flavours and wines made from Pinot Noir as having cherry elements and b) whether we could find a wine made from each of those grapes with relatively high levels of the relevant elements, as rated by novice participants.

Method

Participants. Nineteen first-year Psychology students (16 females, aged 18 to 22, *M* = 18.8, *SD* = 1.2) participated in the experiment for course credit. See page 53 for more information about the participants.

Materials. Twelve different wines were used in this experiment: four Shirazes, four Cabernet Sauvignons and four Pinot Noirs. The four Shirazes were: Lock and Key Shiraz 2009, Yalumba “Galway” Shiraz 2008, Rymill “Yearling” Shiraz 2008 and Redbank “Long Paddock” Shiraz 2007. The four Cabernet Sauvignons were: Rymill “Yearling” Cabernet Sauvignon 2008, Yalumba “Y Series” Cabernet Sauvignon 2007, Yalumba “Mawson’s” Cabernet Sauvignon 2008 and McWilliam’s “Hanwood” Cabernet Sauvignon 2008. The four Pinot Noirs were: Yering Station “Mr Frog” Pinot Noir 2008, Josef Chromy “Pepik” Pinot Noir 2009, Oyster Bay Pinot Noir 2008 and De Bortoli “Windy Peak” Pinot Noir 2008. All of the wines were priced between $10 and $20 per bottle and were chosen because the winemakers’ descriptions featured the element that was related to the grape, referred to below as “relevant” odours (i.e. pepper for Shiraz, blackcurrant and/or blueberry for Cabernet Sauvignon, cherry for Pinot Noir), without containing any of the other odours. That is, none of the Shirazes were described as having blackcurrant, blueberry or cherry aromas by the winemaker. See page 55 for information about how the wine samples were prepared and presented.

Prior to smelling the wine samples, the participants also smelled odour samples of pepper, blackcurrant, blueberry and cherry. All were presented in screw-top containers on make-up removal pads. Ribena served as the blackcurrant odorant, blueberry was an odour sample provided by Quest (batch number 14367) and the cherry odour was made using cherry-flavoured jelly crystals (Aeroplane Jelly Dark Cherry flavour). All of these were dissolved in water and soaked into make-up removal pads, along with black food colouring to disguise any visual differences. The pepper odour involved black pepper cracked onto a make-up removal pad that had been soaked in water with black food colouring and the grains of pepper were hidden from view. Where possible, these odours were chosen based on the standards proposed by Noble et al. (1987), with the intention that they would be used for subsequent experiments. As no standard for blueberry was proposed, an artificial additive was used instead. The proposed standard for cherry was not available at the time of testing, so a substitute was used, derived from pilot testing.

Procedure. Participants attended a one-hour session with up to two other participants. Each participant then sniffed, in series, sixteen of the odour samples (four trials of each of the four odours: pepper, blackcurrant, blueberry, cherry) in a quasi-random order (for logistic reasons all received the same order, but starting at different points in the order), with the constraint that no more than two consecutive samples were the same. The response options were blackcurrant, blueberry, cherry and pepper and, for each trial, the participants responded on a sheet of paper which odour they thought they had smelled. The aim of this was to determine whether the participants could identify these particular odours. The intertrial interval was 1min, with a five minute break after the fifth and tenth trial.

Participants then smelled and rated each of the twelve wines in a quasi-random order (all received the same order, but starting at different points in the order) with the constraint that no three consecutive wines were of the same grape. The same twelve wines were then smelled and rated a second time by the same participants in a different order. The participants were unaware of how many different wines were involved, what they were made from or that each wine was rated twice. Participants rated each wine in turn and then moved on to the next wine, without being able to go back and smell a previous wine sample. They were also not able to smell the odour samples again during the wine task. The intertrial interval for the wine task was 1min with a 5-min break observed after the twelfth trial. Water was available throughout and participants were encouraged to drink between each trial.

As a distractor task, the participants rated each wine in terms of hedonics; irritation (described as “that feeling you get in your nose when you sniff hot mustard or wasabi”); sweetness; strength; and the relative intensity of blackcurrant, blueberry, cherry or pepper notes in each wine. All ratings were measured on 7-point line scales ranging from 1 (‘not at all”) to 7 (“extremely”), except for the hedonics scale which was measured on an 11-point line scale ranging from -5 (dislike extremely) to +5 (like extremely). Every wine was rated on all of these scales.

Analyses. Mean performance in the odour trials was compared to chance. As there were four response options, the chance performance was one trial correct out of four trials for each odour.

The ratings of hedonics, irritation, sweetness and strength for the wines were irrelevant for this project and were included as distractors. No significant differences were observed in these ratings, so they are not presented below.

The ratings for pepper, blackcurrant, blueberry and cherry were analysed in two ways. Firstly, the ratings for the "relevant" odour for wines from each grape (i.e. pepper for Shiraz, blackcurrant and blueberry for Cabernet Sauvignon and cherry for Pinot Noir) were compared to the "non-relevant" odour ratings. That is, for the wines made using the Shiraz grape, the ratings for pepper were compared using pairwise tests to each of the other ratings (blackcurrant, blueberry and cherry) for each wine to determine whether any of the Shirazes were perceived as particularly peppery. A similar procedure was used for the Cabernet Sauvignons and Pinot Noirs.

Secondly, the ratings for the "relevant" odours for each of the wines was compared to the same rating of the wines made from other grapes using *t*-tests. That is, the pepper rating for each of the Shirazes was compared to the pepper rating for each of the non-Shiraz wines, to determine whether novice participants perceive wines made from Shiraz to be more peppery than the wines made from Cabernet Sauvignon and Pinot Noir. A similar procedure was used for the "relevant" odours of the Cabernet Sauvignon and Pinot Noir wines. Due to the large number of comparisons, a Bonferroni correction was applied.

Results

Odour pretest. The average number of correct responses from four trials of each of the odours were: Blackcurrant = 3.32 (*SD* = 0.67), Blueberry =

3.0 (*SD* = 1.05), Cherry = 3.47 (*SD* = 0.90) and Pepper = 3.74 (*SD* = 0.45), all well above chance of one correct trial for each, with the smallest *t*-value of $t(18) = 8.27$, $p < 0.001$.

Wine trials. The mean ratings for each of the twelve wines in terms of the strength of pepper, blackcurrant, blueberry and cherry odours are presented in Table 11. More information about the first and second ratings of each wine is presented in Table A29 in Appendix L.

No wine was found to be significantly higher on the relevant odour compared to any of the non-relevant odours ($p > 0.05$). That is, no Shiraz was observed to have significantly higher pepper ratings compared to the blackcurrant, blueberry and cherry ratings, nor were any of the Cabernet Sauvignons found to be particularly blackcurrant-like or blueberry-like. Similarly, the Pinot Noirs were not rated as being particularly cherry-like.

Furthermore, none of the Shirazes were rated as being significantly more peppery than any of the Cabernet Sauvignons or Pinot Noirs, nor were the Cabernet Sauvignons rated as being significantly more blackcurrant-like or blueberry-like than any of the Shirazes or Pinot Noirs. Finally, none of the Pinot Noirs were rated as being significantly more cherry-like than any of the Shirazes or Cabernet Sauvignons (all $p > 0.05$). Furthermore, these differences were not due to a strict alpha, as all comparisons were non-significant using an alpha of 0.05.

Table 11. Mean (and SD) ratings of the strength of pepper, blackcurrant, blueberry and cherry odours in the twelve wines in Experiment 11.

Wine	Pepper	Blackcurrant	Blueberry	Cherry
Shiraz 1	**2.55 (1.59)**	2.86 (1.44)	2.42 (1.60)	2.39 (1.17)
Shiraz 2	**2.21 (1.33)**	3.03 (1.22)	2.76 (1.57)	2.55 (1.58)
Shiraz 3	**2.50 (1.62)**	2.92 (1.34)	2.26 (1.35)	2.47 (1.53)
Shiraz 4	**2.15 (1.44)**	2.63 (1.23)	2.63 (1.55)	2.55 (1.35)
Pinot 1	2.36 (1.45)	2.58 (1.26)	2.76 (1.68)	**2.76 (1.57)**
Pinot 2	2.21 (1.26)	2.74 (1.44)	2.50 (1.40)	**2.68 (1.46)**
Pinot 3	2.21 (1.21)	2.87 (1.73)	2.53 (1.58)	**2.45 (1.45)**
Pinot 4	2.11 (1.10)	2.82 (1.57)	2.71 (1.51)	**2.58 (1.10)**
Cabernet 1	2.55 (1.49)	**2.50 (1.38)**	**2.26 (1.32)**	2.55 (1.54)
Cabernet 2	2.89 (1.85)	**2.63 (1.09)**	**2.34 (1.23)**	2.47 (1.30)
Cabernet 3	2.50 (1.46)	**2.89 (1.65)**	**2.21 (1.28)**	2.37 (1.34)
Cabernet 4	2.18 (1.40)	**3.13 (1.41)**	**2.61 (1.39)**	2.29 (1.21)

Note: Ratings were collected using a 7-point line scale ranging from 1 (not at all) to 7 (extremely). Scores in bold represent the appropriate odour for each wine. Shiraz 1: Lock and Key, Shiraz 2: Yalumba “Galway”, Shiraz 3: Rymill “Yearling”, Shiraz 4: Redbank “Long Paddock”. Pinot 1: Yering Station “Mr Frog”, Pinot 2: Josef Chromy “Pepik”, Pinot 3: Oyster Bay, Pinot 4: De Bortoli “Windy Peak”. Cabernet 1: Rymill “Yearling”, Cabernet 2: Yalumba “Y Series”, Cabernet 3: Yalumba “Mawson’s”, Cabernet 4: McWilliam’s “Hanwood”.

Discussion

Despite being able to identify the relevant odours very well in isolation, the participants were not able to identify the relevant odours in the wines. That is, the participants did not appear to perceive Shirazes as being particularly peppery, nor did they perceive Cabernet Sauvignons as having blackcurrant or blueberry aromas, nor did they perceive Pinot Noirs as having cherry-like aromas compared to the wines made from the other varietals.

While the wines are discriminable (see experiments in Chapters 6 and 7), the odours specific to each grape are clearly very subtle and therefore difficult for them to either detect or to identify within a wine context. The unique odour in each of the odour samples used in Chapter 5 was also disguised with a strong common element, but wine samples are much more complex. Furthermore, these wines were only smelled, whereas the descriptions by the winemakers are typically produced after tasting the wine, indicating that it could well be the case that isolating the appropriate samples is not possible based on smell alone. One final consideration is that the descriptions by the winemakers may not discriminate between the samples very well, or simply inaccurate. This is discussed further in Chapter 11.

The results here indicate that profiling descriptors that are presumed to be present in, and relevant to, a wine can be very difficult. The next experiment aims to determine whether artificially altering the wines can result in changes to the odour profiles of the wines that are detectable by novices.

The wines with the highest relevant ratings were thus selected for the next experiment and the final experiment in this chapter, with the exception of the Shiraz. The Shiraz that was most peppery at any time (Shiraz 3 – Rymill

“Yearling”) was also rated relatively highly on the blackcurrant scale, so the second highest rating was chosen instead (Shiraz 1 – Lock and Key).

Experiment 12 – Testing the Perceived Intensity of Wine Notes When They Are Enhanced Using Additives

The aim of Experiment 12 was to determine which concentrations of the additives would enhance the relevant element of each wine (e.g. pepper for Shiraz) for novices, using a procedure similar to studies by Noble et al. (1987) and Chollet et al. (2005).

Method

Participants. Twenty-four first-year Psychology students (16 females, aged 18 to 24, *M* = 19.3, *SD* = 1.7) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. Three different wines were used in this experiment, based on the results from Experiment 11: Lock and Key Shiraz 2009, McWilliam’s “Hanwood” Cabernet Sauvignon 2008 and Yering Station “Mr Frog” Pinot Noir 2008. See page 55 for general information about how the wine samples were prepared and presented.

Prior to smelling the wine samples, the participants also smelled the same odour samples of pepper, blackcurrant, blueberry and cherry described in Experiment 11 above.

These odour samples also served as the additives for the wines. Concentrations of these additives are listed in Table 12. Pepper was added to the Shiraz, cherry to the Pinot Noir and the Cabernet Sauvignon was presented with one of two different additives: either blackcurrant or blueberry, as there was uncertainty as to which would be easier for participants to detect. As the Cabernet Sauvignon was presented with blackcurrant or blueberry, there were essentially four wines, each presented with three different levels of additive (none, low, high). At least three levels of additive were required for the next experiment (Experiment 13) and pilot testing indicated that there was little difference between ratings of further levels of additive. Where possible, the standards proposed by Noble et al. (1987) formed the basis of the odours and additives used in the experiments in this chapter. As no standard for blueberry was proposed, an artificial additive was used instead. The proposed standard for cherry was not available at the time of testing, so a substitute was used.

Each participant rated each wine sample twice for a total of 24 rated wines.

Table 12. Concentration of additives in wine samples in Experiment 12.

Odour name	Low concentration	High concentration
Pepper	2g:200mL wine	8g:200mL
Blackcurrant	1:100	1:33.3
Blueberry	1:1250	1:625
Cherry	1:417	1:156

Note: All odorants added to 200mL of wine.

Procedure. As in Experiment 11, each wine was rated on line scales (1 = “Not at all” to 7 = “Extremely”) for pepper, blackcurrant, blueberry and cherry flavours. Ratings for liking, irritance, sweetness and strength were also obtained as a distractor task, but were not analysed.

Analyses. Each combination of wine and additive concentration was presented twice. For all but the low and high concentrations of cherry in the Pinot Noir sample, the ratings over the repetitions were correlated (with the smallest *t*-value of $r = 0.44$, $p = 0.034$) and were therefore averaged to form one rating per participant for each additive concentration of each wine. For the cherry additives, the correlation coefficients were approximately 0.3 indicating that there was still some level of agreement over time in terms of the ratings. This finding is discussed further below.

As the purpose of the experiment was to test whether the concentrations of the additives were detectable, the rating of each level of the relevant additive for each wine was compared to the ratings for all other levels of additives for that wine. That is, for the Shiraz, the high level of pepper was compared to the low level of pepper and the no added pepper Shiraz samples in terms of their rated pepperiness. Furthermore, the low level of pepper and no added pepper Shiraz samples were also compared. This was also done for the Pinot Noir samples in terms of the ratings on the cherry scale and for the Cabernet Sauvignon on the blackcurrant and blueberry scales.

Results

Odour pretest. The average number of correct responses from four trials of each of the odours were: Blackcurrant = 3.00 (*SD* = 0.93), Blueberry = 2.42 (*SD* = 1.32), Cherry = 3.17 (*SD* = 0.82) and Pepper = 3.13 (*SD* = 0.74), all significantly higher than chance performance of one correct trial for each, with the smallest *t*-value of $t(23) = 5.27$, $p < 0.001$.

Wine ratings. The ratings for the none, low and high additive concentrations on the relevant dimension (Table 13) were all significantly different from each other for the Shiraz with pepper and Cabernet Sauvignon with blueberry samples (with the smallest *F*-value of $F(1,23) = 4.93$, $p = 0.037$ for the low vs high concentration of blueberry in Cabernet Sauvignon). When a Bonferroni correction was applied, the latter was the only difference that was no longer statistically significant.

For the Cabernet Sauvignon with blackberry and the Pinot Noir with cherry, the only significant difference was for the comparison of low vs high concentrations of cherry additive for the Pinot Noir ($F(1,23) = 6.88$, $p = 0.015$), which was not significant after a Bonferroni correction. No other comparisons were statistically significant.

Table 13. Mean (and SD) ratings for each of the twelve wines in terms of strength of pepper, blackcurrant, blueberry and cherry odours in Experiment 12.

		Amount of additive		
Wine	Additive	None	Low	High
Shiraz	Pepper	2.31 (1.21)	3.56 (1.91)	4.90 (2.11)
Cabernet Sauvignon	Blackcurrant	2.83 (1.49)	2.58 (1.32)	3.06 (1.72)
Cabernet Sauvignon	Blueberry	2.56 (1.51)	4.90 (1.64)	5.40 (1.73)
Pinot Noir	Cherry	2.75 (1.50)	2.63 (1.23)	3.31 (1.27)

Discussion

The ratings in Table 13 indicate that the pepper additive in the Shiraz was quite successful in terms of creating three different levels of additive, while the blueberry in Cabernet Sauvignon was also successful in this regard. However, the blackcurrant additive appears to have been too weak for the participants to detect. The results for cherry indicate that the particular cherry additive used here was also not useful as an additive.

However, the results for the Shiraz with pepper and the Cabernet Sauvignon with blueberry indicated that the participants were able to use the scales in an appropriate fashion, indicating that they could not only detect, but identify the additives in the wines.

A further experiment using a similar procedure was conducted in order to determine the appropriate concentration for the cherry additive in the Pinot

Noir. This is reported in Appendix I and the concentrations from this subsequent experiment were used for the low and high concentration conditions with the Pinot Noir in the following experiment.

These results build upon the methodology used by Noble et al. (1987) by determining not just one level of perceivable additives, but two, that can then be used to address the aim of the next experiment. The technique used in this experiment, whereby participants rated each wine in terms of the intensity of each of the possible additives, served as a measure that was not only sensitive enough to detect the differences between the concentrations of additives, but to also rank them in the correct order. It is acknowledged that the differences between each concentration were not detected for every additive and so the subsequent test (Appendix I) was necessary. A further benefit of this approach is that it is economical in terms of the number of participants and the amount of time required to complete the task, although the procedure may not be as rigorous as that used elsewhere (e.g. Goodstein, Bohlscheid, Evans, & Ross, 2014).

Experiment 13 – Can the Easy-to-Hard Effect Be Utilised as a Training Method for Novices?

The aim of Experiment 13 was to determine if making the wine identification task easier in initial training (by adding the relevant elements to the wine samples) resulted in a transfer of learning to the more difficult version of the task (where additives were no longer present). In the learning literature, this effect is known as the “easy-to-hard” effect, or “transfer along a continuum” (Mackintosh, 1975; Pavlov, 1927).

Another way of thinking about this experiment is that the labels on which the participants were trained required them to identify pepper in a Shiraz, blueberry in a Cabernet Sauvignon and cherry in a Pinot Noir. The results from Experiment 11 suggested that the participants did not consider the wines made from one grape to be higher on the odour relevant to that wine than wines made from another grape. From a signal detection perspective, it appears that the participants were unable to detect the relevant signal. By using additives, the signal (e.g. pepper in Shiraz) should stand out more from the noise (the other elements in the wine, particularly those that are in all of the wines, such as alcohol). By increasing the strength of the signal, the participants should have a better chance of detecting the signal amongst the noise and this should result in them learning what the signal smells like amongst the noise. That is, increasing the signal should help the participants to learn which components of the odour to attend to. A similar rationale has been used with olive oil (Paredes-Olay, Moreno-Fernandez, Rosas, & Ramos-Alvarez, 2010). In this study, participants tasted samples of oils, which were mostly sunflower oil with varying concentrations of olive oil. Their task was to determine whether the sample contained olive oil. Varying concentrations of olive oil were tested in order to test whether signal detection theory was appropriate for olive oil testing.

In the present experiment, half of the participants were trained on wines that included decreasing amounts of added odours (the easy-to-hard group) while the other half were trained on unadulterated wines throughout. The aim was to determine whether being trained on wines with added elements aided later identification performance.

Method

Participants. Forty-eight first-year Psychology students (32 females, aged 18 - 49, *M* = 20.2, *SD* = 4.8) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. Three different wines were used in this experiment, based on the results from Experiment 11: Lock and Key Shiraz 2009, McWilliam's "Hanwood" Cabernet Sauvignon 2008 and Yering Station "Mr Frog" Pinot Noir 2008. See page 55 for general information about how the wine samples were prepared and presented.

For the easy-to-hard group, the additives that were determined in Experiment 12 and Appendix I were added in the first two training blocks. Concentrations of the additives are listed in Table 14. Pepper was added to the Shiraz, cherry to the Pinot Noir and blueberry was added to the Cabernet Sauvignon. No additives were used for the hard-to-hard group.

Table 14. Concentration of additives in wine samples in Experiment 13.

Odour name	Low concentration	High concentration
Pepper	2g:200mL wine	8g:200mL
Blueberry	1:1250	1:625
Cherry (Queen Cherry brandy)	1:500	1:333

Note: All odorants added to 200mL of wine.

Procedure. Upon arrival, participants were allocated into one of two groups according to their time of arrival: the hard-to-hard or easy-to-hard group. Up to three participants were tested in each session and all participants in the same session were allocated to the same group.

On each trial participants were asked to identify the wines using the labels pepper, blueberry and cherry. They received three 18-trial training blocks with feedback and one test block without feedback. Each block consisted of six trials of each of the three wines. For the hard-to-hard group, no additives were added to the wines for any of the blocks. Their conditions were thus similar to the Appropriate group in the experiments in Chapter 5. For the easy-to-hard group, all wine samples in the first training block included high concentrations of the relevant additive (i.e. the Shiraz contained pepper, the Cabernet Sauvignon contained blackcurrant and the Pinot Noir contained cherry), while all samples in the second training block included low concentrations of the relevant additive. In the third training and test blocks, the easy-to-hard group received wines that did not contain any additives.

All timing, feedback, stimulus randomisation and data collection was conducted using a custom Inquisit script on PCs for data collection (Millisecond Software LLC, 2011).

Analyses. Each group was compared to chance performance (6 out of 18 items correct) in each block using one-sample *t*-tests. Furthermore, groups were compared to each other at each block using independent samples *t*-tests. A mixed model ANOVA was conducted and linear trend and interaction post-hoc tests were run to determine whether any learning effects were present.

Results

The easy-to-hard group performed significantly better than chance in all blocks except for the third training block (with the smallest *t*-value of $t(25) = 3.24$, $p = 0.003$ for the test block). The hard-to-hard group performed significantly better than chance in the second and third training blocks (with the smallest *t*-value of $t(21) = 2.21$, $p = 0.038$) as well as the test block ($t(21) = 3.85$, $p = 0.001$). When a Bonferroni correction was applied, the hard-to-hard groups were no longer significantly better than chance in the second and third training blocks, while no result changed for the easy-to-hard group (Figure 7).

The groups differed significantly in the first and second training blocks ($t(46) = 2.86$, $p = 0.006$ and $t(46) = 2.07$, $p = 0.044$) although the result for the second block was no longer significant when a Bonferroni correction was applied. The groups did not differ significantly in the third training or test blocks.

A significant linear decrease over blocks was observed for the easy-to-hard group ($F(1,25) = 8.69$, $p = 0.007$) reflecting the decrease in performance as the amount of additive was decreased, while no significant linear trend was found for the hard-to-hard group ($F(1,21) = 1.48$, $p = 0.24$). The linear trends were significantly different from each other ($F(1,46) = 8.97$, $p = 0.004$.

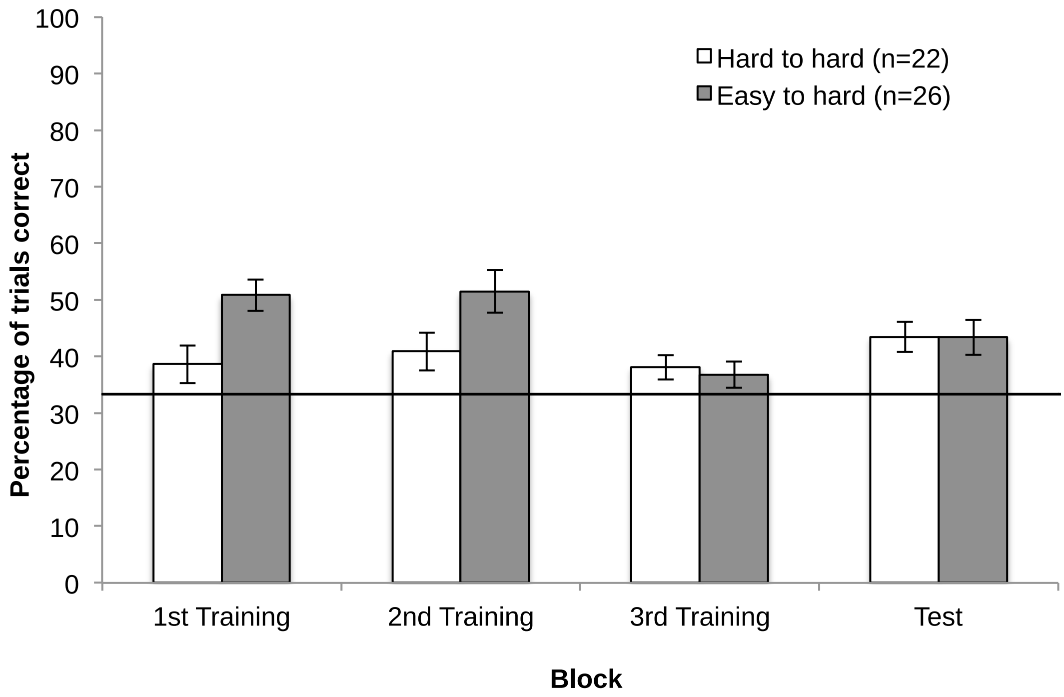

Figure 7. Mean (+/-SE) percentage of trials correct in each training and test block by group in Experiment 13.

Note: Feedback was given in the training blocks, but not the test block. The horizontal line indicates chance level of responding (33%). Participants either received wine with additives in the first two training blocks (easy-to-hard group) or received wines with no additives (hard-to-hard group). Both groups received wines with no additives in the third training block and test.

Discussion

As expected, the additives made the task easier for the easy-to-hard group, although they were unable to transfer this performance to the wine samples when the additives were no longer present. Thus it appears that the participants in the easy-to-hard group were simply naming the additives in the wines, not the actual wines themselves. That they could not transfer this learning to the unadulterated wines in the third training block and test block indicate that, at least in the short term, this may not be a useful training procedure. Furthermore, performance was still relatively poor, even with the highest level of additives.

In contrast with the results from previous experiments in this chapter, the most interesting results from this experiment was that, on test, both groups performed significantly better than chance. If the participants in the easy-to-hard group were only learning to identify the added elements in the wines, then they would have had less training on the actual wines than the hard-to-hard group. This suggests that either the relatively short amount of training received by the easy-to-hard group on the wines (the final training block) was as useful as the three training blocks that the hard-to-hard group received, or that the easy-to-hard group still learned something about the wines during the blocks when the wines contained additives. The finding that the easy-to-hard group did not perform at a level significantly higher than chance in the third block provides more evidence for the first of these hypotheses than it does for the second.

Summary

The results from Experiment 11 indicated that the participants did not perceive the wines as particularly peppery (for Shiraz), cherry-like (for Pinot Noir) or blackcurrant- or blueberry-like (for the Cabernet Sauvignon) when using only via orthonasal olfaction. This finding potentially explains why participants found it so difficult to identify the wines based on these labels.

Participants were clearly better at applying labels to the wines when additives were present, both in being able to use the relevant scales to rate the wines and in identifying the wines in Experiment 13. This indicates that perhaps the relevant components of the wine (e.g. pepperiness in Shiraz) are too subtle for novices to detect, despite being one of the dominant features of the winemaker's description.

However, the participants in both groups in Experiment 13 were able to identify the wines using element labels in the final training block at a level significantly higher than chance. This indicates that at least some level of learning took place and that these labels may be somewhat appropriate for novices to use even when the relevant element is no longer enhanced using additives. This finding is similar to that seen for the Descriptor group in Experiment 5 (Chapter 5).

There was no indication of the "easy-to-hard" or "transfer along a continuum" effect. Due to adaptation effects, only a few trials can be included within a given session. However, this would involve far more training sessions than were practical given the limitations of four hours per participant according to the rules of the subject pool.

CHAPTER NINE – THE USE OF FLAVOUR AND LONGER TRAINING SESSIONS

In all of the experiments reported in Chapters 6, 7 and 8, the participants were only allowed to smell the wines. The general finding from these experiments was that the participants could learn to identify the wines using smell alone under certain conditions, but generally not to the extent seen with the more simple odour samples used in Chapter 5.

While olfaction is an important component of flavour, it is likely that these olfaction-only studies were limited in that the participants did not receive important information from the samples in the form of retronasal olfaction, taste and mouthfeel. To address these limitations, the experiments reported in this chapter involved tasting the wines rather than just smelling them.

Due to the large number of wine samples being tasted, and given that the wine samples contain alcohol, there were breaks of at least 24 hours between blocks. Furthermore, the participants were required to spit out the wine, in order to minimise the effects of alcohol on their performance.

Unlike the experiments reported in Chapter 8, no additional elements were added to the wines. Instead, the aim of the first experiment in this chapter is to repeat the procedure used in Chapter 6, with the only change being that the wines are tasted, instead of just smelled.

Experiment 14 – The Effect of Grape Name vs Descriptor Labels on Wine Flavour Identification

The aim of the current experiment was to determine whether the use of flavour assists novice participants in the identification of wines when they are given short wine descriptions or grape names.

Method

Participants. Twenty-four first-year Psychology students (15 females aged 18 to 45, Mean (*M)* = 21.1, *SD* = 7.5) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. Three different wines were used: Yalumba “Galway” Shiraz 2011, Yering Station “Mr Frog” Pinot Noir 2010 and Donelli Red Lambrusco (NV). All blocks consisted of 18 trials, six of each wine. See page 57 for information about how the samples were prepared and presented.

Procedure. Participants signed up for a one-hour experiment and were randomly allocated to the “grape name” or “two word description” group based on their time of arrival. The experiment consisted of two training blocks and one test block. For each participant, the experiment was split over three days, with one block of stimuli presented on each day. The smallest break between blocks was 24 hours and the longest break between blocks was 5 days. There was no significant difference between groups in terms of intervals between blocks. See page 59 for general information about the procedure used in this experiment.

The task was to identify each sample according to the given labels of that condition. For the grape name group, these labels were: Shiraz, Pinot Noir and Lambrusco. For the two word description group, these labels were: "Spice and chocolate", "Black cherry and gamey" and "Floral and raspberry". Two word descriptors were chosen so that the participants had a choice of words to use. Participants were required to spit out all wine samples after responding, so that no visual cues were received from seeing the wine enter the spittoon. Water was available throughout.

Results

The mean performance in the practice block for the grape group was 3.5 correct out of four items (*SD* = 0.52), which was not significantly different to the description group (M = 3.42, *SD* = 0.67), $t(22) = 0.34$, $p = 0.74$.

Both groups performed significantly higher than chance (six items correct) in both training blocks, with the smallest *t*-value of $t(11) = 6.58$, $p < 0.001$. The description group maintained higher than chance performance (*M* = 10.08, *SD* = 1.93) in the test block, $t(11) = 7.33$, $p < 0.001$. In contrast, the grape group was not significantly higher than chance (*M* = 6.67, *SD* = 1.78), $t(11) = 1.30$, $p = 0.22$. The grape group was also significantly lower than the description group in the test block, $t(22) = 4.52$, $p < 0.001$ (Figure 8).

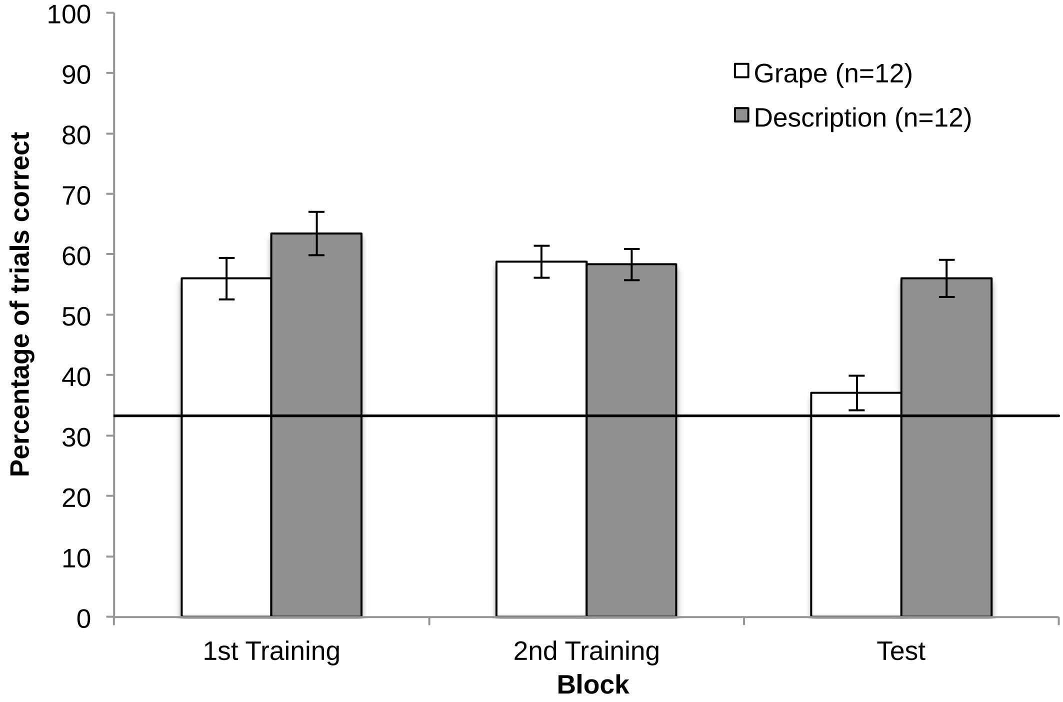

Figure 8. Mean (+/- SE) percentage of trials correct in each training and test block by group in Experiment 14.

Note: Feedback was given in the training blocks, but not the test block. The horizontal line indicates chance level of responding (33%). Participants were given either grape names (Grape name group) or two-word wine descriptions (Descriptor group).

There were no significant differences between the groups in either of the training blocks, nor did either group improve between the training blocks.

When taking the different grapes into account, it appears that the Lambrusco was easy for the participants to identify compared to the Shiraz and Pinot Noir (Table 15). Despite the Lambrusco being somewhat easier to identify during training, the grape group dropped in performance on Lambrusco during test ($F(1,11) = 25.30$, $p < 0.001$), while the description group did not $F(1,11) = 4.77$, $p = 0.052$. The difference between these two quadratic trends is also statistically significant, $F(1,22) = 28.13$, $p < 0.001$. Indeed, the groups differed significantly in their identification of Lambrusco during test ($t(22) = 5.47$, $p < 0.001$), but not in any of the other sessions. Nor did the groups differ in terms of performance of any other wines during any of the sessions (largest $t(22) = 1.74$, $p = 0.095$).

All of the wines were identified at a rate significantly higher than chance (two correct per block) during all blocks by both groups, with the exception of Shiraz during training block 2 for the description group (p = 0.056) and for the Shiraz and Pinot during test for both groups (largest $t(11)$ = 1.59, p = 0.139).

.

Table 15. Mean (and SD) correct for each wine in training and test blocks by group in Experiment 14.

	Training Block 1			Training Block 2			Test Block		
Group	Shiraz	Lambrusco	Pinot	Shiraz	Lambrusco	Pinot	Shiraz	Lambrusco	Pinot
Grape	3.17 (1.64)	4.08 (1.24)	2.83 (1.12)	2.83 (1.20)	5.08 (0.79)	2.67 (0.89)	1.67 (0.78)	2.67 (1.16)	2.33 (0.99)
Description	3.08 (1.44)	4.67 (1.07)	3.67 (1.30)	2.75 (1.22)	4.42 (1.31)	3.33 (0.99)	2.25 (0.87)	5.33 (1.23)	2.50 (1.09)

Note: Each wine was presented six times in each block.

Discussion

While both groups performed similarly during training, performance was only maintained during the test phase by the group given the two-word descriptions. While all of the labels used (i.e. both the grape names and the two-word descriptors) were appropriate for the wines, only the descriptions were meaningful for these participants. It is interesting to note that the grape group could match the grape names to the wines with feedback, but not without. One possible suggestion is that they learned during each training block, but during the test block when no further learning was possible, performance dropped back to chance level.

The results also show that the participants can perform a wine identification task at a level significantly higher than chance, but only when given appropriate labels that are meaningful *for them*, which agrees with the findings from Chapter 5. The participants could not do this task when it was an orthonasal experiment (Chapter 6), but they can do so when the wine is tasted. Tasting a wine also brings additional cues, such as sweetness and carbonation detection (which would help with Lambrusco identification), palate weight and retronasal cues. None of the words in the descriptions referred to any of these cues, so they did not give the description group an unfair advantage. Thus, the difference between the grape and description groups in test must be due to the appropriateness of the labels

The particular case of the Lambrusco demonstrates the usefulness of labels that are appropriate to the respondent in terms of aiding identification, particularly once feedback is no longer present. While the Lambrusco should have been as easy to identify during test for the grape name group, performance on this wine was significantly lower during test, while no other differences were observed between the groups.

Thus, this particular wine may be driving the main differences between the groups, and also the significant drop in performance at test for the grape group.

That all of the wines were identified significantly higher than chance during all blocks with feedback (with the exception of the Shiraz during training block 2 for the description group) once again suggests that feedback is important to maintaining performance. The presence of the easier wine in the line-up and the drop in performance for the grape name group also suggests that the labels must be appropriate for the participants to be able to use them from session to session. However, if the respondents learned to use the carbonation of the Lambrusco to identify the wine, then the fact that the labels do not refer to the carbonation may challenge this assertion, although it is simple to learn that one of three samples has a distinctive feature, even if the label is not relevant.

Finally, the importance of taste (and other oral sensations) in terms of wine identification has been demonstrated in this experiment. When the same wines were used in Chapter 6, the participants could not perform significantly higher than chance on any of the wines, in any of the blocks, even with the same labels. Thus, the old adage that “90% of winetasting is in the nose” (Goodall & Eyres, 2013) may be overstating matters or, at the very least, does not refer solely to orthonasal olfaction.

Experiment 15 – Testing the Amount of Training Required for Novices to Learn to Identify Wine Flavours

The participants in Experiment 14 – and the experiments reported in previous chapters – received a relatively small amount of training. In contrast, wine experts receive many years of training. The next experiment was designed to determine whether more training is beneficial for the participants. A secondary aim was to test how much training the participants need to do well at this task. Thus, we recruited novice participants for a four-hour experiment, which was the maximum time allowed for experiments on first year Psychology students at the time.

Method

Participants. Ten first-year Psychology students (five females, aged 18-42, M = 20.8, SD = 7.5) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. Three different wines were used: Yalumba “Galway” Shiraz 2011, Yering Station “Mr Frog” Pinot Noir 2010 and Donelli Red Lambrusco (NV). These wines were the same as those used in Experiments 4 and 5. All blocks consisted of 18 trials, 6 of each wine. See page 57 for general information about how the samples were prepared.

Procedure. The four-hour experiment was split in 20-min sessions over twelve days, with at least one day between sessions. See page 59 for general information about the procedure used in this experiment.

Participants received twelve blocks of wines. Ten of these were training blocks, during which participants received feedback after their response, which indicated to participants which response was correct. Blocks 6 and 12 were test blocks, during which participants received no feedback, but the procedure was otherwise unchanged. The task was to identify each sample according to the given labels: Shiraz, Pinot Noir and Lambrusco. These labels were chosen as the participants in Experiment 14 were not able to maintain performance above chance with these labels in the test session, so if the present participants were able to do so, then it would be a good indication of learning.

Transfer effect. The last five participants (who were run as a second wave of participants) were also given an extra session where they were tested on an alternate set of wines made from the same grapes. During this final block, the participants were presented with six samples of each of the alternate set of wines (Plunkett Fowles “Stonedweller’s” Shiraz 2008, Bourke Street Pinot Noir 2011 and Luigi Cavalli Red Lambrusco NV) and no feedback was given during this block. Participants were not told that this was a different set of wines.

Results

Figure 9 shows mean number of items correct during each block of 18 trials. Participants were given grape names as labels, rather than descriptors, as this was considered a more difficult task.

Performance in all blocks was significantly higher than chance (six items correct), with the smallest *t*-value of $t(9) = 5.92$, $p < 0.001$.

The Lambrusco was easy to identify for the participants, as shown by a ceiling effect for that wine (Table 16), and, in this experiment, performance on this wine was maintained into the test blocks (blocks 6 and 12). However, performance on the Shiraz and Pinot also improved over the sessions. When compared to a chance score of two correct per block, performance on both wines was significantly better than chance on many blocks (Table 16). Note that the statistic (compared to chance) for Lambrusco could not be calculated in some blocks due to zero variance in the data.

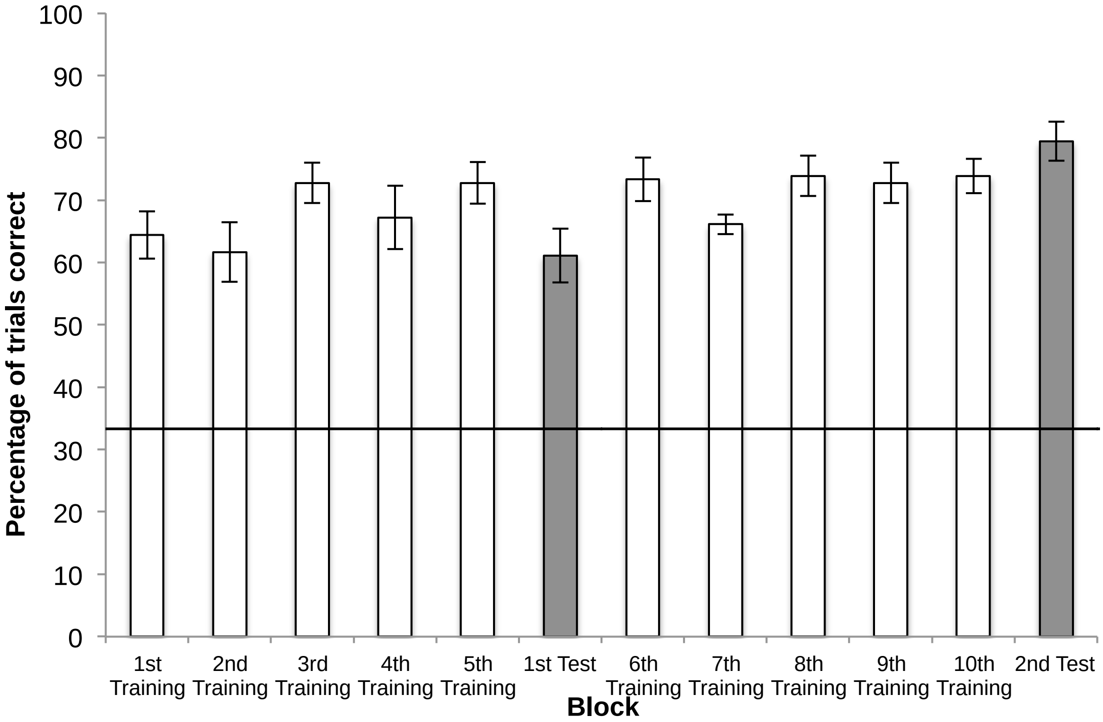

Figure 9. Mean (+/- SE) percentage of trials correct in each training and test block in Experiment 15.

Note: Feedback was given in the training blocks, but not in test blocks (darker bars). The horizontal line indicates chance level of responding (33%). Participants were asked to match grape names to the samples.

Due to the ceiling level of performance on the Lambrusco stimulus, the Pinot Noir and Shiraz were also compared to a chance score of three out of six correct. In this case, most blocks were not significantly higher than chance performance for either of these two wines, although in the final test block, both were significantly better than chance ($t(9) = 3.34$, $p = 0.009$ for Pinot Noir and $t(9) = 2.70$, $p = 0.024$ for Shiraz). Thus, the fact that the results reached significance suggests that the previous findings may not solely be a function of wine type.

Table 16. Mean (and SD) items correct by wine by block in Experiment 15.

Block	Shiraz	Lambrusco	Pinot Noir
Training Block 1	3.30* (1.34)	4.80* (1.03)	3.50* (1.27)
Training Block 2	3.30* (1.25)	4.90* (0.99)	2.90 (1.29)
Training Block 3	3.70* (0.48)	5.80* (0.42)	3.60* (1.65)
Training Block 4	2.90* (1.10)	5.40* (0.84)	3.80* (1.75)
Training Block 5	3.50* (0.85)	5.90* (0.32)	3.70* (1.57)
Test Block 1	2.40 (0.70)	5.80* (0.63)	2.80 (1.40)
Training Block 6	3.50* (1.35)	5.90* (0.32)	3.80* (0.79)
Training Block 7	2.60 (1.08)	6.00* (0.00)	3.30* (0.95)
Training Block 8	3.70* (1.06)	5.80* (0.42)	3.80* (1.03)
Training Block 9	3.70* (0.95)	5.90* (0.32)	3.50* (1.65)
Training Block 10	3.70* (1.06)	6.00* (0.00)	3.60* (1.43)
Test Block 2	4.10* (1.29)	6.00* (0.00)	4.20* (1.14)

Note: Each wine was presented six times in each block. An asterisk (*) indicates that performance for that wine was significantly higher than chance (two correct) or where all trials were correct in that block.

The training blocks were analysed with a trend analysis and showed a significant linear trend ($F(1,9) = 8.57$, $p = 0.017$), suggesting that learning did indeed take place over the course of the training blocks. No other interpretable trends were significant. However, when the same linear trend was tested individually for each wine, it was only significant for the Lambrusco, suggesting that most of the learning effect may have been due to this grape. Thus, with the benefit of hindsight, the use of this wine was possibly a mistake. This was addressed in the next experiment.

The change in performance between the test blocks was also analysed. On average, participants scored 11.0 items correct in the first test block (block 6, *SD* = 2.45), compared to 14.3 items correct in the last test block (block 12, *SD* = 1.77). This difference is statistically significant, $t(9) = 3.85$, $p = 0.004$. Performance on the Lambrusco samples did not improve significantly as performance was already particularly high. However, the increase in performance between the two test blocks (i.e. blocks 6 and 12) was statistically significant for both Shiraz, $t(9) = 4.02$, $p = 0.003$, and for Pinot Noir, $t(9) = 2.41$, $p = 0.039$.

Transfer effect. In the transfer test block, all of the five participants who completed this stage of the experiment were correct for all Lambrusco trials, despite the fact that this was tested on a different Lambrusco. In contrast, performance on the other new wines was 2.40 correct (*SD* = 1.67) for Shiraz and 3.60 trials correct (*SD* = 1.34) for Pinot Noir. Participants were not significantly higher than chance for Shiraz or Pinot Noir ($p = 0.621$ and 0.056 respectively) while the test for the Lambrusco could not be run due to lack of variance.

Performance in this additional block was also compared to performance in the previous block (i.e. block 12). On average, participants were significantly worse at the new Shiraz than the old Shiraz, $t(4) = 3.14$, $p = 0.035$, but not for the Pinot Noir.

Discussion

Despite the fact that the Lambrusco appears to be an easy stimulus to detect, a learning effect was found with these participants, as highlighted by the statistically significant linear trend. The results suggest a learning effect that was clearly driven by the increase in performance for the Lambrusco stimulus. This is most likely driven by the sweetness and slight spritz of this particular wine – elements that were not detectable by smell alone in previous experiments. Furthermore, this learning transferred to the new Lambrusco stimulus for the five participants that were given the extra test block.

Performance on the Pinot Noir and Shiraz improved between the two test blocks (blocks 6 and 12). This indicates that the extended training not only made the Lambrusco particularly easy to identify, but also improved test performance on the other wines in the study. Furthermore, it may indicate that repeated testing may be useful during training. Written and oral feedback from participants after previous experiments suggested that they used strategies during training that were not particularly useful during test, such as counting how many of each stimulus they had already experienced. This strategy is useless in the test condition in the absence of feedback. Thus, given extra testing, the participants may have had an opportunity to realise this and thus adjust their strategy.

However, there is evidence that this learning did not generalize to other wines made from the same grapes, with the exception of the Lambrusco. The

Lambrusco included features such as carbonation and sweetness that were probably more salient and easier to recognize compared to more subtle differences between the Shiraz and Pinot Noir. However, the result still indicates that certain elements of the wine can be learned and applied to a second wine from the same category.

The result is also somewhat contrary to that seen in Experiment 14 in that these participants did learn to use the word Lambrusco and maintained this into test. However, these participants had experienced more training before being tested, which may partially explain this result.

Learning to identify wine samples is generally seen as something that requires a lot of time. However, these results suggest that some basic learning can take place in a relatively short period of time, particularly for wines that are very different to others in the line-up.

Experiment 16 – Can Novices Transfer Their Learning of Identification of Wine Flavour to New Wines Made from the Same Grape?

Given that the results from these last two experiments may have been mostly driven by the Lambrusco, the aim of the next experiment was to determine whether similar results could be found using a wine that was more similar to the other two wines. As such, an experiment was devised where the Lambrusco was replaced with a drier, non-sparkling red wine. As Experiment 15 was both time-consuming and costly to run, Experiment 16 reverted to the task of applying short, two-word descriptions to the wines instead of grape names. An extra training block was added compared to Experiment 14 as a compromise. This procedure also served as a test of the validity of the labels used in the winemakers' tasting notes.

The secondary aim of this experiment was to test whether participants could generalize their knowledge to new wine samples that are made of the same grapes and fit the same two-word descriptions.

Method

Participants. Twenty-seven first-year Psychology students (12 females, aged 18 to 35, *M* = 19.6, *SD* = 3.3) participated in the experiment for course credit. See page 53 for general information about the participants. Two participants failed to attend all of the sessions of the experiment. Where possible their data are included in the following analyses.

Design. The experiment was a counterbalanced within-subjects design over three training blocks and one test block. Participants were trained on either of two sets of wines and then tested on both, so that one was a familiar set of wines and the other was novel, but related in that the same grapes were used.

Materials. Two sets of red wines were used. Each set contained a Shiraz, a Pinot Noir and a Cabernet Sauvignon. As the two-word descriptors were used, the wines were chosen so that the same descriptors appeared in the winemakers' descriptions of both wines. For example, both Shirazes were described as spicy and chocolate, with neither wine containing the descriptors appearing in the descriptions of either of the Pinot Noirs or Cabernet

Sauvignons. The two-word descriptors were: spicy and chocolate (Shiraz), blackcurrant (or cassis) and blueberry (Cabernet Sauvignon) and black cherry and gamey (Pinot Noir).

The first set of wines consisted of the Yalumba “Galway” Shiraz 2011, Yering Station “Mr Frog” Pinot Noir 2010 and McWilliam’s “Hanwood” Cabernet Sauvignon 2009. The second set of wines included the Plunkett Fowles “Stonedweller’s” Shiraz 2008, Bourke Street Pinot Noir 2011 and Lindeman’s “Bin 45” Cabernet Sauvignon 2011.

All blocks consisted of 18 trials, six of each wine. All trials consisted of approximately 10mL of one of the wines, presented in black plastic cups sealed with aluminium foil to prevent evaporation and to eliminate visual cues. See page 57 for general information about the preparation of samples used in this experiment.

Procedure. Participants signed up for a one-and-a-half-hour experiment, held over four separate days with a break of no more than four days between sessions. See page 59 for general information about the procedure used.

All participants received the same instructions and response options (the same two-word descriptors). Approximately half of the participants were trained on one of the sets, while the other half were trained on the other set. Participants were randomly allocated to one of these counterbalanced groups.

Participants received four blocks of wines. The first three blocks were training blocks, during which participants received feedback after their response, which indicated to participants which response was correct. The final block was a test block, during which participants received no feedback. The task was to identify each sample according to the given labels, which were: “Spice and chocolate” for the Shiraz, “Black cherry and gamey” for the Pinot Noir and “Blackcurrant and blueberry” for the Cabernet Sauvignon.

During the test block, both groups were tested on all wines used in the experiment. Of the 18 test trials, each of the six wines was presented three times. Thus, for each participant, half of their test trials were on the wines on which they were trained and half were on the alternate set.

Results

No significant learning effect was found over the training blocks, $F(1,25) = 2.67$, $p = 0.12$. However, performance in all blocks was significantly higher than chance, with the smallest *t*-value of $t(26) = 2.79$, $p = 0.01$ for block 1.

Performance on Pinot Noir was significantly better than chance throughout training and test, with the smallest *t*-value of $t(26) = 2.47$, $p = 0.021$, while performance on Shiraz was significantly higher than chance only in the final training block and the test block, $t(25) = 3.23$ and $t(24) = 4.11$ respectively, both $p < 0.01$ (Figure 10). Performance on Cabernet Sauvignon was not significantly higher than chance in any of training blocks (Table 17).

In the final test block, participants were significantly better than chance for both the Shiraz and Pinot Noir on which they had been trained, $t(24) = 3.06$ and 3.92 respectively, $p < 0.01$, but not on the alternate Shiraz and Pinot Noir, $p > 0.05$. Participants were significantly better than chance for the

Cabernet on which they were trained ($t(24) = 2.06$, $p < 0.05$), but not on the "new" Cabernet, $p > 0.05$ (Table 18).

However, the difference between the old and new stimuli was only significant for the Pinot Noir, with respondents performing significantly better on the "old" stimuli than the "new" stimuli, $t(24) = 3.00$, $p = 0.006$.

Table 17. Mean (and SD) items correct by wine by training block in Experiment 16.

Training Block	Shiraz	Cabernet	Pinot
1	2.41 (1.05)	2.15 (1.10)	2.48* (1.01)
2	2.56 (1.48)	2.48 (1.34)	3.15* (1.46)
3	2.88* (1.40)	2.31 (1.52)	2.77* (0.95)

Note: Six samples of each wine were presented each block. * indicates that performance for that wine was higher than chance (two correct) in that block.

Table 18. Mean (and SD) items correct by wine during the test block in Experiment 16.

Set	Shiraz	Cabernet	Pinot
Old	1.56* (0.92)	1.28* (0.68)	1.76* (0.97)
New	1.20 (0.76)	1.04 (0.89)	1.00 (0.82)

Note: Three samples of each wine were presented each block. The "old" set of wines is the one on which the respondents were trained, while the "new" set of wines is the one on which the respondents have not seen before. *

indicates that performance for that wine was higher than chance (one correct) in that block.

Figure 10. Mean (+/- SE) percentage of trials correct in training and test blocks by wine in Experiment 16.

Note: Feedback was given in the training blocks, but not in the test block. The horizontal line indicates chance level of responding (33%). Participants were asked to match descriptors to the samples. The results for the test block are split into old stimuli (wines on which the participants were trained) and new stimuli (wines on which the participants were not trained).

Discussion

While participants were significantly better than chance throughout, especially with the Shiraz and Pinot Noir, performance was not phenomenal.

This is a particularly difficult task for novices, even with the two-word descriptions (which was more useful for them in Experiment 14) instead of the grape names.

The finding that performance was above chance could be driven by the fact that participants find two of the wines quite similar and one relatively dissimilar and therefore easier to identify. In this case, the wine that was identified correctly the most was the Pinot Noir. However, in Experiments 14 and 15, the same Pinot Noir was confused with the same Shiraz used here. This suggests that the context of the other wines in the set may be an important factor when learning about wines.

The descriptors used were chosen on the basis of one person’s description of each wine and the present experiment serves as a test of the validity of these labels, similar to the experiments in Chapter 5. The results suggest that the labels used by the winemakers may not be particularly appropriate for the wines, at least for novices, given that they could not transfer the use of the labels to the new wines, despite learning to use them for the samples on which they were trained.

The winemaker’s tasting notes was chosen as, presumably, they know their wine quite well and have tasted them many times. However, the winemaker also has an interest in selling their wine and may therefore describe the wine in a way that is relevant for wines made of that grape but not necessarily of that particular wine. To counter this potential problem, independent descriptions of each wine (tasting notes written by prominent Australian wine critics) were also checked to ensure that each wine was considered to be a good representation of wines made from that grape.

Admittedly, the wines used were relatively cheap due to budget restrictions and may thus not have been the best examples of their style.

Despite these potential problems with the samples, the participants were still able to perform significantly better than chance and to maintain this into test despite only having three training sessions.

Summary

The experiments in this chapter demonstrate that the use of flavour is important when it comes to wine training. The problem with doing so is that training takes much longer to complete, as the alcohol in the wine has a cumulative effect on the taster, even if the wine is spat out during tasting.

Furthermore, extended periods of training do appear to lead to an increase in performance, but that increase is not very large apart from one sample that appears to have been quite easy to learn when flavour was involved in the experiment.

Finally, whether or not the participants can transfer learning to a new set of wines appears to depend on how well they have learned salient features of the original wines. In Experiment 15, the participants were very good at identifying the Lambrusco wine and what they learned of that wine transferred to a different Lambrusco. However, this was not the case for either of the other wines in that experiment, nor was it the case for any of the wines in Experiment 16. Thus any transfer finding is probably due to the fact that Lambrusco has very easy to detect features such as sweetness and carbonation, while the differences between the other wines are not as easy to detect.

CHAPTER TEN – SENSORY TRAINING WITH AND WITHOUT CONTEXTUAL INORMATION

Experiment 17 – The Effect of Training Using Contextual Wine Information on Wine Flavour Identification

As described in Chapter 5, people can generally identify odour mixtures using labels for a unique element in the mixture, but only when the label is appropriate for that unique element (Chapter 5). When tested on the same procedure using wine samples (Chapter 6), the participants were able to identify wines using either labels that refer to the unique elements (e.g. spicy for Shiraz, blackcurrant for Cabernet Sauvignon) or using labels that require some knowledge about wine (e.g. Shiraz, Cabernet, Pinot Noir), but not to the same extent seen with odour samples. The exception was when the participants were tested using flavour on a set with a particularly easy wine: Lambrusco. This wine stood out, as it was sweeter and slightly spritzy. Despite this, with only a relatively short amount of training, the participants could not put the word Lambrusco to this wine at a level above chance once feedback was no longer present.

The sensory labels used in the previous experiments only referred to specific odours and flavours (e.g. chocolate, blackcurrant, raspberry, dark cherry) but not other sensory aspects of a wine, such as mouthfeel information (astringency, acid) or “length” (duration of aftertaste). For one group, these sensory elements were added to the feedback in this experiment.

But detecting these particular sensory cues in a stimulus as complex as wine is still a much more difficult task than detecting the unique odour in a binary odour mixture, such as those used in Chapter 5. As discussed in Chapter 3, many wine training courses also include conceptual knowledge about wines, including information about wine regions, styles, grapes and winemaking techniques. The question then is whether conceptual information is useful in terms of wine training.

Experts are able to draw on their conceptual knowledge of wine in general. However, this does not always work to the expert's advantage. Morrot et al. (2001) found that if white wines are presented with red colouring, experts are more likely to describe them using adjectives that relate to red wines, whereas novices do not do so. In this case, it appears that the expectations of the experts influenced their perception through top-down processes. In that particular study, the aim was to determine whether conflicting sensory information had an effect on wine descriptions. However, it raises the question of whether non-conflicting conceptual information may be useful to participants who are attempting to learn perceptual information.

Previous studies have found that conceptual knowledge has an influence on perceptual processing in a visual paradigm (Curby & Gauthier, 2010; Gauthier, James, Curby, & Tarr, 2003). Gauthier et al. (2003) arbitrarily assigned artificial semantic concepts to novel visual objects and found a direct effect of these semantic associations on visual object recognition. This effect has also been studied in a wine context by LaTour et al. (2011), where participants were given perceptual training (smelling wines with aroma additives and identifying the most prominent odour) and conceptual training in the form of a Powerpoint slideshow containing topics about wine production and discussion of wines made from the grape in question. They found that conceptual training was important because it provided the first steps for their novices in terms of learning the particular vocabulary involved in wine tasting.

One group in the present experiment was given conceptual information, such as the region and vintage from which the wine was sourced and a description of the hallmark characteristics of wines from that particular region and vintage, in an effort to aid the participants. They were also given additional sensory information as described earlier, such as information about aftertaste persistence or weight of body (see Appendix B).

As the identification task used in previous chapters had indicated that the participants found it relatively difficult to apply labels to the wine samples, a non-verbal categorization task was also used in this study.

Method

Participants. Thirty-five first-year Psychology students (22 females, aged 18 to 27, *M* = 18.9, *SD* = 1.9) participated in the experiment for course credit. See page 53 for general information about the participants.

Design. The experiment was a 2 (group) x (5) (block) mixed design consisting of four training blocks and one test block split over five days for the identification task. The experiment was also a 2 (group) x (2) (test/re-test) design for the categorization task. These two sessions took place on the first and last day of testing. Participants were trained on two wines for the first four blocks, while in the final session (the test block), participants were tested on those two wines and two similar but novel wines. The wines on which they were trained were counterbalanced across participants.

Materials. Two sets of red wines were used. Each set consisted of a Shiraz and a Pinot Noir. Both Shirazes (Yalumba “Galway” Shiraz (2011) and Tanunda Hill Shiraz (2011)) were from the same region (the Barossa Valley), while the Pinot Noirs (Yering Station “Mr Frog” Pinot Noir (2012) and Punt Road “Emperor’s Prize” Pinot Noir (2010)) were from the same region as each other (the Yarra Valley), but not the same region as the Shirazes. The Shirazes were selected on the basis that the winemaker’s description featured similar sensory information (odours and flavours, mouthfeel characteristics) to each other. The Pinot Noirs were selected on the same basis and also under the constraint that the sensory information contained in the winemakers’ descriptions was different to that in the descriptions of the Shirazes. All trials consisted of approximately 10mL of one of the wines, presented in black plastic cups sealed with aluminium foil to prevent evaporation and to eliminate visual cues.

The timing, response collection, feedback and quasi-random order of trials (no stimulus was presented more than three times in a row) were determined by a computer script running in Inquisit 3.0 (Millisecond Software LLC, 2011). Participants were required to give their answer before spitting each sample out, so that they did not receive visual cues from the spittoon.

Procedure. Participants signed up for a two-hour experiment, held over five separate days. Up to three participants attended each session. Participants completed consent paperwork, the AWKQ and a brief olfactory screening task (Flowers or Pear – four trials) during the first session, as in previous experiments.

Prior to any wine training, participants completed a wine categorization task. Respondents were given 8 wines (two of each of the four wines) and asked to categorise them into two categories based on flavour. No other information was given; including how many wines belonged in each category. Respondents assigned a wine to a category before spitting it out and were not allowed to reassign the wine to another category after spitting, due to the possible confounding problem of visual cues from the spittoon. The first wine that they tasted was automatically assigned to category A. The participants were then required to assign subsequent wines either to category A or category B.

Participants were randomly allocated to one of two groups: the "sensory-only information" group and the "sensory and conceptual information" group. Half of the participants within each group were trained on one of the sets, while the other half were trained on the other set. Participants were randomly allocated to one of these counterbalanced groups. Both groups and both counterbalanced conditions were tested on both sets of wines in the final test block.

Participants then received the first of five blocks of wines. The first four blocks were training blocks, during which participants received feedback after their response to the question "Which wine did you just taste? Shiraz or Pinot Noir?", which indicated to participants which response was correct. Each training block consisted of 20 trials: 10 of each of the two wines.

Feedback was different for the groups. The “sensory only information” group received feedback such as “Wrong! That was Shiraz, the spicy and chocolatey wine” where the only information they received was the label with which they responded. The “sensory plus conceptual information” group received more detailed feedback, including information about the region in which the grapes were grown, the vintage conditions, and other information. The feedback options were all presented in each training session in a random order. See Appendix B for a list of the feedback used throughout the training blocks.

The final block was a test block, during which participants received no feedback. During the test block, both groups were tested on all four wines (both Pinot Noirs and both Shirazes) used in the experiment five times each, for a total of 20 trials. Thus, for each participant, half of their test trials were on the wines on which they have been trained and half were on the alternate set.

During the final session, participants also performed the wine categorization task a second time as a retest, to determine whether performance improved.

Scoring for the training and test blocks was how many of each wine were given the correct label (Shiraz or Pinot Noir). Scoring for the categorization task was based on how many wines in each category were of the same grape.

Analyses. For the identification task, the groups were compared in terms of number correct in each block. Each group was also compared to chance (ten correct) at each block. A 2 x (4) mixed model ANOVA (with appropriate simple effect contrasts) tested for differences between the groups over training, an overall training (linear and/or quadratic) effect, whether each group improved between training blocks and whether the change in performance between training blocks was different for the groups.

The main findings of interest were a) whether participants can do the identification task with either of the sets of information (compared to chance), b) whether participants could learn to do the task over time, given more training than in most of the previous experiments, c) whether one group of participants learned at a greater rate than the other and d) whether participants could generalize their knowledge from the set on which they were trained to a new set.

Results

Compared to chance. Both groups performed significantly better than chance (10 items correct) in the fourth training block and in the test block (with the smallest t-value of $t(16) = 3.31$, $p < 0.01$ for the sensory information only group in block four). In addition, the sensory information only group was also significantly higher than chance in the third session ($t(16) = 3.77$, $p = 0.002$) and the sensory & contextual information group were significantly higher than chance in the second training session ($t(17) = 4.42$, $p < 0.001$). All of these results remained significant when a Bonferroni correction was applied.

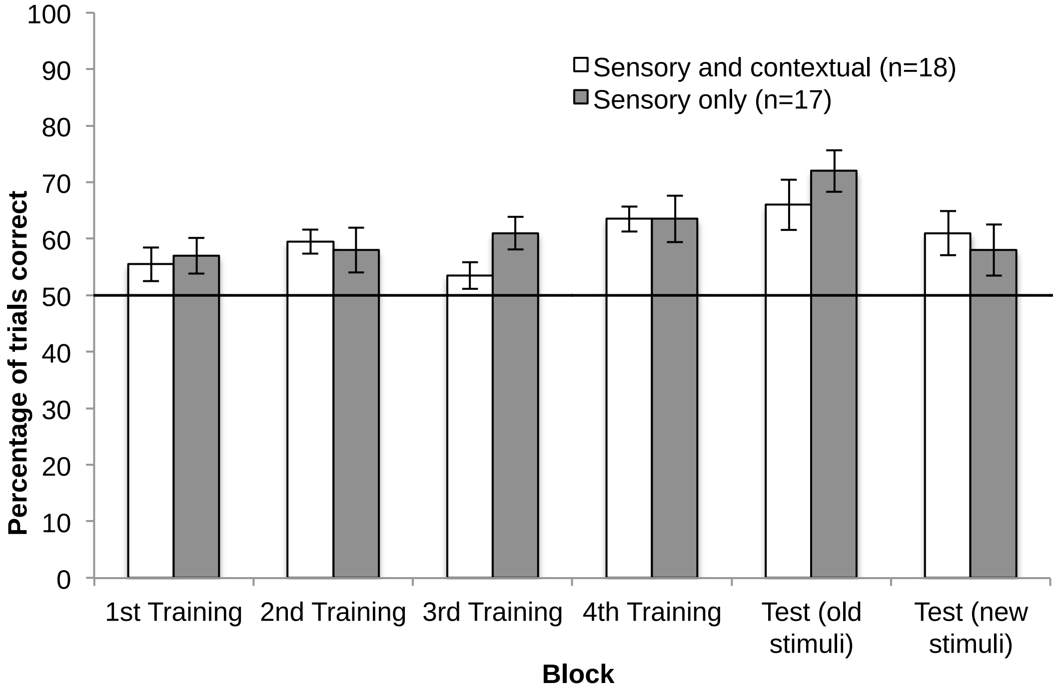

Figure 11. Mean (+/- SE) percentage of trials correct in each training and test block by group in Experiment 17.

Note: Feedback was given in the training blocks, but not the test block. For the test block, the old and new stimuli are presented separately. The horizontal line indicates chance level of responding (50%). Participants were given either sensory only information (Sensory only group) or sensory and contextual information (Contextual and sensory group).

Learning effect. A 2 (group) x (5) (block) ANOVA was conducted to determine whether performance improved over time. A significant positive linear trend was found averaged over both groups ($F(1,33) = 10.33$, p = 0.003) indicated that, overall, participants in the experiment displayed some learning. Tests of simple effects showed that this linear trend was significant for both groups ($F(1,17) = 5.94$, p = 0.026 for sensory and contextual information and *F*(1,16) = 6.48, *p* = 0.022). The linear trend did not differ significantly between the groups (*F*(1,33) = 0.03, *p* = 0.88) and no other trends below a cubic result were significant over the whole experiment or for either group. Finally, no significant difference was found between the groups averaged over block (*F*(1,33) = 0.52, *p* = 0.48).

When the test block is removed from the analysis, the same pattern of results as above was found, except for the significant linear trend for the sensory and contextual information group.

Transfer effect. Both groups performed significantly better than chance in the test on the wines on which they had been trained (with the smallest *t*-value of *t*(17)3.63, *p* = 0.002). The group trained on sensory and contextual information also performed significantly better than chance (five trials correct) on the wines on which they had not been trained (*M* = 6.06, *SD* = 1.66), *t*(17) = 2.70, *p* = 0.015, while the sensory information only group did not perform significantly better than chance on the wines on which they were not trained (*M* = 5.76, *SD* = 1.86), *t*(16) = 1.70, *p* = 0.109. Thus, the sensory and contextual information group showed a transfer of knowledge to the new stimuli, while the sensory only group did not (Figure 11). However, the groups did not differ significantly on performance on either the new or old stimuli.

Categorisation task. The pre- and post-training categorisation tasks were scored separately. Each task involved the sorting of eight wines (two different Shirazes and two different Pinot Noirs, each presented twice). A pair-wise dissimilarity estimate (distance) was calculated for each possible wine pair based on how many of the 35 respondents placed the two wines in the pair in different categories. This method is essentially similar to that used by Ballester et al. (2008), except that dissimilarity estimates were used instead of similarity estimates. A matrix of distance estimates between each possible pairing of wines were then run through a multidimensional scaling procedure (PROXSCAL), treating the data as ordinal.

The scree plot and fit statistics (stress, Dispersion Accounted For and Tucker's Coefficient of Congruence) indicated that there were two dimensions in the dissimilarity matrix for both pre- and post-training categorisation tasks. The fit statistics for the pre-training categorisation task were normalized raw stress of 0.013, Dispersion Accounted For (D.A.F.) of 0.987 and Tucker's Coefficient of Congruence of 0.993. For the post-training categorisation task, the normalized raw stress was 0.003, D.A.F was 0.997 and Tucker's Coefficient of Congruence was 0.999. Good fit is indicated by stress values closer to 0 and lower than 0.15 is considered acceptable (Borg, Groenen, & Mair, 2012). Conversely, D.A.F. and Tucker's Coefficient of Congruence values closer to 1 indicate good fit (Lorenzo-Seva & ten Berge, 2006).

Each wine was then plotted according to the coordinates on each dimension based on the results of the cluster analysis. These are shown in Figure 12 and Figure 13. Two subsequent hierarchical cluster analyses were conducted on these coordinates, one for each of the pre-training and post-training categorisation tasks, using squared Euclidean distances. Post-training, the cluster analysis clearly indicated two clusters, which perfectly split the Pinot Noir and Shiraz wines. The pre-training cluster analysis was less clear and indicated either two or three clusters. In the two-cluster solution, one cluster included three wines (all Shirazes) while the other cluster includes five wines (the remaining Shiraz and the four Pinot Noirs).

Taken together, these results suggest that the participants could mostly separate the wines into two different categories based on grape type (dimension 1) prior to training, and this result was more clear after training. However, a second dimension was also present in the data, indicating a secondary source of variance.

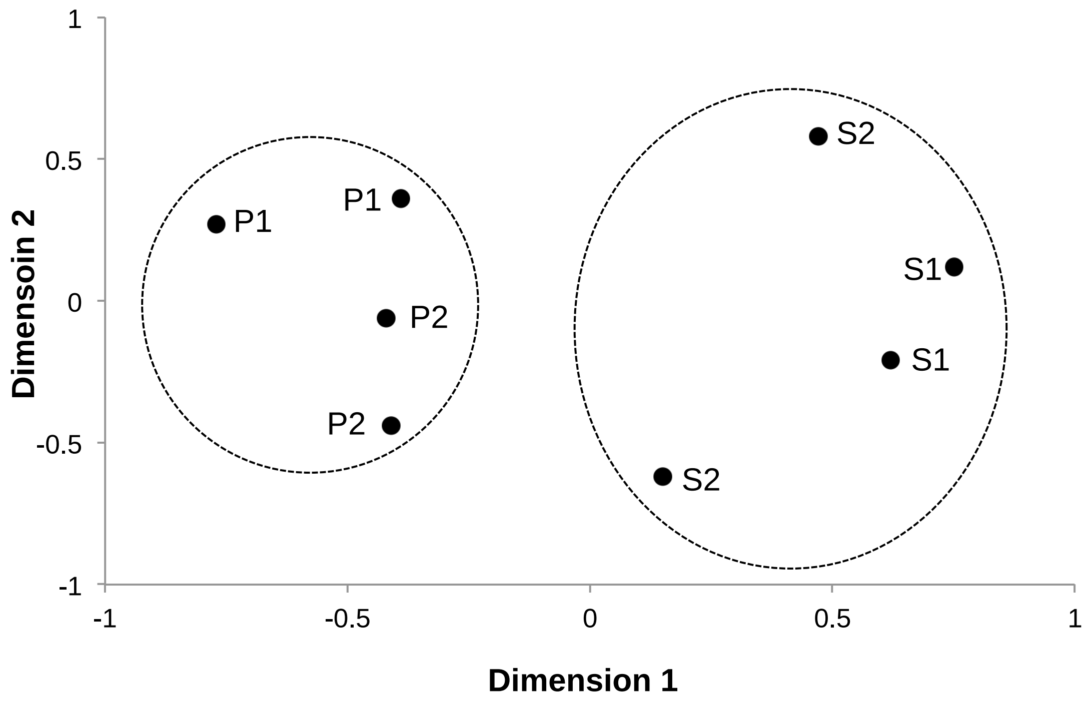

Figure 12. Location of each wine prior to training in the two dimensions indicated by multidimensional scaling in Experiment 17.
Note: Dotted lines indicate clusters according to a subsequent hierarchical cluster analysis. S1 and S2 are the Shirazes and P1 and P2 are the Pinot Noirs (each wine was identified twice).

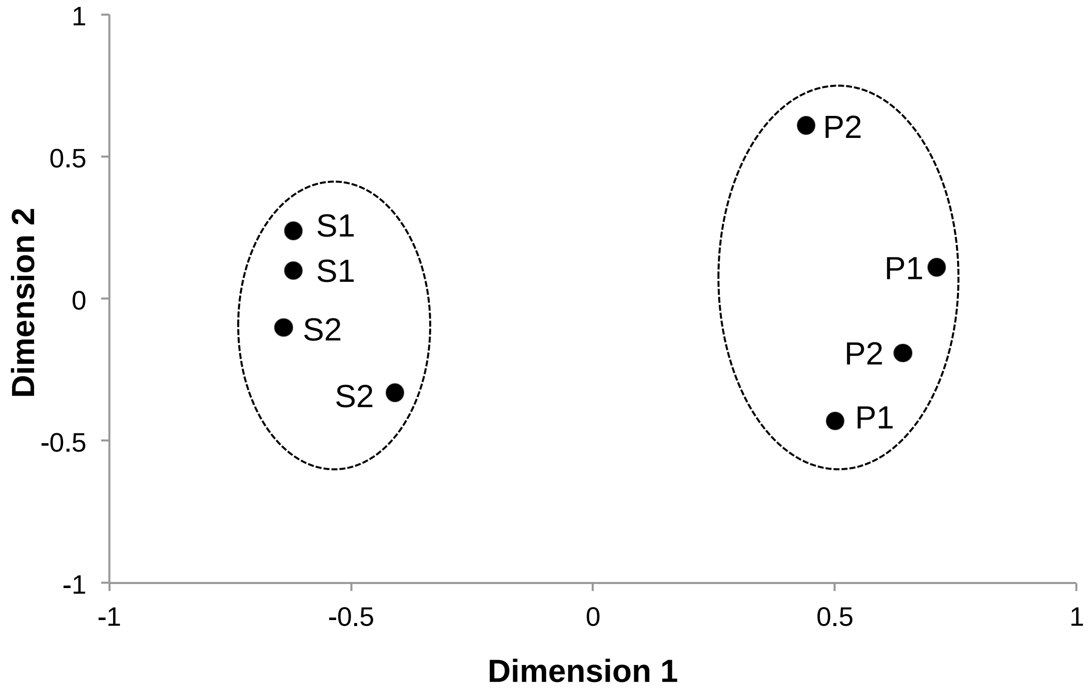

Figure 13. Location of each wine after training in the two dimensions indicated by multidimensional scaling in Experiment 17.

Note: Dotted lines indicate clusters according to a subsequent hierarchical cluster analysis. S1 and S2 are the Shirazes and P1 and P2 are the Pinot Noirs (each wine was identified twice).

Discussion

The significant learning effect found in this experiment indicates that novices are able to learn to identify wines based on flavour using the given labels. The response options were the same for each group (spicy and chocolate for the Shiraz and black cherry and earthy for the Pinot Noir).

However, no significant differences were found between the conditions in any of the blocks and, overall, the pattern of results was very similar in both groups, suggesting that giving contextual information made virtually no difference to performance. This information included grape names, the name of the region from which the grapes were grown, information about the region (e.g. climate) and the effects that each climate has on the final flavour of the wine.

The participants showed some learning in terms of using these labels for the wine samples, suggesting that the use of sensory descriptors is actually a useful tool, compared to all of the other experiments in which grape names were used. In these other experiments, learning was only found when 3.5 hours of training was administered. Thus, there is some evidence that these may be useful labels for training novices.

Furthermore, both groups performed significantly better than chance in either the second or third training session and both were significantly better than chance in the fourth session, indicating that the participants started to learn to apply these labels in a relatively short period of time.

However, the categorisation task added new information not seen in previous experiments. The pre-training multidimensional scaling and hierarchical cluster analysis indicate that, despite being novices and having no prior exposure to these wines within the experiment, the participants could categorise the wines based on grape type. After training, the wines formed perfect clusters based on grape type based on the participants' categorisation. This task was a non-verbal task. This result, taken with the findings from the wine discrimination experiments in Chapters 5 and 6 indicate that the participants were able to not just discriminate between wines, but also categorise them based on grape type. Despite this, learning to apply a label for wines in each cluster is a task that novices do not find as easy.

Summary

The additional information given to the sensory plus contextual group made little difference to their performance compared to the sensory group. However, both groups managed to learn to apply the given labels (Shiraz and Pinot Noir) to the wines at a rate above chance by the end of training and maintained performance into test, indicating some learning of the connection between these labels and the wine odour properties. Furthermore, these labels appear to be appropriate for novices given that they did learn to use them, unlike the Irrelevant or Inappropriate groups in Chapter 5. This task was somewhat easier than previous experiments, as only two different types of wine were used.

Some evidence of transfer of learning to the new wines was found amongst those who received sensory and contextual information, as that group performed significantly higher than chance on the new wine stimuli, while the sensory only group did not. However, as the groups did not differ significantly on performance on the new stimuli, this result is less clear. It should be noted that the “new” stimuli had been previously encountered in the pre-training categorization task, although they had not been trained on these wines and had only tasted each one twice.

The categorization task provided a new perspective on the results. While it was not the case for every participant, it appears that over the whole sample, these participants were able to sort the wines into clusters based on grape type, even prior to any training. Taken together with the triangle test results from Chapters 6 and 7, the results suggest that novices are able to both discriminate between and categorise wines. Furthermore, given enough training, they can learn to apply labels to those categories, even if the labels refer to the grape names and mean relatively little to them. Although the participants still made errors, the results suggest that the restriction for novice performance on a wine identification task may not be a perceptual task, but instead a task of learning to use the established language used by experts. Thus, the perceptual properties of the wines appear to be perceivable by novices. The task for novices is to learn which words are used to apply to these perceptions.

CHAPTER ELEVEN – GENERAL DISCUSSION

Overview of results

The overall aim of these experiments was to examine potentially important factors that affect training of identification of wine samples. The first step in doing this was to determine which words/labels can be used for identification purposes.

The first three experiments reported in Chapter 5 used binary odour samples to determine the effects of labels in an identification task. In these experiments, it was found that the connection between the label and the odour was an important factor in identification. Those in the Appropriate group (where the veridical label was associated with the stimulus) performed well above chance throughout the three experiments. Other label types included Self-generated (Experiment 1), Inappropriate (where the correct labels were associated with the incorrect stimulus, Experiment 2) and Irrelevant (where unrelated labels were associated with the stimuli, Experiment 3). The Self-generated labels were used with above-chance performance by novices after training, while the Inappropriate and Irrelevant groups performed significantly better than chance in some blocks. In general the Appropriate groups performed better than the other groups (with the exception of the self-generated group in Experiment 1 after one training block). Thus, the connection between the odour and the label is an important factor in terms of odour identification, in agreement with Cain (1979), although it also appears that participants can sometimes be trained to use labels that are not directly related to odours for identification, even with relatively little training.

The novel results from the odour experiments in Chapter 5 raised the question of which labels would be appropriate for novices to use in a wine identification task. This was addressed in Chapter 6, where the first step was to find a set of wines that the novices could discriminate (Experiments 4 and 6). This was followed by identification experiments using the same procedure as the odour experiments in Chapter 5 (Experiments 5 and 7). In all of these experiments, the participants smelled the wine samples, without tasting them. The two label types used in Experiment 5 were grape names (Shiraz, Pinot Noir and Lambrusco) or descriptors (spice and chocolate, black cherry and gamey, floral and raspberry). Both groups were able to perform at a level significantly higher than chance by the second training block, but did not perform at the same level as that of the Appropriate groups in Experiments 1-3. In Experiment 7, the task was to create a label for the wine using the Wine Aroma Wheel (Noble et al., 1987), either by comparing two samples and attempting to describe the unique odours from each using the wheel as a prompt, or by simply choosing words appropriate for each wine without any direct comparison. Very few of the descriptions were considered to be "accurate" (i.e. match descriptors in the winemakers' descriptions) and neither of the groups could use their own labels to identify the wines, even after training. It appears that applying labels to wines is more difficult than applying labels to odours. This appears to be the case both when the labels refer to the stimulus as a whole (grape names) or to an important component of the stimulus (descriptors of elements, e.g. pepper). Furthermore, participants appear to be unable to describe wines in a way that they can then use to identify the wines, consistent with previous non-wine studies (e.g. Lawless, 1984).

The aim of the experiments reported in Chapter 7 was to determine whether another set of wines could be found that were easier for the participants to discriminate using the triangle test procedure. A set of white wines (Experiment 8) and a set consisting of both red and white wines (Experiment 9) were tested in the absence of visual cues and found to have similar or worse discriminability to the wines used in Experiments 4 through 7. One particularly interesting finding was in Experiment 9, where participants made the most discrimination errors on trials involving a red and white wine and the least errors on trials involving two white wines, indicating that they were unable to discriminate between red and white wines based on smell alone. One final experiment in Chapter 7 (Experiment 10) used a similar methodology to Experiments 8 and 9 (the triangle test) to determine whether the mere presence of the appropriate labels also had an effect on discrimination, as well as identification performance, using the odour samples from Chapter 5. One group was told the identity of the unique element in each of the odour samples while the other was not. No significant difference was observed between groups. These experiments, and other similar experiments reported in the appendices, indicate that the wines used in Chapter 6 were one of the most discriminable sets tested and that the use of any other set would most likely have resulted in similar or worse identification performance. Furthermore, while the presence of Appropriate labels improves identification performance, they do not improve discrimination performance, indicating that there are some limitations to their top-down influences.

A series of experiments is reported in Chapter 8, where the aim was to determine whether a) the commonly accepted varietal characteristics of the wine samples (e.g. that a Shiraz is peppery) was evident to novices (Experiment 11), b) increasing the concentrations of these elements increased their identification using line scales (Experiment 12) and c) whether initially presenting the participants with “enhanced” samples increased their ability to identify the wines using descriptor labels once the enhancements were removed (Experiment 13). The results suggested that a) the participants did not appear to identify the commonly accepted varietal characteristics (that is, they did not rate the Shiraz as significantly more peppery than any of the other red wines and so on for the other wines), b) that increasing the concentration of these elements using additives did result in significant and predictable changes in how the wines were rated and c) that enhancing the odours of the wines did not lead to any increase in performance in the test session compared to those who had been trained on unadulterated wines throughout.

The experiments in Chapter 9 included numerous alterations to the training procedure used in previous chapters. Notably, participants were now able to taste the wine, rather than just smell the samples, as had been the case in previous chapters. Experiment 14 tested the same question as Experiment 5, whether grape names (Shiraz, Lambrusco, Pinot Noir) or abbreviated winemaker descriptions (spice and chocolate, floral and raspberry, black cherry and gamey) were appropriate labels for participants in terms of wine identification. During training, both groups performed at a similar level, but in test, only the Description group performed significantly higher than chance (and significantly higher than the Grape Name group). Experiment 15 addressed the question of whether more training was helpful when learning to use grape names, with results suggesting that the participants were able to use the grape names at levels above chance with less than four hours of training. This experiment included a test session after half of the training blocks as well as a second test session at the end of training. Performance significantly improved from the first to second training session, indicating that multiple testing sessions may be a useful technique. Furthermore, in both of these experiments, the participants were able to transfer at least some of their learning to new stimuli. Experiment 16 used an alternate set of wines in order to address the potential problem that one of the wines (the Lambrusco) was too easy to identify by flavour in Experiments 14 and 15. The participants in Experiment 16 were also able to identify this new set of wines at a rate significantly better than chance using descriptors over three training sessions and in test, but were not able to transfer their learning to a new set of wines.

Finally, Experiment 17 (reported in Chapter 10) tested whether training on other aspects of the wine, such as the region from which it came and aspects of the wine that are typical of wines from that region (contextual information) along with the descriptors (sensory information) was superior to being trained on sensory information alone. Both groups were able to use grape names by the fourth (final) training session and in test, with some evidence to suggest that this learning transferred to new wines, at least for the sensory and contextual information group. No significant differences were found between the groups. However, a secondary task, in which the participants were asked to categories the two Shirazes and two Pinot Noirs into two groups before and after training suggest that the participants could indeed discriminate the wines based on grape type, even before training had commenced.

Discussion of results

Wine discrimination performance

The aim of the discrimination experiments was not to determine whether training could improve discrimination performance. Participants were not given any feedback during these experiments. The main aim of these experiments was to determine whether the participants could discriminate between the wines. If they could not do so, then identification of the wines would not have been possible. All discrimination studies were conducted orthonasally.

In Experiment 4 (Chapter 5), the participants could discriminate between any combination of the three red wines at a level above chance, indicating that they were suitable for further studies. However, when three white wines were used (Experiment 8), the participants could not discriminate between them. Furthermore, when a mixture of red and white wines was used (Experiment 9), it appears that most of the errors occurred for trials involved the Chardonnay and Shiraz – a white and red wine, respectively. Furthermore, the other white wine used in Experiment 9 (Sauvignon Blanc) appears to have been discriminable from both the other white wine and the red wine. This finding is most likely due to perceptual similarities between the Chardonnay and Shiraz that arose from similar winemaking techniques, such as the use of oak in both of these wines, but not in the Sauvignon Blanc. Thus, it is possible that the elements that were the most salient to novices were present in both of these wines, but not in the Sauvignon Blanc. Furthermore, this challenges the findings by Ballester et al. (2009), who found that novices, as well as experts, could categorise (and thus discriminate between) red and white wines when visual cues were removed. This suggests that any such findings may be specific to the stimuli used and any perceived perceptual overlap between them. Unlike stimuli in other senses, perceptual overlap cannot easily be quantified and must be tested using humans, indicating the importance of pilot testing.

In Experiment 10, the aim was to determine whether the presence of correct labels could enhance discrimination performance via top-down processes. The results suggest that the presence of labels made virtually no difference to discrimination performance, indicating that the presence of labels most likely does not guide the participants' olfactory search of the odour mixture and therefore does not help a discrimination task, which is not a verbal task. Instead, the presence of the labels appears to help only with identification performance, as discussed below.

Thus it appears that discrimination training may not be particularly helpful to novices, as it appears that they already possess the ability to discriminate between wine samples.

Wine categorisation performance

In Experiment 17, the participants were able to categorise the wines based on grape type, even before any training. This finding is in contrast to

Ballester et al. (2008), who found that novices were unable to sort 10 Chardonnays and 10 wines made from the Melon de Bourgogne grape into categories based on grape type, whereas their experts could. One possible explanation for this is that the grapes used by Ballester et al. (2008) share some common characteristics and may thus be quite similar. Wines made with the Chardonnay and Melon de Bourgogne grapes are often made using similar techniques, such as lees aging (Robinson, Harding, & Vouillamoz, 2012). In contrast, the wines in Experiment 17 were made from Shiraz and Pinot Noir, which are generally considered to produce quite distinct wine styles. Furthermore, Ballester et al. (2008) required categorisation of 20 wines (10 made from each grape) into as many categories as the participant liked, whereas the participants in Experiment 17 only tasted eight samples (two different Shirazes and two different Pinot Noirs, each presented twice) and were instructed to sort them into two categories.

Another difference between the experiments is that Ballester et al. (2008) participants only smelled the wines, whereas the participants in Experiment 17 tasted the wines. Given these differences, the results may not be directly comparable, but the present finding that the participants could sort wines into categories based on grape type, even before any training, adds further weight to the conclusion that novices can detect differences between wines and that expert performance in wine identification is not merely due to perceptual advantages.

The implications of this finding for the learning of expertise is that it appears that the novices can categorise at least some wines based on grape without any training type and that, when the difference between the sensory profiles of the grape types is large enough, the varietal-based categorization that experts appear to use (e.g. Hughson & Boakes, 2002a) may also be useful to novices.

The appropriate label effect

The label applied to an odour can have strong effects in terms of hedonics and recognition memory, as discussed at the end of Chapter 5. Furthermore, the presence of labels in visual perception experiments (e.g. ‘young girl’ or ‘saxophone player’) can, via top-down processes, bias perception towards one of the interpretations of an ambiguous figure, perhaps due to the label influencing attention (Cavanagh, 1999). The results from Experiment 10 suggest that the mere presence of labels may not bias attention in the same way for olfactory stimuli, at least not for discrimination tasks. Instead, it appears that the labels make the identification task easier by giving the respondents options to choose from, rather than guiding the participants to the salient elements of the odour mixtures. This is further supported by the Inappropriate group (Experiment 2), who were given the correct labels, but trained to match them to the incorrect samples. If the mere presence of the appropriate (a.k.a. canonical, veridical) labels guided the participants to the relevant elements of the stimuli, as they appear to do for a visual search task, then the participants could have made the connection between the labels and the stimuli and realised the deception involved. From the post-experimental questionnaires, this was only the case for one of the participants.

Taken together, the results from Experiments 1 to 3 suggest that the connection between the odour and the label is an important aspect of odour identification, in agreement with the three keys to successful odour identification (Cain, 1979) and other studies that highlight the importance of the connection between an odour and its name (Jehl, Royet, & Holley, 1997; Jonsson, Tchekhova, Lonner, & Olsson, 2005). While it may be unsurprising that the participants were able to match appropriate labels to the stimuli better than inappropriate and irrelevant labels, it is still interesting to note that performance on the inappropriate and irrelevant labels was significantly higher than chance in some blocks and that the requirement of a long-standing connection between a stimulus and its name can be overcome with relatively little training.

The question of what constitutes an appropriate label for wine stimuli for novice participants was addressed in Chapter 6. It was expected that grape names would not be as useful as the description-based labels for the participants as they had not developed a connection between the words (e.g. Shiraz) and the sensations in the wine. Thus, the description-based labels (e.g. spice and chocolate) were expected to be more useful for the participants. This was not the case when the wines were smelled, with participants able to use both kinds of label at a level significantly greater than chance. The only experiment to find a difference between Description and Grape Name groups was Experiment 14, where the wines were tasted. In this experiment, the groups did not differ during training, with both groups performing significantly better than chance. However, in the test session, the Grape Name group could no longer perform at a level significantly higher than chance, while the Descriptor group could. Furthermore, the groups significantly differed in the final test block, suggesting that the descriptions may be more useful than the grape names as they carry more meaning for a novice than the words “Shiraz” and “Pinot Noir”.

All participants in all experiments were wine novices according to their answers on the Australian Wine Knowledge Questionnaire (Hughson & Boakes, 2001) and would thus not have made any connection between the grape names and the stimuli. Thus, the finding that they could learn to use the grape names at a level significantly greater than chance in some experiments once again suggests that the requirement of a long-standing connection between the odour and its name (Cain, 1979) can be overcome.

However, the results from Experiment 11 indicate that the participants did not seem to detect the elements within each wine to which these descriptions refer, although they did when the elements were enhanced in Experiments 12 and 13. Thus it is unclear how the participants were using the labels and none were able to elucidate the issue in post-experimental questionnaires. One potential explanation is that the grape names and descriptor labels do not have a previous connection to an olfactory sensation *in a wine context* and it was thus possible to at least start to create an association within the time allotted, rather than having to ignore a previously learned odour-label connection, as was the case for the Inappropriate and Irrelevant groups in Chapter 5.

Description of wines and matching descriptions back to wines

When asked to describe the odour samples, the vast majority of participants were unable to give the veridical name for the unique element. These participants were not given any prompts in doing so and many used autobiographical terms, such as “Aunt Lily’s clothes”, which could not be checked for accuracy. However, they were able to use these labels at a level significantly higher than chance, suggesting that their self-generated labels had some connection to the stimuli *for them*.

The same was not true for the wine samples. Participants were asked to use their own words to describe wine samples (Appendix I) and their answers on post-experimental questionnaires suggested that they found this an extremely difficult task. Thus, in Experiment 7, the participants were given prompts in the form of the Wine Aroma Wheel (Noble et al., 1987). The words chosen by the novices did not match those in the winemakers’ descriptions in most cases, although it is acknowledged that wines are complex stimuli and that any number of descriptors could be considered correct. However, it is clear that the novices were unable to match their own descriptions back to the wine samples. This finding is in agreement with Chollet et al. (2005) work with beer matching tasks and also with Lawless (1984), whose novices were unable to match their own descriptions, or descriptions written by other novices, back to the wine samples. The difference between these previous experiments and those reported here is that the novices were trained in the present experiments. This did not appear to make a difference.

Two possible explanations exist to explain why novices could not match their own descriptions back to the wine samples, despite training. The first is that the descriptions were not accurate representations of the attributes of the stimuli and thus did not actually relate to any sensation within the wine, consistent with Gawel (1997) and Solomon (1988). Thus, despite the finding that labelling an odour can facilitate subsequent recognition of that odour (Rabin & Cain, 1984), this does not appear to be the case when the applied label is inaccurate. This is also consistent with Cain (1979), whose participants performed poorly in terms of odour samples when using self-generated labels that were either near misses or far misses.

The second possible explanation is that the participants were unable to identify unique elements within the wines to describe, even when asked to compare and contrast the wines side-by-side when generating their descriptions. This theory is consistent with the work with beer experts by Chollet et al. (2005). In their experiment, when the beers were supplemented, or enhanced with additives, the novices were only able to match descriptions written by experts, under the presumption that these expert descriptions contained these enhanced elements in them. When the elements were not enhanced in the beers, the novices were unable to use the expert descriptions, suggesting that part of the challenge for novices is identifying the attribute in the first place.

However, one limitation is that the participants in Experiment 7 were limited to using the terms on the Wine Aroma Wheel, in an attempt to ensure they used standardized wine terminology. This may have forced them to use terms that, although meaningful in a non-wine context, had no connection to any sensation within the wine for them. Given the finding by Lehrer (1983) that different people find different aspects of wines salient, this may have forced the participants to use words that they would not otherwise use. However, the results from Experiment 7 also agree with other studies where novices cannot match their own descriptions back to wines (e.g. Lawless, 1984).

More recently, alternative methodologies for untrained consumers, such as the “check all that apply” approach, have been reported (e.g. Ng, Chaya, & Hort, 2013). These were not explored here, but could be considered for future experiments.

While it appears that the language of wine is a useful communication tool for experts, it may not actually mean very much to novice consumers, presumably because they lack the connection between the label and the odorant in this wine context. It may be useful to explore other means of communication, such as visual depictions or cross-modal correspondences (Spence, 2010).

Feedback as a training mechanism

The aim of the repeated training sessions was to train the participants to identify the odour or wine samples. In most experiments, the participants were only given a relatively small amount of training and no significant learning effect was found over training in most experiments, either because the participants were already performing significantly better than chance in the first training block or because they could not do the task in the first place and this did not improve over time (e.g. Experiment 7). The results suggest that active training using feedback was not a particularly good training mechanism over this period of time.

However, feedback did appear to be important in Experiment 14, where the grape name group dropped back to chance level of performance in test. This suggests that feedback was important at least in this case for maintaining performance on labels that otherwise have little sensory meaning to the participants.

Direct comparisons as a means of identifying unique elements

As discussed in Chapter 3, intermixed trials are an important component in terms of perceptual learning. Intermixed trials were used in all training and test blocks, although some adjacent trials were the same stimulus. The aim of this was not to test theories of perceptual learning, but to use this effect to allow participants to compare and contrast the stimuli throughout their training, in order to highlight the unique elements of each stimulus.

In experiments where participants have described multiple samples, such as the beer experiments by Chollet et al. (2005), the participants were not permitted to smell or taste more than one sample at once, limiting the opportunity to compare and contrast the samples. Thus, Experiment 7 was novel in that half of the participants were able to compare and contrast the wines during this description phase. It appears that this procedure did not help the participants to identify the characteristics that are unique to each grape. Furthermore, neither group could match their descriptions back to the wines at a level above chance, even after training. Thus, even if the compare and contrast method during description helps participants *detect* the unique elements in a stimulus, it still does not help them *identify* the elements.

Multiple testing sessions as a training mechanism

In Experiment 15, participants were given ten training sessions and two test sessions, one of which occurred after the fifth training session (i.e. halfway through training). A significant increase in performance was found between the test sessions. This may be reflective of the general increase in performance across the training sessions, but it also raises some potential points for training.

Repeated testing sessions in this experiment may have allowed the participants to evaluate their identification strategy and adjust it where necessary to improve in future training and test sessions. The finding is not unlike that found in the very different domain of memory tests, known as test-enhanced learning (McDaniel & Fisher, 1991; Roediger & Karpicke, 2006). While the findings in the memory literature are usually applied to scholarly education, it may also be a useful finding for other types of education, including wine education, although a much more rigorous design than that used in Experiment 15 should be employed to rule out general learning effects over time.

Tasting the wines

Previous wine studies have used olfactory-only designs (e.g. Ballester et al., 2008) and found that their participants were able to perform the tasks required, so the same methodology was used in most of the present experiments (reported in Chapters 5 to 8) to reduce any concerns about the effect of alcohol on performance. However, the participants in the experiments in Chapters 9 and 10 were asked to taste (and spit out) the wines, in order to determine whether there was any improvement in performance over the olfactory-only experiments.

While not directly tested in an experiment, it appears that allowing the participants to taste the wines aided identification performance, which is in agreement with Small and Prescott (2005). This is particularly the case for the Lambrusco wine, as seen in Experiments 14 and 15 (Chapter 9). This finding highlights the importance of gustatory and somatosensory information when tasting wine samples. However, when the participants were only allowed to smell the wine, they were still able to use grape name or abbreviated winemakers' descriptions to identify the wines, suggesting that a lack of gustatory and/or somatosensory information is not necessarily a major problem unless there is an important and unique component to the wine that can only be detected in the mouth. The notion is most likely captured by the old, but unattributed, adage that "90% of winetasting is in the nose".

In visual perception experiments, hundreds or stimuli can be presented per hour. This is not the case for chemosensory stimuli, where adaptation is a concern. For wine stimuli, there is also a concern about alcohol consumption during training. The experiments in Chapter 6 indicate that the participants were able to gain enough information from olfaction alone to be able to do the task at a rate significantly higher than chance. Thus, while adaptation is still a concern whether the wine samples are smelled or tasted, training involving only smelling the wine may be a useful way of avoiding alcohol consumption related concerns.

The Easy-to-Hard effect and the use of additives in the wines

The easy-to-hard effect is a well established finding in the learning literature, where learning on an easier version of the task transfers to a more difficult version of the task, resulting in better performance than for those who have been trained on the difficult version of the task (Mackintosh, 1975; Pavlov, 1927). Furthermore, additives have been used on numerous occasions to highlight the important attributes of beers (Chollet et al., 2005; Meilgaard et al., 1982) and wines (LaTour et al., 2011; Noble et al., 1987).

One possible explanation for the poor identification performance displayed in the wine experiments was that the participants found it difficult to detect the relevant element within the wine sample. When these elements were increased with additives, the participants were able to use the appropriate terms either to rate the wines on line scales for those attributes (Experiment 12) or to identify the wines using Appropriate labels (Experiment 13). When the additives were no longer present, the wines were rated the same on all scales (Experiment 12). Similarly, in the training experiment (Experiment 13), the removal of any additives resulted in a marked drop in performance.

These results suggest that the participants found it difficult to detect the relevant elements in the wines (e.g. pepper in Shiraz) when no additives were present. It may have been the case that the added elements (pepper, blueberry and cherry) were not an accurate representation of the natural pepper, blueberry and cherry odours within the wine samples, although they were used after consultation with experts. Or, it is possible that the added elements were an accurate representation of the odours within the wines, but that when the wines are smelled without the additives, the relevant elements are more difficult to detect amongst the hundreds of other odours within the wine. One final possibility is that the participants were simply able to identify the additives and were able to identify the additives only but were unable to transfer this recognition ability to the wine samples themselves.

The “easy” stimuli were created in a manner similar to the reference standards of Noble et al. (1987), although in the present experiments, two different levels of additives were used in an attempt to create stimuli that were of “easy” and “medium” naming difficulty. The aim of the wine standards made by Noble et al. (1987) was “to assist international usage and understanding of flavor terminology.” However, the present results suggest that, at least for novices, any identification advantage of the standards may not carry over to unadulterated wines, which is also in agreement with the findings by Chollet et al. (2005) with beer samples. That is, the standards may have limited usefulness in terms of terminology training when it comes to testing with unadulterated wines, at least in the case of short-term training.

Length of training

It is important to note that the participants in most of these experiments received relatively little training. The aim was not to make the participants into experts within a short space of time, but to determine whether any particular variables increased performance in the short term. Learning effects were only observed in a small number of the experiments, as already discussed.

The aim of Experiment 15, where participants received ten training blocks and two test blocks (all blocks separated by at least 24 hours) was to determine whether the participants could improve and, if so, how much training was necessary to do so. Identification performance was already quite high in the first training session as the wines were tasted and included the sweet, semi-sparkling Lambrusco wine. Despite this, a learning effect was still found, although there was no clear plateau to suggest a point at which no further learning could take place, suggesting that continued training is important. This finding may be unsurprising to most, but it is important to note that the experiment was an expensive and time-consuming experiment to run. Combined with the fact that one can only taste or smell so many wine samples in a session, high quality wine is a relatively expensive commodity for a learning experiment. As such, it was impossible to run experiments involving this much training throughout all of the present experiments. However, previous experiments have used wines of similar quality, including the same Lambrusco (e.g. Hughson & Boakes, 2009).

Training on contextual information

Experiment 17 (Chapter 10) aimed to determine whether the provision of some relevant conceptual information during perceptual training was useful for the participants. Previous studies using wine stimuli (LaTour et al., 2011) have found that conceptual training was an important factor for novices. In Experiment 17, no significant differences were found between the group who received sensory-only information and those who received sensory and conceptual information during feedback, indicating that this particular information was not useful for these participants in terms of identifying the wines.

The conceptual information given to the participants was not explicitly tied to a sensory aspect of the wine. For example, explaining that the Shiraz came from the 2009 vintage does not relate to any sensory attribute of the wine for a novice. However, using abstract visual stimuli, Gauthier et al. (2003) and Curby and Gauthier (2010) found that semantic information can increase recognition, even if that information is completely arbitrary. No such similar effect was found in this particular study. Thus, while conceptual training is often used in some wine training courses, it does not appear to be immediately useful to novices.

Transfer of knowledge

The aim of any wine identification training program is not to train participants to be able to identify a very particular set of wines, but to train participants on a set of salient attributes that define the particular wine style, which they can then use to identify other examples of the wine. For example, learning that a particular Shiraz tends to be peppery is not a particularly useful rule unless pepperiness is an attribute found in other Shirazes.

This transfer of learning was tested in two ways. The first was in the easy-to-hard experiments, where participants were presented with wines with additives to enhance the important elements. Participants were able to identify the important elements when enhanced, but could not transfer this learning to the same elements in the wines when the additives were removed. This finding has already been discussed above.

The second way in which transfer of learning was tested was in Experiments 15, 16 and 17. In these experiments, the participants were tested using the wines on which they had been trained, as well as a new set of wines that shared similar characteristics to the training wines. In Experiment 15, where participants tasted the wines, all five participants who took part in this task were able to correctly identify every single trial of the new Lambrusco, suggesting that they had learned important sensory information about that wine compared to the others. This particular style of wine is sweet and slightly spritzy, so this learning may have been mostly due to non-olfactory cues. However, these non-olfactory cues are still an important part of wine perception. This suggests that, when important discriminating cues are particularly salient, novices are able to use these cues.

In Experiment 16, no such transfer was found in that the participants could not identify the new samples at a rate that was significantly higher than chance. All of the wines in this experiment were still, red table wines and the participants received much less training. However, they were able to identify the wines on which they were trained at a level significantly greater than chance in the test session, suggesting that what they had learned during training did not transfer to these new stimuli. One possible explanation for this is provided by the assertion by Ballester et al. (2009) that the sensory boundaries between wines made of different grapes are not clear-cut. While every effort was made to ensure that the trained and untrained versions of the stimuli were similar in terms of the winemakers' descriptions, this may not have been an ideal selection procedure. This is discussed further below in the limitations section. A second possible explanation is that the participants may not have been able to detect or identify the attributes in the given descriptors (e.g. spice and chocolate) and instead attempted to learn some other sensory characteristic of the wines apart from that described by the labels that were given to them. This would be analogous to the Irrelevant groups in Chapter 5, who were asked to apply labels to the samples when there was no detectable sensory attribute in the stimuli that matched those labels, which would explain the poor performance.

In Experiment 17, a transfer effect was found for the group that received sensory and contextual information during training. However, their performance in the test session was not significantly different to those who received sensory information only and this latter group did not perform significantly better than chance. Taken together, these results indicate that the observed transfer effect in the former group was relatively weak. However, as these wines were still, red table wines and did not appear to include any elements such as the sweetness and spritz of the Lambrusco (as used in Experiment 15), it is still interesting that one group performed significantly better than chance on these new stimuli with only four sessions of training. Thus it appears that novices can start to learn rules that transfer to other wines made with the same grapes during this amount of training.

Limitations

Amount of training

The major limitation of this research is the relatively small amount of training that the participants received. The participants were first year Psychology students who were sourced from a participant pool. While these participants were obtained at no cost, there is a limit to the number of hours available to each experimenter.

As discussed in Chapter 3, trained panels undergo many hours of training, often requiring at least 60 hours or more. In most cases, this training does not end after one particular batch of products has been evaluated and instead continues indefinitely. However, these panelists are trained to detect, identify and objectively quantify multiple attributes of a particular stimulus or product. In the experiments reported here, the participants were trained on a maximum of three different odour or wine samples and each received at least twelve trials on each sample with feedback. Other experiments have trained participants on far more odours using far fewer training sessions (e.g. Cain, 1979).

The other limitation on the amount of training was that wine is a relatively expensive stimulus to use, which is why odour samples were used in some of the earlier studies. Given the limited resources in terms of access to participants and the use of relatively expensive stimuli, it was necessary to limit the amount of training available.

In Experiment 15, where participants received much more training, a linear increase in performance was observed across training blocks. However, participants in many of the short experiments were still able to perform at levels greater than chance, indicating that while the task was not impossible, any increase in performance may take much longer than the time allocated here and may not be observable in a laboratory setting without more resources.

Furthermore, the training was condensed into a relatively short amount of time due to time limitations. This may have resulted in some level of adaptation, which may have countered any learning effects.

The use of first-year Psychology students as participants

Undergraduate Psychology students were chosen on the basis that they participated at no cost and were unlikely to have extensive experience with wine due to their age (the legal drinking age in Australia is 18). Some participants indicated that they had drunk some wine in the past, but their results on the AWKQ (Hughson & Boakes, 2001) indicated that they had little to no conceptual knowledge of wine. While these participants may have had some perceptual knowledge of wine, exclusion of their data made no difference to the pattern of results in any studies.

However, there may be little agreement in the literature as to what constitutes a wine novice. Parr et al. (2002) defined novices as those who drink wine regularly but had little to no formal wine training. In contrast, Melcher and Schooler (1996) used a group of “non-red wine drinkers” as novices, as they had virtually no experience with the stimulus. Thus, while the novices in the experiments reported in this thesis may have had a small amount of wine experience, only a very small number drank wine regularly.

One more potential limitation in terms of the participants was that there was no requirement that the participants were wine drinkers. However, the participants voluntarily signed up for the experiments through a system and were aware that the task was a wine task. According to the Australian Wine Knowledge Questionnaire, the vast majority of the participants had at least some exposure to wine before taking part in the experiment, although some were not regular wine drinkers. However, the experimenter (AR) noted that there was a high level of motivation amongst almost all of the participants, thus alleviating some concerns about this limitation.

Limited access to participants and potential adaptation

All participants were sourced from a subject pool, where each researcher is allocated up to 100 hours per semester of testing, with a maximum of 4 hours per participant. It was necessary to be economical in terms of the number of participants per experiment and the amount of testing time per participant due to these limitations. As mentioned in Chapter 4, relatively small numbers of participants are often used in this area of research, although this may have resulted in limited power for some experiments. Thus, future researchers with greater resources may wish to revisit some of these experiments with a larger sample size.

Another concern about these limitations was that there was a relatively large number of samples that were presented per hour per participant, which may have resulted in some level of adaptation. All efforts were made to avoid this, such as breaks between blocks of stimuli and the provision of water for all participants. The use of small numbers of participants per experiment and possible adaptation were a necessary trade-off given the limited access to participants.

The use of winemakers' tasting notes as "appropriate" labels

In the odour sample studies reported in Chapter 5, the "appropriate" labels were veridical labels. In order to be correct, the participants would identify the sample containing the vanilla odour using the label "vanilla". For the wine samples, the “appropriate” labels were taken from the winemakers’ descriptions and were chosen on the basis that they only appeared in the description of that wine and not in any of the other wines. The chosen descriptors were also chosen on the basis that they were typical of wines made of that particular grape (e.g. a correct term for the Shiraz samples was “pepper” or “spicy”), although it is acknowledged that wines made from a grape vary in perceptual characteristics based on region, variety, winemaking techniques and many other factors.

A limitation of this approach is that the terms appear to have been chosen on the basis of the descriptions of people who would be considered experts by most standards (see Chapter 3). While it is possible that a winemaker may describe his or her Shiraz as peppery because that is what a Shiraz is supposed to smell like, rather than on the basis that it does actually smell peppery, efforts were made to corroborate these terms. The descriptions of these wines written by prominent wine critics were also checked. There was some variation between how the wines were described and some of the terms that were unique in the winemakers’ descriptions appeared in the critics’ descriptions of the other wines. However, the terms chosen are generally recognised as typical of wines made from those grapes and were retained on this basis.

Stimulus presentation

The wines were initially presented on make-up removal pads, in order to prevent the participants from attempting to drink the samples. It is entirely possible that this may have altered the odour of the wines. Furthermore, these samples were reused between participants (refreshed every two hours) in order to maximise the number of participants that could be tested on each testing day. Thus the wine samples may have changed between the two people who smelled each sample. However, the wines would still have been as constant as possible for each participant. Thus, while this consideration may have reduced the external validity of the experiments, the internal validity was likely to be less effected.

Finally, the wines were presented in black plastic cups for the taste experiments, rather than wine glasses. This was because when the wines were swirled, or otherwise moved around, in the glasses, visual differences were apparent, whether or not food colouring was used to disguise these differences. Thus black (or otherwise opaque) wine glasses were required. Furthermore, given that scores of samples were required for each testing session, obtaining the required number of glasses at low cost was not possible.

The wines used were relatively cheap

Due to funding constraints, the wines that were used generally retailed between A$8 and A$22 per bottle. These wines were therefore relatively cheap and may not have been of the best quality. However, 94% of bottled wine purchased and consumed in Australia costs less than $20 per bottle and this figure increases to 97% if non-bottled wine is included (Australian Bureau of Statistics, 2013), so these particular wines may have been representative of the typical wines drunk by most wine consumers.

Every effort was made to ensure that the wines used were sufficiently representative of their style, including checking tasting notes written by winemakers and critics. Furthermore, four experienced wine retailers, each with over 20 years of experience, were asked to taste the wines and determine whether they were acceptable representations of wines made from that grape. All reported that they thought the wines were varietal. They were also asked whether the descriptors used were representative of the wines and all reported that they were. Thus, while the tasting notes from the winemakers were a potential weakness of the study, efforts were made to confirm these tasting notes by critics and experienced wine retailers.

Conclusion

Different wine training courses may have different goals, such as teaching novices that not all wines taste the same (discrimination) or how to describe or identify wines based solely on perception and not by reading the label (identification). In terms of discrimination, the present results indicate that novices can discriminate many wines based on smell alone and may also be able to categorise some wines into groups based on grape type, in contrast with the previous literature (e.g. Ballester et al., 2008). Thus, discrimination training may not be particularly helpful to novices as it appears that they already possess this ability, at least for some sets of wines.

Previous studies have suggested that a major difference between experts and novices is their ability to use a specific lexicon to describe wines (Gawel, 1997; Solomon, 1988), which requires identification of the elements within the wine. The first requirement of being able to identify an element of a wine is the ability to detect it in the wine. The present studies suggest that novices may not be able to detect the important, salient elements of wine samples, as they were unable to rate wines differently on the important elements that are generally believed to differentiate wines made from different grapes (e.g. pepperiness in a Shiraz). In order to improve this ability, previous studies (Chollet et al., 2005; LaTour et al., 2011) have used adulterated samples to enhance these elements, but the present results suggest that this learning does not transfer to unadulterated versions of the same wines, indicating that this may not be a very useful training method.

The aim of most wine training books and online courses is to improve conceptual knowledge, such as how wines are made. The results here suggest that, while such training has been found to be useful in vision research (Curby & Gauthier, 2010; Gauthier et al., 2003), it met with limited success in Experiment 17, indicating that the main focus of wine training books may not help novices to understand the perceptual aspects of wine. This is in keeping with the study by LaTour et al. (2011), who found that those trained on conceptual aspects of wine did not perform as well as those trained on the perceptual aspects of wine.

Books may describe the important elements of wines made from different grapes (e.g. Shiraz tends to be peppery), but it is up to the reader to determine what these elements smell like within wine samples. There is generally no practical test of how much a reader has learned and thus the usefulness of these particular methods is never tested. In contrast, practical wine training courses include actual perceptual training, but also do not include any form of testing, so their effectiveness has also not been evaluated. The present results suggest that training a novice to identify important elements in wine samples can succeed in the short term in some situations, particularly when taste is involved. However, in the present studies, training participants to identify samples based on the grape from which they were made met with less success, indicating that descriptors may be more useful for novices compared to grape names, as the ability to identify wines by grape names is likely to involve a cognitive component (Hughson & Boakes, 2002a).

The ability to generate a description of a wine appears to be a very difficult task for novices. This conclusion is drawn from Experiment 7, where participants were unable to learn to use their own descriptions to identify the wine samples. Whereas this has been found in previous research (e.g. Valentin et al., 2003), the novel finding in the present study is that this was the case even after training. In contrast, the participants in Experiment 1 were able to use their own labels for the odour mixtures, suggesting that there may be something unique to wine language that makes this a particularly difficult task. Furthermore, comparing wines side-by-side did not help the novices to generate descriptions that they would later find useful. These comparisons were similar to those used in the discrimination experiments, where participants were able to discriminate between the wines. Thus, while the comparisons allowed the participants to detect the unique elements of each wine, they did not help them to apply useful labels to these elements.

Wine expertise is not a perceptual advantage. Wine expertise is the ability to describe wines using a specific lexicon that other experts can use. This ability can be learned through extensive training, although the present results suggest that novices can start to learn to use important wine descriptors within a relatively short space of time. Conceptual aspects of wine expertise, such as learning information about wine regions and winemaking techniques, do not appear to help very much with these particular abilities. Finally, it appears that being trained on the same wine examples does not always lead to learning about other wines made from the same grape, so training on multiple types of each wine may be more useful.

On the basis of these experiments, it appears that none of the methods serve as a shortcut to wine expertise. Training methods that use adulterated samples may help participants to learn the lexicon, but this does not appear to transfer to unadulterated stimuli. Learning concepts may help people to learn about wine, but it is unlikely to help them with the particular skills that generally separate experts from novices. Instead, it appears that attaining these skills requires a considerable amount of time, but the present results show that some learning can occur over a relatively short amount of time (less than four hours) given intense training.

As discussed at various points in the thesis, the practical constraints generated by subjects pool limitations and costs meant that many of the experiments were not as productive as they may have been had those limitations not been present. Future researchers with more resources may wish to revisit some of these experiments, particularly the easy-to-hard effect and the use of contextual as well as sensory training.

APPENDIX A

Descriptions of odour samples by participants in Experiment 1

Vanilla + Citral (A)

"Only lemon smell"

"Detergent, Aunt Lily's clothes, strong lemon smell from the detergent dad used"

"Lollies, candy cane but not mint, Movieworld"

"Lavender, toilet cleaner, refreshing (air fragrance), smells good (musky good), soothing on nose"

"Cheap lemon lollies, very lemony"

"Lemon throat lollies, strong, smells sour"

"Lolly pop (sunkiss), stronger than B (lemon)"

"Floweriness, medicine"

"Sherbert lemons, high, sharper, flowery mellow butter"

Melon + Citral (B)

"Stronger smell"

"Dad's car, slightly sweeter than A"

"Lighter, little orange, lavender"

"Lemon, with smell diff, rusty, old vomit, stings top of nose, less potent than C"

"Midori, mixed with lemonade, less lemony"

"Sweet scent, softer"

"Grapey odour"

"Flatter smell than odour A, sweet, B and C seem alike but B's sweeter"

"Washing powder, sharper"

Banana + Citral (C)

"Can smell banana"

"My old diary, sweet"

"Stronger orange than B, very similar to B but a bit stronger, hint of sandalwood"

"Vomity, stings top of nose"

"Vanilla, stronger Midori, bubblegum (unlike B)"

"Similar to A, but less intense, doesn't smell as sour, smell banana"

"Toiletry smell, pear, flowery smell, old smell"

"Woodier, earthiness, spicy, pungent, camphor"

"Pear drops, banana"

APPENDIX B

Examples of Feedback in Experiment 17

For the "sensory-only information" group

- Correct! That was Shiraz, the spicy and chocolatey one.
- Incorrect! That was Pinot Noir, the black cherry and gamey one.

For the "sensory plus conceptual information" group

- Correct! That was Shiraz, the spicy and chocolatey one.
- Incorrect! That was Pinot Noir, the lighter-bodied wine.
- Correct! That was Pinot Noir, the wine from the cooler climate.
- Incorrect! That was Shiraz, the wine from the 2009 vintage.
- Incorrect! That was Shiraz, the wine from the Barossa Valley.
- Correct! That was Shiraz, the wine with the more persistent aftertaste.
- Incorrect! That was Pinot Noir, the more delicate style.
- Correct! That was Pinot Noir. Pinot Noirs from that region are generally noted for their earthy flavours.
- Incorrect! That was Shiraz. Shirazes from that region are noted for their peppery and spicy characteristics.

APPENDIX C

Modified Australian Wine Knowledge Questionnaire

Please answer the following questions as best you can. You are required to answer each question, if you don't know, just guess.

Date:______________ Time:________________

Participant Number: ______ Age:_______ Gender:________

1. Indicate the traditional colour of the following varieties of wine

a. Chardonnay White Red
b. Shiraz White Red
c. Merlot White Red
d. Chambourcin White Red
e. Riesling White Red
f. Semillon White Red
g. Gewurztraminer White Red
h. Grenache White Red

2. How do botrytis wines differ from standard wines?

A. Sugar is added to standard still wine to increase sweetness.
B. Grapes are infected by a fungus called botrytis.
C. Grapes of the variety botrytis are used.
D. Botrytis fermentation techniques are used.
E. None of the above.

3. What is the main grape variety used in "Grange"?

A. Semillon.
B. Chardonnay.
C. Cabernet.
D. Shiraz.
E. Pinot Noir.

4. What type of oak is Grange primarily matured in?

A. American.
B. French.
C. Spanish.
D. Australian.
E. English.

5. What is the distinction between aroma and bouquet?

A. Bouquet is produced by red grapes and aroma by white grapes.
B. Bouquet occurs only in sparkling wines and aroma occurs only in still wines.
C. Aroma is based on climate, bouquet on soils.
D. Bouquet comes from fermentation procedures whereas aroma has origins in the grape alone.
E. Bouquet fades with bottle age whereas aroma does not.

6. What style is typical Hunter Valley Semillon?

A. Dry and Unwooded.
B. Sweet and Unwooded.
C. Sweet and Heavily Oaked.
D. Dry and Heavily Oaked.
E. Dry and Sweet.

7. What grapes is traditional champagne made with?

A. Riesling and Chardonnay.
B. Shiraz and Cabernet.
C. Chardonnay and Pinot Noir.
D. Grenache and Semillon.
E. Sauvignon Blanc.

8. What colour is the flesh of a Pinot Noir grape?

A. Red.
B. White.
C. Pink.
D. Purple.
E. Yellow.

9. How often do you drink wine?

A. Everyday.
B. At least once a week.
C. Once or more a month.
D. Less than once per month.

10. How many years have you been a regular wine drinker (at least twice per week)?

A. More than 10 years
B. 5 – 10 years
C. 1 – 5 years
D. Less than one year
E. Am not a regular wine drinker

11. How much have you read about wine?

A. 3 or more books or articles
B. 1 – 3 books or articles.
C. Less than 1 book.
D. Only labels.

12. How knowledgeable would you say you are about wine?

A. Expert level
B. I know more than average
C. I know a little bit
D. I drink it but I don’t know much about it
E. I don’t know much about wine and I don’t drink it.

APPENDIX D

Experiment A1 – Pilot Study Testing the Identification of Elements in Binary Odour Mixtures

This experiment was run prior to the first experiment reported in the body of the thesis. It was the first attempt to train participants on the binary odour mixtures used throughout Chapter 5.

Method

Participants. Twelve first-year Psychology students (six female) took part in the experiment for course credit. One participant reported that they were suffering from rhinitis during the experiment and her data were excluded. See page 53 for general information about the participants.

Materials. Unlike the experiments in the rest of the thesis, the participants were trained on two of the odour mixtures, not three. These were Vanilla + Citral and Melon + Citral (AX and BX, as is often used in perceptual learning experiments). Furthermore, the concentrations of the additives were determined by mixing odours so that the strength of the Citral odour was identical in both mixtures and approximately equal to the strength of the Vanilla or Melon (Table A19). Thus, the Citral was not as strong as the odour mixtures used in Chapter 5. The olfactory stimuli were soaked into cotton buds and placed in opaque Décor plastic sauce bottles.

Table A19. Concentration and supplier of odours in Experiment A1.

Odour name	Dilution Ratio	Supplier	Order number
Citral	1:1800	Perfume Manufacturers	N/A
Vanilla	1:600	Tastemaster	080820
Melon	1:725	Quest	AP05403

Note: All odorants dissolved in water.

Procedure. Participants attended a single one-and-a-half-hour session, consisting of a practice block of eight trials using the same flower and pear stimuli as in other experiments. Participants then received two training blocks and one test block, each consisting of 18 trials (nine trials of Vanilla + Citral and nine of Melon + Citral). All participants were asked to identify the unique odour in each mixture using the labels Vanilla or Melon (similar to the Appropriate groups in Chapter 5) and were told to ignore the common lemon odour. That is, there was only one group of participants in the experiment. Feedback was given during the training sessions immediately after each trial, but not during the test session. The intertrial interval was 1min with a 5-min break observed between blocks. Participants were encouraged to drink water after each trial.

The stimuli were randomized with the only constraint being that no odour appeared more than three times in a row. Participants were not informed about this constraint, nor were they informed that there were nine of each odour mixture in each block.

Only one participant attended each session and responses were recorded on answer sheets.

Results

The participants were able to identify the odours at a level significantly better than chance (with the smallest *t*-value of $t(10) = 5.93$, $p < 0.001$ for the first training block). There was no significant improvement in performance between training blocks, $t(10) = 2.21$, $p = 0.052$ (Figure A14).

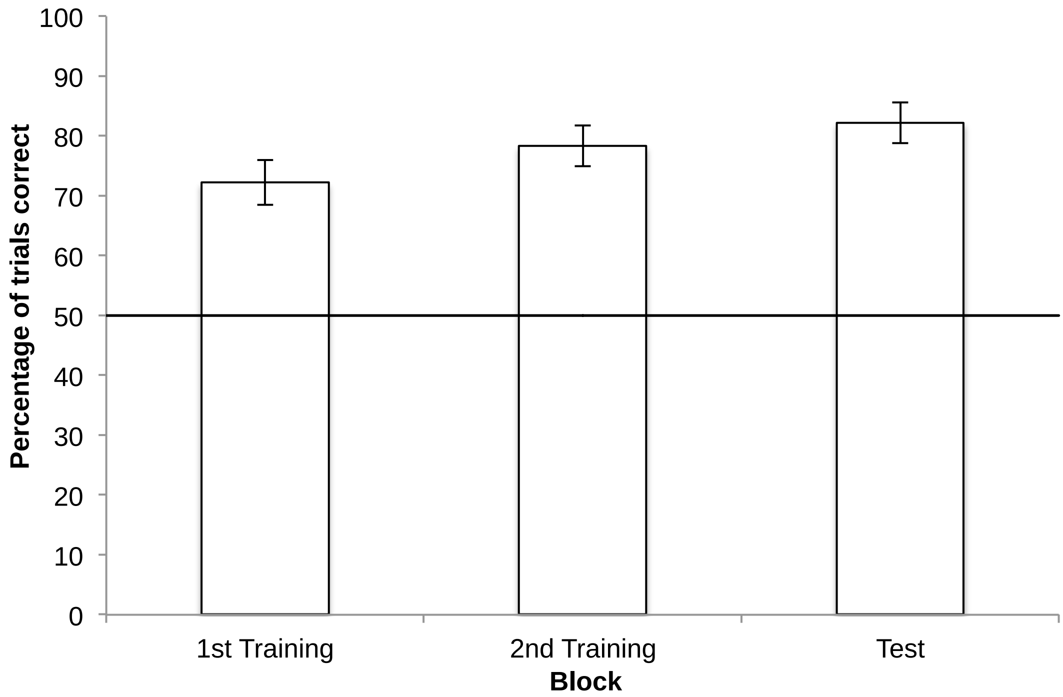

Figure A14. Mean (+/-SE) percentage of trials correct in each training and test block in Experiment A1.

Note: Feedback was given in the training blocks, but not in the test block. The horizontal line indicates chance level of responding (50%).

Discussion

This experiment served as a formalized pilot experiment to test the concentrations of the odours before starting the experiments in Chapter 5. The results suggested that the odours were too easy to identify and that more difficult stimuli would be needed in order to observe any learning effects. Thus, in order to make the task more difficult, the concentration of Citral in all odours was increased. Furthermore, a third odour mixture was added to make the experiments more difficult for the participants.

APPENDIX E

Experiment A2 – Testing Discrimination of an Alternate Set of White Wines

This experiment was part of the series of experiments reported in Chapter 7 and Appendices F and G. The aim of these experiments was to find a set of wines that were easier to discriminate than those reported in Experiment 4 (Chapter 6).

Method

Participants. Nine first-year Psychology students (eight female, aged between 18 and 29, *M* = 19.9, *SD* = 3.5) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. Three whites wines were used in this experiment. They were Yalumba “Y Series” Sauvignon Blanc (2008), Yering Station “Mr Frog” Chardonnay (2007) and Jacob’s Creek Riesling (2008). See page 57 for general information about preparation of the samples. The odour samples were not used in this experiment.

Procedure. Trials consisted of three stimuli, two of which were identical and one different. The participants sniffed all three wines in a trial and were asked to identify the unique stimulus. No feedback was given. Each trial consisted only of odour stimuli or of wine stimuli. See page 58 for general information about the procedure used in this experiment.

Results

On average, the participants were correct in 10.0 out of 18 trials (*SD* = 2.60). This was significantly higher than chance, $t(8) = 4.62$, $p = 0.002$. When each specific combination of wines was analysed, only two were significantly better than chance (both $t(8) = 3.41$, $p = 0.009$). These are indicated by asterisks in Table A20.

Table A20. Summary of errors by target and foil combination for wine samples in Experiment A2.

Target	Sauvignon Blanc		Chardonnay		Riesling	
Foils	Chard-onnay	Riesling	Sauvignon Blanc	Riesling	Sauvignon Blanc	Chard-onnay
Mean errors (/6)	1.33	1.11*	1.11*	1.67	1.33	1.44

Discussion

While the participants were able to discriminate between the wines at a rate that was greater than chance, the mean performance was still 10 out of 18 trials, which was similar to the performance on the red wine samples used in Experiment 4. Because this set of wines was not easier to discriminate than those in other experiments, they were not used any further.

APPENDIX F

Experiment A3 – Testing Discrimination of an Alternate Set of White Wines

This experiment was part of the series of experiments reported in Chapter 7 and Appendices E and G. The aim of these experiments was to find a set of wines that were easier to discriminate than those reported in Experiment 4 (Chapter 6).

Method

Participants. Eight first-year Psychology students (all female, aged 18 to 21, *M* = 19.4, *SD* = 1.1) participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. Three whites wines were used in this experiment. They were Yalumba “Y Series” Sauvignon Blanc (2008), Yering Station “Mr Frog” Chardonnay (2007) and Hugel “Gentil” (2008). The Hugel Gentil is made from a blend of white grapes. The odour samples were the same as in the experiments in Chapter 5, namely vanilla + Citral, melon + Citral and banana + Citral. The odours were included as a control to serve as a check for motivation and perceptual ability amongst the participants. See pages 55 and 57 for general information about preparation of the samples.

Procedure. The stimuli were presented in alternating blocks of wine and odour stimuli, where each block consisted of six trials. Trials consisted of three stimuli, two of which were identical and one different. The participants sniffed all three odours or wines in a trial and were asked to identify the unique stimulus. Each trial consisted only of odour stimuli or of wine stimuli. No feedback was given. See page 58 for general information about the procedure used in this experiment.

Results

On average, the participants were correct in 8.88 out of 18 wine trials (*SD* = 2.47) and 9.25 (*SD* = 1.67) of the 18 odour trials. This was significantly higher than chance, $t(8) = 4.62$, $p = 0.002$. Both were significantly greater than the chance score of 6 correct ($t(7) = 3.29$, $p = 0.013$ and $t(7) = 5.51$, $p = 0.001$ respectively), but not significantly different to each other, $t(7) = 0.32$, $p = 0.76$. For the wines, only one combination of target and foil was discriminated at a rate significantly better than chance and this was detecting Chardonnay when the foils were the Gentil, $t(7) = 3.74$, $p = 0.007$. The errors for each type of stimulus are presented in Table A21 and Table A22.

Table A21. Summary of errors by target and foil combination for odour samples in Experiment A3.

Target	Vanilla		Melon		Banana	
Foils	Melon	Banana	Vanilla	Banana	Vanilla	Melon
Mean errors (/6)	1.63	1.25	1.63	1.63	1.38	1.25

Table A22. Summary of errors by target and foil combination for wine samples in Experiment A3.

Target	Sauvignon Blanc		Chardonnay		"Gentil"	
Foils	Chard-onnay	"Gentil"	Sauvignon Blanc	"Gentil"	Sauvignon Blanc	Chard-onnay
Mean errors (/6)	2.00	1.63	1.63	1.00	1.38	1.50

Discussion

While the participants were able to discriminate between the wines at a rate that was greater than chance, the mean performance did not exceed 10 out of 18 trial. Because this set of wines was not easier to discriminate than those in other experiments, they were not used any further.

APPENDIX G

Experiment A4 – Testing Discrimination of an Easier Set of Binary Odour Mixtures

The aim of the experiments reported in Chapter 7 and Appendices E and F was to find a set of wines that were easier to discriminate than the wines used in Chapter 6. The highest number of triangle trials correct in any experiment was around 10 out of 18. The aim of the present experiment was to determine whether any set of stimuli would yield a higher discrimination score than 10 out of 18. Thus, an easier set of odour stimuli were created by decreasing the amount of Citral in the mixtures, based on the concentrations used in Appendix D.

Method

Participants. Nine first-year Psychology students (seven female, aged 18 to 23, *M* = 19.0, *SD* = 1.7) took part in the experiment for course credit. See page 53 for general information about the participants.

Materials. The stimuli for this experiment were based on the odour mixtures used in Chapter 5, except that the concentration of Citral was reduced, in order to make the task easier. Concentrations are presented in Table A23. See page 55 for general information about preparation of the samples.

Table A23. Concentration and supplier of odours in Experiment A4.

Odour name	Dilution Ratio	Supplier	Order number
Citral	1:1800	Perfume Manufacturers	N/A
Vanilla	1:600	Tastemaster	080820
Melon	1:725	Quest	AP05403
Banana	1:725	Quest	CD53114

Note: All odorants dissolved in water.

Procedure. The procedure was the same as that used for other experiments involving triangle tests. Participants were presented with three blocks of six trials. Each trial consisted of three stimuli, two of which were identical and one different. The participants were asked to smell each stimulus in order to determine which was the unique sample. No feedback was given throughout the experiment. Water was freely available and participants were asked to smell and drink the water regularly. See page 58 for general information about the procedure used in this experiment.

Results

Out of a possible 18 trials, the mean performance was 13.44 (SD = 2.01), which was significantly higher than chance, $t(9) = 11.13$, $p < 0.001$. Each possible combination of target and foil combinations was also discriminated at a rate significantly higher than chance, with the smallest t-value of $t(9) = 2.87$, $p = 0.021$ (see Table A24).

Table A24. Summary of errors by target and foil combination for odour samples in Experiment A4.

Target	Vanilla		Melon		Banana	
Foils	Melon	Banana	Vanilla	Banana	Vanilla	Melon
Mean errors (/6)	0.78	0.67	0.78	0.56	0.67	1.11

Discussion

The results suggest that performance at a level greater than 10 out of 18 correct in triangle tests is possible if the stimuli are easy enough to discriminate. This highlights the finding that the participants in the wine experiments were able to discriminate between the samples, but not to a high level, indicating that the stimuli were still quite confusable. Furthermore, the finding in previous experiments that discrimination performance rarely exceeded 10 out of 18 correct does not appear to be an artifact of the type of discrimination test employed.

APPENDIX H

Experiment A5 – Comparing the Sensitivity of Triangle Tests and Same/Different Tasks

The experiments in Chapter 7 (and Appendices E and F) used triangle tests as a measure of discrimination. Apart from the easy odour triangles (Appendix G), the mean number of correct trials was never much higher than 10 trials out of 18. I was concerned that the triangle tasks may not have been a sensitive measure of discrimination performance.

An alternate discrimination task is the same/different task, where participants smell two samples in succession and judge whether the two samples contain the same odour, or two different odours. The present experiment used both triangle and same/different tasks to determine whether the same/different task may be a more sensitive measure of discrimination.

Method

Participants. Forty first-year Psychology students (31 females, aged 18 to 38, M = 20.6, *SD* = 4.7) participated in the experiment for course credit. See page 53 for general information about the participants.

Apparatus. The same odour and wine stimuli were used for the triangle tests and the same/different judgements. Only two wines and two odour stimuli were used for this experiment. The odours were Vanilla + Citral and Melon + Citral and the wines were the Yalumba “Galway” Shiraz (2008) and the Yering Station “Mr Frog” Pinot Noir (2008). All stimuli were judged orthonasally. See pages 55 and 57 for general information about preparation of the samples.

Procedure. All participants attended a single one-hour session. Participants were allocated to either the triangle test group, or the same/different judgement group so that participants in adjacent sessions were in different groups. Up to three participants were present in any one session and all participants in the same session were allocated to the same group.

For both groups, participants were allowed to smell each sample for a maximum of 3s. Water was available throughout the experiment and participants were asked to sniff and drink the water throughout the experiment. All responses were recorded using paper and pen and no feedback was given throughout the experiment.

Triangle test group. The participants in the triangle test group received four blocks, each with nine trials. Two of the blocks consisted of wine stimuli and two of odour stimuli and participants started with either a wine block or odour block (counterbalanced design). Thus all essentially followed the same order, but started at different points. Participants alternated between the stimuli types with each subsequent block. Trials consisted of three stimuli, two of which were the same and one different. The location of the different sample in each trial was counterbalanced. Thus, each participant smelled 108 samples and made 36 judgements. The inter-trial interval was 1min, with a 5-min break observed between blocks.

Same/different group. The participants in the same/different group also received four blocks of odours, two of which were wine stimuli and two were odour stimuli. The starting block for each participant was counterbalanced and each participant then alternated between stimuli with each subsequent block. Thus all essentially followed the same order, but started at different points.

Each block consisted of 17 stimuli that were smelled in order. As the participants smelled each stimulus, their task was to determine whether each stimulus was the same or different to the previous stimulus. Thus, in each block, the participants smelled 17 stimuli and made 16 same/different judgements. Over the course of the experiment, they smelled 68 stimuli and made 64 same/different judgements.

The interval between stimuli was also a variable of interest and were alternated between 0s and 45s, timed by the experimenter. An equal number of same and different judgements were presented over long and short intervals. An example of a block of stimuli is:

AX (0s) AX (45s) BX (0s) AX (45s) AX (0s) BX (45s) BX (0s) AX (45s) BX (0s) BX (45s) AX (0s) BX (45s) BX (0s) BX (45s) AX (0s) AX (45s) AX

Analyses. For the triangle group, the percentage correct was compared to chance (33.3%) using a one-sample t-test. For the same/different group, d` was calculated for each stimulus and each interval length using the following formula:

(p(hits) - p(false alarms)) / (1 – p(false alarms).

These scores were analysed using one-sample t-tests for both wine and odour samples to test for discrimination levels above chance.

Results

Triangle tests. The average number of correct wine trials was 7.39 (*SD* = 2.37) out of 18, while the average correct number of odour trials was 8.17 (*SD* = 2.50). Both of these figures were significantly higher than the chance level of six items correct, $t(22) = 2.82$, $p = 0.010$ and $t(22) = 4.17$, $p < 0.001$ respectively. Furthermore, there was no significant difference in terms of discrimination performance between the stimuli, $t(22) = 0.92$, $p = 0.37$.

Same different tests. The only d-prime that was significantly higher than 0 was that for the odours over the short interval t(21) = 3.19, p = 0.004 (see Table A25). A (2x2) ANOVA did not reveal a significant main effect for stimulus type (F(1,21) = 2.76, p = 0.11) or interval (F(1,21) = 0.20, p = 0.66). The stimulus by interval interaction term was also not statistically significant, $F(1,21) = 2.31$, $p = 0.14$.

The only significant simple effect was between the odours and wines over a short-term interval. According to this measure, the odours were more discriminable than the wines, but only over a short interval, $t(21) = -2.12$, $p = 0.046$.

Table A25. d-prime scores by stimulus and interval length in Experiment A5.

Interval length	**Wines**	**Odours**

Short (15s)	.021	.435
Long (45s)	.202	.120

Discussion

While the triangle test found that both types of stimuli were discriminable, the same-different task did not, indicating that the triangle test is a more sensitive test. Even though the triangle tests are less efficient in terms of the number of samples required per comparison, for these stimuli, it appears to be the more informative test.

It was expected that the shorter intervals in the same/different task would result in greater discrimination performance and this was observed for the odours. It was not observed for the wines, but this may be because the wines were not discriminable according to this test.

The results from this experiment, along with the results from Experiment A4 (Appendix G), suggest that the triangle test is a sensitive measure of discrimination.

Appendix I

Experiment A6 – Testing the Perceived Intensity of Wine Notes When They Are Enhanced Using Alternative Additives

The experiments in Chapter 8 describe an attempt to highlight important odours within wine samples by increasing their concentration using additives, similar to the procedure described by Noble et al. (1987). The results for Experiment 12 (Chapter 8), where some additives were tested, showed that the particular cherry additive used in the Pinot Noir sample was not particularly suitable for this purpose. The aim of Experiment A6 was to test a different cherry additive in the same Pinot Noir. A different wine and additive combination was also tested to determine whether it could be used in case this second cherry and Pinot Noir combination still failed to increase the perceived cherry-ness of the wine.

Method

Participants. Fifteen postgraduate Psychology students (12 females, aged 22 to 35, *M* = 26.1, *SD* = 3.4) participated in the experiment. All drank wine at least once a month, although their low scores on the AWKQ indicated that none of them were experts. Postgraduate students were used instead of undergraduate students as this experiment took place between teaching semesters. See page 53 for general information about the participants.

Materials. Two different wines were used in this experiment. They were the Yering Station “Mr Frog” Pinot Noir (2008) and the Yering Station

"Mr Frog" Chardonnay (2008). See page 57 for general information about how the wines samples were prepared and presented.

The odours that were added to the wines were vanilla (Tastemaster, order number 080820) and cherry (Queen cherry brandy). Each additive was tested at four different concentrations (none, low, medium and high) and these concentrations are outlined in Table A26. The vanilla was added to the Chardonnay and the cherry to the Pinot Noir.

Procedure. Each participant rated each combination of wine and additive twice, for a total of 16 wines. As in Experiment 12, each wine was rated in line scales (1 = "Not at all" and 7 = "Extremely") for vanilla and cherry. Ratings for blueberry, pepper, liking, irritance, sweetness and strength were also collected but were not analysed. The participants were also asked to describe the wines using any words that they liked. All samples were presented orthonasally and water was available throughout. A 1-min inter-trial interval was observed throughout the experiment.

Table A26. Concentration of additives in wine samples in Experiment A6.

Odour name	Low concentration	Medium concentration	High concentration
Vanilla	1:2000	1:1000	1:600
Cherry	1:1000	1:500	1:333

Note: All odorants added to 200mL of wine.

Analysis. The purpose of the experiment was to determine whether the different levels of additive were detectable in the wine. Thus, each wine that contained additives was compared to the same wine without additives on the scale of cherry or vanilla, depending on which had been added to the wine.

Results

Ratings. When the Pinot Noir contained either the medium or high level of cherry additive, it was rated as having a significantly stronger cherry odour than the base wine, $t(14) = 4.16$, $p = 0.001$ and $t(14) = 5.69$, $p < 0.001$ respectively. Furthermore, the medium and high concentrations differed significantly, $t(14) = 2.70$, $p = 0.017$.

Any level of vanilla added to the Chardonnay resulted in the wine being rated as having a stronger vanilla odour compared to the base wine, with the smallest *t*-value of $t(14) = 5.57$, $p < 0.001$ for the low concentration. The mean ratings are presented in Table A27.

Table A27. Mean (and SD) ratings for both of the wines in terms of strength of vanilla and cherry odours in Experiment A6.

		Amount of additive			
Wine	Additive	None	Low	Medium	High
Chardonnay	Vanilla	2.00 (1.18)	4.33 (1.73)	4.67 (1.44)	4.67 (1.44)
Pinot Noir	Cherry	2.75 (1.50)	3.00 (1.57)	4.17 (1.51)	5.33 (1.22)

Descriptions. Most of the descriptors from the participants were words such as “sweet” or “yucky”. There were very few attempts to describe the particular flavours of the wines. In the post-experiment questionnaire, the participants described this as a particularly difficult task.

Discussion

The new cherry additive resulted in an increase in the perceived level of cherry odour in the Pinot Noir, indicating that it was suitable for use in the easy-to-hard experiments. Furthermore, the medium and high levels of cherry additive appeared to enhance that element of the wine to difference degrees, indicating that they would be useful for the gradual change of concentration in the easy-to-hard experiment. This additive at these concentrations was then used in Experiment 13.

APPENDIX J

Experiment A7 – Can the Easy-to-Hard Effect Be Utilised as a Training Method for Novices using Odour Samples?

In order to study the easy-to-hard effect and whether it is useful as a wine training technique, a pilot experiment was run using the odour stimuli instead of the wines, in order to reduce expenses. The rationale for using the easy-to-hard effect is described in Chapter 8.

Method

Participants. Thirty-eight first-year Psychology students (26 females, aged 18 to 44, *M* = 19.5, *SD* = 4.4) from the University of Sydney participated in the experiment for course credit. See page 53 for general information about the participants.

Materials. The odour mixtures described on page 55 were used in this experiment. However, the ratio of unique element to Citral was adjusted in order to create stimuli of three different levels: easy, medium and hard. The easy samples used the same concentrations as those used in previous chapters, while the medium and hard samples contained less of the unique odorant. The Citral concentration was the same in all samples. Concentrations are outlined in Table A28.

Table A28. Concentration and supplier of odours used in Experiment A7.

Odour name	Dilution Ratio			Supplier	Order number
	Easy	Medium	Hard		
Rose (Flowers)	1:600	N/A	N/A	Quest	AP07749
Pear	1:600	N/A	N/A	Quest	AP06882
Citral	1:500	1:500	1:500	Perfume Mfrs	N/A
Vanilla	1:600	1:857	1:1500	Tastemaster	080820
Melon	1:725	1:1160	1:1933	Quest	AP05403
Banana	1:725	1:1160	1:1933	Quest	CD53114

Note: All odorants dissolved in water.

Procedure. The general procedure outlined on page 59 was followed for this experiment, except that participants received three training sessions and a test session, instead of two training sessions and a test session as used in previous experiments.

Participants were allocated into either the “hard-to-hard” group or the “easy-to-hard” group. The “hard-to-hard” group was trained and tested on the hard samples throughout all blocks, while the easy-to-hard group was trained on the easy samples in their first training block, the medium samples in their second training block and the hard samples in their third training block. Both groups were tested on the hard samples. All participants completed the experiment in a single one-and-a-half-hour session, with 5-min intervals between blocks to reduce adaptation.

Results

Both groups scored significantly above chance performance of 6 trials correct per block in all blocks (with the smallest *t*-value of $t(18) = 3.53$, $p = 0.002$ for the hard to hard group in the second training block). The only significant difference between the groups was in the second training block, where the easy-to-hard group performed significantly better than the hard-to-hard group, $t(36) = 2.18$, $p = 0.036$. See Figure A15.

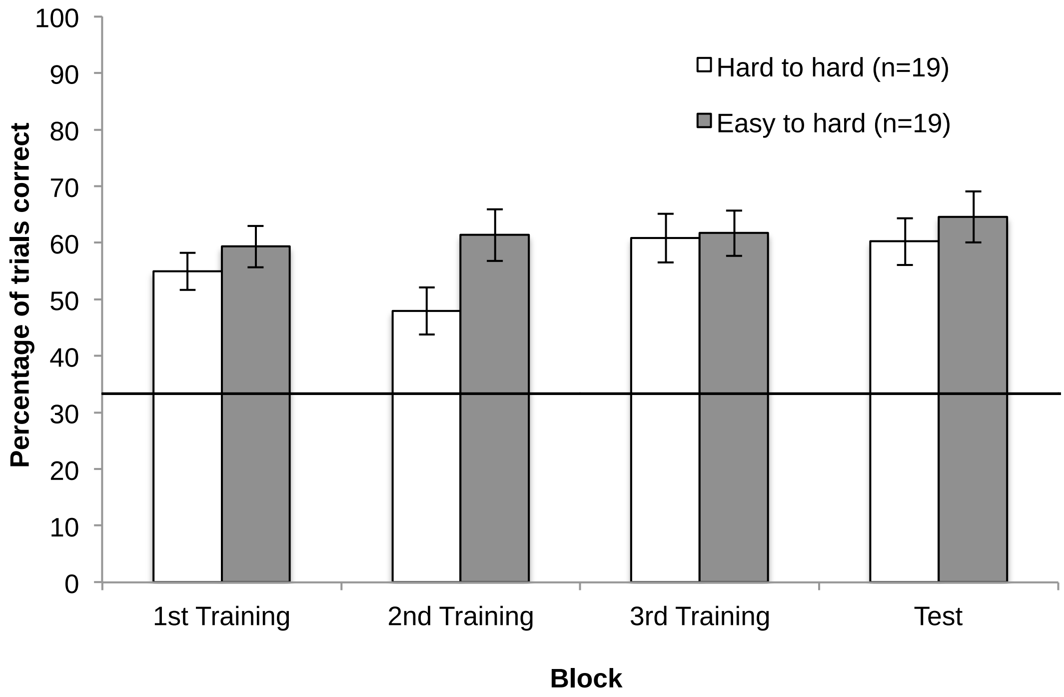

Figure A15. Mean (+/- SE) percentage of trials correct by block in Experiment A7.

Note: No feedback was given in the test block. The horizontal line indicates chance level of responding (33%). Participants were asked to match grape names to the samples.

Discussion

The results suggest that the hard condition was not difficult enough. That is, that any level of unique odorant was detectable and identifiable at a rate significantly greater than chance. This highlighted the need to test the concentrations before running the same experiment using expensive wine stimuli. This testing is outlined in Chapter 8

APPENDIX K

Wines used in each experiment

Experiment 4 – conducted in April 2009

Yalumba ‘‘Galway Vintage’’ Shiraz (2007) – Barossa Valley

Yering Station ‘‘Mr. Frog’’ Pinot Noir (2007) – Yarra Valley

Donelli Reggiano Lambrusco (non-vintage) – Emilia-Romagna, Italy

Experiment 5 – conducted in May 2009

Yalumba ‘‘Galway Vintage’’ Shiraz (2007) – Barossa Valley

Yering Station ‘‘Mr. Frog’’ Pinot Noir (2007) – Yarra Valley

Donelli Reggiano Lambrusco (non-vintage) – Emilia-Romagna, Italy

Experiment 6 – conducted May 2011

Moppity “Lock and Key” Shiraz (2009) – Hilltops

Yering Station ‘‘Mr. Frog’’ Pinot Noir (2009) – Yarra Valley

McWilliam’s “Hanwood” Cabernet Sauvignon (2009) – Multi-regional

Experiment 7 – conducted September 2011

Moppity “Lock and Key” Shiraz (2009) – Hilltops

Yering Station ‘‘Mr. Frog’’ Pinot Noir (2009) – Yarra Valley

McWilliam’s “Hanwood” Cabernet Sauvignon (2009) – Multi-regional

Experiment 8 – conducted October 2009

Yalumba ‘‘Mawson’s” Sauvignon Blanc (2008) - Wrattonbully

Yering Station ‘‘Mr. Frog’’ Chardonnay (2008) – Yarra Valley

Leasingham “Exclusive Release” Riesling (2007) – Clare Valley

Experiment 9 – conducted September 2009

Yalumba ‘‘Y Series” Sauvignon Blanc (2008) – South Australia

Yering Station ‘‘Mr. Frog’’ Chardonnay (2008) – Yarra Valley

Yalumba “Galway” Shiraz (2008) – Barossa Valley

Experiment 11 – conducted April 2010

Moppity “Lock and Key” Shiraz (2009) – Hilltops

Yalumba “Galway” Shiraz (2008) – Barossa Valley

Rymill “Yearling” Shiraz (2008) – Coonawarra

Redbank “Long Paddock” Shiraz (2007) – Victoria

Rymill “Yearling” Cabernet Sauvignon (2008) – Coonawarra

Yalumba “Y Series” Cabernet Sauvignon (2007) – South Australia

Yalumba “Mawson’s” Cabernet Sauvignon (2008) – South Australia

McWilliam’s “Hanwood” Cabernet Sauvignon (2008) – Multi-regional

Yering Station “Mr Frog” Pinot Noir (2008) – Yarra Valley

Josef Chromy “Pepik” Pinot Noir (2009) – Tasmania

Oyster Bay Pinot Noir (2008) – Marlborough, New Zealand

De Bortoli “Windy Peak” Pinot Noir (2008) – Yarra Valley

Experiment 12 – conducted August 2010

Moppity “Lock and Key” Shiraz (2009) – Hilltops

McWilliam’s “Hanwood” Cabernet Sauvignon (2008) – Multi-regional

Yering Station “Mr Frog” Pinot Noir (2008) – Yarra Valley

Experiment 13 – conducted August 2011

Moppity “Lock and Key” Shiraz (2009) – Hilltops

McWilliam’s “Hanwood” Cabernet Sauvignon (2008) – Multi-regional

Yering Station “Mr Frog” Pinot Noir (2008) – Yarra Valley

Experiment 14 – conducted May 2012

Yalumba “Galway” Shiraz (2011) – Barossa Valley

Yering Station “Mr Frog” Pinot Noir (2010) – Yarra Valley

Donelli Red Lambrusco (non-vintage) – Emilia-Romagna, Italy

Experiment 15 – conducted August 2012

Yalumba “Galway” Shiraz (2011) – Barossa Valley

Yering Station “Mr Frog” Pinot Noir (2010) – Yarra Valley

Donelli Red Lambrusco (non-vintage) – Emilia-Romagna, Italy

Experiment 16 – conducted August 2012

Yalumba “Galway” Shiraz (2011) – Barossa Valley

Yering Station “Mr Frog” Pinot Noir (2010) – Yarra Valley

McWilliam’s “Hanwood” Cabernet Sauvignon (2009) – Multi-regional

Plunkett Fowles “Stonedweller’s” Shiraz (2008) – Strathbogie Ranges

Bourke Street Pinot Noir (2011) – Canberra

Lindeman’s “Bin 45” Cabernet Sauvignon (2011) – South Eastern Australia

Experiment 17 – conducted May 2013

Yalumba “Galway” Shiraz (2011) – Barossa Valley

Tanunda Hill Shiraz (2011) – Barossa Valley

Yering Station “Mr Frog” Pinot Noir (2012) – Yarra Valley

Punt Road “Emperor’s Prize” Pinot Noir (2010) – Yarra Valley

Experiment A2 – conducted August 2009

Yalumba “Y Series” Sauvignon Blanc (2008) – South Australia

Yering Station “Mr Frog” Chardonnay (2007) – Yarra Valley

Jacob’s Creek Riesling (2008) – South Eastern Australia

Experiment A3 – conducted October 2009

Yalumba “Y Series” Sauvignon Blanc (2008) – South Australia

Yering Station “Mr Frog” Chardonnay (2007) – Yarra Valley

Hugel “Gentil” (2008) – Alsace, France

Experiment A5 – conducted October 2009

Yalumba “Galway” Shiraz (2008) – Barossa Valley

Yering Station “Mr Frog” Pinot Noir (2008) – Yarra Valley

Experiment A6 – conducted May 2010

Yering Station “Mr Frog” Pinot Noir (2008) – Yarra Valley

Yering Station “Mr Frog” Chardonnay (2008) – Yarra Valley

APPENDIX L

Ratings for first and second presentation of wines in Experiment 11

Table A29. Mean ratings for each of the twelve wines in terms of strength of pepper, blackcurrant, blueberry and cherry odours in Experiment 11.

	First rating				Second rating			
Wine	Pepper	Black-currant	Blue-berry	Cherry	Pepper	Black-currant	Blue-berry	Cherry
Shiraz 1	**2.53**	2.84	2.37	2.47	**2.58**	2.89	2.47	2.32
Shiraz 2	**2.37**	2.79	2.68	2.58	**2.05**	3.26	2.84	2.53
Shiraz 3	**2.32**	2.79	2.11	2.21	**2.68**	3.05	2.42	2.74
Shiraz 4	**2.05**	2.53	2.63	2.58	**2.26**	2.74	2.63	2.53
Pinot 1	2.21	2.68	2.79	**2.58**	2.53	2.47	2.74	**2.95**
Pinot 2	2.00	2.53	2.47	**2.95**	2.42	2.95	2.53	**2.42**
Pinot 3	2.16	3.11	2.37	**2.32**	2.26	2.63	2.68	**2.58**
Pinot 4	2.00	3.00	2.89	**2.32**	2.21	2.63	2.53	**2.84**
Cabernet 1	2.89	**2.68**	**1.89**	2.42	2.21	**2.32**	**2.63**	2.68
Cabernet 2	2.74	**2.84**	**2.68**	2.37	3.05	**2.42**	**2.00**	2.58
Cabernet 3	2.47	**3.05**	**2.11**	2.26	2.53	**2.74**	**2.32**	2.47
Cabernet 4	2.11	**3.47**	**2.58**	1.84	2.26	**2.79**	**2.63**	2.74

Note: Ratings were collected using a 7-point line scale ranging from 1 (not at all) to 7 (extremely). Scores in bold represent the appropriate odour for each wine. Shiraz 1: Lock and Key, Shiraz 2: Yalumba “Galway”, Shiraz 3: Rymill “Yearling”, Shiraz 4: Redbank “Long Paddock”. Cabernet 1: Rymill “Yearling”, Cabernet 2: Yalumba “Y Series”, Cabernet 3: Yalumba “Mawson’s”, Cabernet

4: McWilliam’s “Hanwood”. Pinot 1: Yering Station “Mr Frog”, Pinot 2: Josef Chromy “Pepik”, Pinot 3: Oyster Bay, Pinot 4: De Bortoli “Windy Peak”.

.

References

Albrecht, J., Anzinger, A., Kopietz, R., Schoepf, V., Kleemann, A. M., Pollatos, O., & Wiesmann, M. (2008). Test-retest reliability of the olfactory detection threshold test of the sniffin' sticks. *Chemical Senses, 33*(5), 461-467. doi: 10.1093/chemse/bjn013

American Society for Testing and Materials. (1981). Guidelnes for the Selection and Training of Sensory Panel Members, ASTM STP 758 *Committee E-18 on Sensory Evaluation of Materials and Products* (pp. 1-35): American Society for Testing and Materials.

Amerine, M. A., Pangborn, R. M., & Roessler, E. B. (1965). Principles of sensory evaluation of food *Food Science and Technology Monographs* (pp. 338-339). New York, New York: Academic Press.

Atanasova, B., Thomas-Danguin, T., Chabanet, C., Langlois, D., Nicklaus, S., & Etiévant, P. (2005). Perceptual Interactions in Odour Mixtures: Odour Quality in Binary Mixtures of Woody and Fruity Wine Odorants. *Chemical Senses*. doi: 10.1093/chemse/bji016

Australian Bureau of Statistics. (2013). Australian Wine and Grape Industry, 2012-13. Canberra, Australia: Australian Bureau of Statistics.

Aznar, M., Lopez, R., Cacho, J. F., & Ferreira, V. (2001). Identification and quantification of impact odorants of aged red wines from Rioja. GC-olfactometry, quantitative GC-MS, and odor evaluation of HPLC fractions. *Journal of Agricultural and Food Chemistry, 49*(6), 2924-2929. doi: 10.1021/jf001372u

Baccino, T., Cabrol-Bass, D., Candau, J., Meyer, C., Scheer, T., Vuillaume, M., & Wathelet, O. (2010). Sharing an olfactory experience: The impact of oral communication. *Food Quality and Preference, 21*(5), 443-452. doi: 10.1016/j.foodqual.2009.11.001

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*(4), 203-213. doi: 10.1007/s12078-009-9058-0

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*(4), 351-359. doi: 10.1016/j.foodqual.2004.06.001

Ballester, J., Patris, B., Symoneaux, R., & Valentin, D. (2008). Conceptual vs. perceptual wine spaces: Does expertise matter? *Food Quality and Preference, 19*(3), 267-276. doi: 10.1016/j.foodqual.2007.08.001

Bende, M., & Nordin, S. (1997). Perceptual Learning in Olfaction: Professional Wine Tasters versus Controls. *Physiology & Behavior, 62*(5), 1065-1070. doi: 10.1016/S0031-9384(97)00251-5

Bensafi, M., Rinck, F., Schaal, B., & Rouby, C. (2007). Verbal cues modulate hedonic perception of odors in 5-year-old children as well as in adults. *Chemical Senses, 32*(9), 855-862. doi: 10.1093/chemse/bjm055

Berg, H. W., Filipello, F., Hinreiner, E., & Webb, A. (1955). Evaluation of thresholds and minimum difference concentrations for various constituents of wine. *Food Technology, 9*, 23-26.

Betts, R. (2013). *The Essential Scratch and Sniff Guide to Becoming a Wine Expert: Take a Whiff of That*. Boston, MA: Houghton Mifflin Harcourt Publishing Company.

Bey, C., & McAdams, S. (2002). Schema-based processing in auditory scene analysis. *Perception & Psychophysics, 64*(5), 844-854. doi: 10.3758/bf03194750

Bitnes, J., Ueland, O., Moller, P., & Martens, M. (2008). Reliability of sensory assessors: Issues of retention and learning. *Journal of Sensory Studies, 23*(6), 852-870. doi: 10.1111/j.1745-459X.2008.00191.x

Blackman, J., Saliba, A., & Schmidtke, L. (2010). Sweetness acceptance of novices, experienced consumers and winemakers in Hunter Valley Semillon wines. *Food Quality and Preference, 21*(7), 679-683. doi: 10.1016/j.foodqual.2010.05.001

Borg, I., Groenen, P. J. F., & Mair, P. (2012). *Applied Multidimensional Scaling*: Springer.

Bregman, A. S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, Massachusetts: Bradford Books.

Cain, W. S. (1977). Differential sensitivity for smell: "noise" at the nose. *Science, 195*(4280), 796-798. doi: 10.1126/science.836592

Cain, W. S. (1979). To know with the nose: Keys to odor identification. *Science, 203*(4379), 467-470. doi: 10.1126/science.760202

Calkin, R. R., & Jellinek, J. S. (1994). *Perfumery: Practice and Principles*. New York, New York: Wiley.

Carlucci, A., & Monteleone, E. (2008). A procedure of sensory evaluation for describing the aroma profile of single grape variety wines. *Journal of Sensory Studies, 23*(6), 817-834. doi: 10.1111/j.1745-459X.2008.00189.x

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*(7-8), 562-571. doi: 10.1016/j.foodqual.2006.03.020

Cavanagh, P. (1999). Top-down processing in vision. In A. W. Robert & C. K. Frank (Eds.), *MIT Encyclopedia of Cognitive Science* (pp. 844-845). Cambridge, MA: MIT Press.

Chase, W., & Simon, H. (1973). Perception in Chess. *Cognitive Psychology, 5*, 55-81.

Chollet, S., Valentin, D., & Abdi, H. (2005). Do trained assessors generalize their knowledge to new stimuli? *Food Quality and Preference, 16*(1), 13-23. doi: 10.1016/j.foodqual.2003.12.003

Clarke, A. D. F., Elsner, M., & Rohde, H. (2013). Where's Wally: The influence of visual salience on referring expression generation. *Frontiers in Psychology, 4*. doi: 10.3389/fpsyg.2013.00329

The Complete Wine Course. (2008). from http://www.freewinecourse.com

Curby, K. M., & Gauthier, I. (2010). To the trained eye: Perceptual expertise alters visual processing. *Topics in Cognitive Science, 2*(2), 189-201.

Dacremont, C., & Valentin, D. (2004). Mobilization of short-term memory capacity for odors in discriminative tests: implication for assessors' selection. *Food Quality and Preference, 15*(7-8), 735-742. doi: 10.1016/j.foodqual.2004.02.005

Dalton, P., Doolittle, N., & Breslin, P. A. S. (2002). Gender-specific induction of enhanced sensitivity to odors. *Nature Neuroscience, 5*, 199-200.

Davis, R. G. (1981). The role of non-olfactory context cues in odor identification. *Perception & Psychophysics, 30*(1), 83-89. doi: 10.3758/bf03206139

de Groot, A. (1978). *Thoughts and choice in chess (2nd ed.)*. The Hague, Netherlands: Mouton (originally published in 1946).

Desor, J. A., & Beauchamp, G. K. (1974). Human capacity to transmit olfactory information. *Perception & Psychophysics, 16*(3), 551-556. doi: 10.3758/bf03198586

Doty, R. L., Huggins, G. R., Synder, P. J., & Lowry, L. D. (1981). Endocrine, Cardiovascular, and Psychological Correlates of Olfactory Sensitivity Changes During the Human Menstrual Cycle. *Journal of Comparative and Physiological Psychology, 95*(1), 45-60.

Doty, R. L., Shaman, P., & Dann, M. (1984). Development of the University of Pennsylvania Smell Identification Test - a standardized microencapsulated test of olfactory function. *Physiology & Behavior, 32*(3), 489-502. doi: 10.1016/0031-9384(84)90269-5

Dwyer, D. M., Hodder, K. I., & Honey, R. C. (2004). Perceptual learning in humans: Roles of preexposure schedule, feedback, and discrimination assay. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology, 57*(3), 245-259. doi: 10.1080/02724990344000114

Edwards, F. (1986). The marketing of wine from small wineries: Managing the intangibles. *International Journal of Wine Marketing, 1*, 1-6.

Engen, T. (1982). *The Perception of Odors*: Elsevier Science.

Engen, T., Kuisma, J. E., & Eimas, P. D. (1973). Short-term memory of odors. *Journal of Experimental Psychology, 99*(2), 222-225. doi: 10.1037/h0034645

Engen, T., & Ross, B. M. (1973). Long-term-memory of odors with and without verbal descriptions. *Journal of Experimental Psychology, 100*(2), 221-227. doi: 10.1037/h0035492

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*(4), 533-541. doi: 10.1016/j.foodcont.2009.08.013

Fernando, A. B., Murray, J. E., & Milton, A. L. (2013). The amygdala: securing pleasure and avoiding pain. *Front Behav Neurosci, 7*, 190. doi: 10.3389/fnbeh.2013.00190

Findlay, C. J., Castura, J. C., & Lesschaeve, I. (2007). Feedback calibration: A training method for descriptive panels. *Food Quality and Preference, 18*(2), 321-328. doi: 10.1016/j.foodqual.2006.02.007

Findlay, C. J., Castura, J. C., Schlich, P., & Lesschaeve, I. (2006). Use of feedback calibration to reduce the training time for wine panels. *Food Quality and Preference, 17*(3-4), 266-276. doi: 10.1016/j.foodqual.2005.07.005

Frank, M. E., Goyert, H. F., & Hettinger, T. P. (2010). Time and intensity factors in identification of components of odor mixtures. *Chemical Senses, 35*(9), 777-787. doi: 10.1093/chemse/bjq078

Frank, R. A., Rybalsky, K., Brearton, M., & Mannea, E. (2011). Odor Recognition Memory as a Function of Odor-Naming Performance. *Chemical Senses, 36*(1), 29-41. doi: 10.1093/chemse/bjq095

Galton, F. (1869). *Hereditary genius*. London: Macmillan and Company.

Gauthier, I., James, T. W., Curby, K. M, & Tarr, M. J. (2003). The influence of conceptual knowledge on visual discrimination. *Cognitive Neuropsychology, 20*(3-6), 507-523.

Gawel, R. (1997). The use of language by trained and untrained experienced wine tasters. *Journal of Sensory Studies, 12*(4), 267-284. doi: 10.1111/j.1745-459X.1997.tb00067.x

Gawel, R., & Godden, P. W. (2008). Evaluation of the consistency of wine quality assessments from expert wine tasters. *Australian Journal of Grape and Wine Research, 14*(1), 1-8. doi: 10.1111/j.1755-0238.2008.00001.x

Gawel, R., Royal, T., & Leske, P. (2002). The effect of different oak types on the sensory properties of Chardonnay. *Australian & New Zealand Wine Industry Journal, 17*(1), 14-20.

Gilbert, A. (2008). *What the Nose Knows: The Science of Scent in Everyday Life*. New York, New York: Crown Publishing Group.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585-612. doi: 10.1146/annurev.psych.49.1.585

Goodall, J., & Eyres, H. (2013). *The Bluffer's Guide to Wine*. London: Bluffers Guides Limited.

Goodstein, E. S., Bohlscheid, J. C., Evans, M., & Ross, C. F. (2014). Perception of flavor finish in model white wine: A time-intensity study.

Food Quality and Preference, 36(0), 50-60. doi: 10.1016/j.foodqual.2014.02.012

Gottfried, J. A. (2010). Central mechanisms of odour object perception. *Nature Reviews Neuroscience, 11*(9), 628-641. doi: 10.1038/nrn2883

Hanni, T., & Utermohlen, V. (2010). Beverage preferences, attitudes, & behavior of ‘sweet’ vs.‘tolerant’ wine consumers. Retrieved 23/8/2014, from http://www.timhanni.com/ConsumerStudySummary.pdf

Hayes, J. E., & Pickering, G. J. (2012). Wine Expertise Predicts Taste Phenotype. *American Journal of Enology and Viticulture, 63*(1), 80-84. doi: 10.5344/ajev.2011.11050

Hedner, M., Larsson, M., Arnold, N., Zucco, G. M., & Hummel, T. (2010). Cognitive factors in odor detection, odor discrimination, and odor identification tasks. *Journal of Clinical and Experimental Neuropsychology, 32*(10), 1062-1067. doi: 10.1080/13803391003683070

Henning, H. (1916). *Der Geruch*. Leipzig: J. A. Barth.

Hersleth, M., Mevik, B. H., Naes, T., & Guinard, J. X. (2003). Effect of contextual factors on liking for wine - use of robust design methodology. *Food Quality and Preference, 14*(7), 615-622. doi: 10.1016/s0950-3293(02)00190-8

Herz, R. S., & von Clef, J. (2001). The influence of verbal labeling on the perception of odors: Evidence for olfactory illusions? *Perception, 30*(3), 381-391. doi: 10.1068/p3179

Heymann, H., Hopfer, H., & Bershaw, D. (2014). An Exploration of the Perception of Minerality in White Wines by Projective Mapping and

Descriptive Analysis. *Journal of Sensory Studies, 29*(1), 1-13. doi: 10.1111/joss.12076

Hopfer, H., & Heymann, H. (2014). Judging wine quality: Do we need experts, consumers or trained panelists? *Food Quality and Preference, 32*, 221-233.

Howell, K. S., Swiegers, J. H., Elsey, G. M., Siebert, T. E., Bartowsky, E. J., Fleet, G. H., . . . de Barros Lopes, M. A. (2004). Variation in 4-mercapto-4-methyl-pentan-2-one release by Saccharomyces cerevisiae commercial wine strains. *FEMS Microbiology Letters, 240*, 125-129.

Hughson, A. L. (2003). Wine expertise: current theories and findings regarding its nature and bases. *Food Australia, 55*(5), 193-196.

Hughson, A. L., & Boakes, R. A. (2001). Perceptual and cognitive aspects of wine expertise. *Australian Journal of Psychology, 53*(2), 103-108. doi: 10.1080/00049530108255130

Hughson, A. L., & Boakes, R. A. (2002a). The knowing nose: the role of knowledge in wine expertise. *Food Quality and Preference, 13*(7-8), 463-472.

Hughson, A. L., & Boakes, R. A. (2002b). The role of knowledge in the descriptive expertise of wine experts. *Australian Journal of Psychology, 54*(1), 54-54.

Hughson, A. L., & Boakes, R. A. (2009). Passive perceptual learning in relation to wine: Short-term recognition and verbal description. *Quarterly Journal of Experimental Psychology, 62*(1), 1-8. doi: 10.1080/17470210802214890

International Organization for Standardization. (2012). Sensory analysis - General guidelines for the selection, training and monitoring of selected assessors and expert sensory assessors *8586:2012(en)*. Geneva, Switzerland: International Organization for Standardization.

Jackson, R., & Jackson, R. S. (2009). *Wine Tasting: A Professional Handbook*. New York: Elsevier Science.

James, W. (1890). *The Principles of Psychology*. New York: H. Holt.

Jehl, C., Royet, J. P., & Holley, A. (1997). Role of verbal encoding in short- and long-term odor recognition. *Perception & Psychophysics, 59*(1), 100-110. doi: 10.3758/bf03206852

Jiang, B., & Zhang, Z. (2010). Volatile compounds of young wines from cabernet sauvignon, cabernet gernischet and chardonnay varieties grown in the loess plateau region of china. *Molecules, 15*(12), 9184-9196. doi: 10.3390/molecules15129184

Jinks, A., & Laing, D. G. (1999a). A limit in the processing of components in odour mixtures. *Perception, 28*(3), 395-404. doi: 10.1068/p2898

Jinks, A., & Laing, D. G. (1999b). Temporal processing reveals a mechanism for limiting the capacity of humans to analyze odor mixtures. *Cognitive Brain Research, 8*(3), 311-325. doi: 10.1016/s0926-6410(99)00034-8

Jinks, A., & Laing, D. G. (2001). The analysis of odor mixtures by humans: evidence for a configurational process. *Physiology & Behavior, 72*(1–2), 51-63. doi: 10.1016/S0031-9384(00)00407-8

Jonsson, F. U., Tchekhova, A., Lonner, P., & Olsson, M. J. (2005). A metamemory perspective on odor naming and identification. *Chemical Senses, 30*(4), 353-365. doi: 10.1093/chemse/bji030

Kemp, S., Hollowood, T., & Hort, J. (2011). *Sensory Evaluation: A Practical Handbook*: Wiley.

King, E. S., & Heymann, H. (2014). The Effect of Reduced Alcohol on the Sensory Profiles and Consumer Preferences of White Wine. *Journal of Sensory Studies, 29*(1), 33-42. doi: 10.1111/joss.12079

Koza, B. J., Cilmi, A., Dolese, M., & Zellner, D. A. (2005). Color enhances orthonasal olfactory intensity and reduces retronasal olfactory intensity. *Chemical Senses, 30*(8), 643-649. doi: 10.1093/chemse/bji057

Kreider, P. (2011). *Uncorked: The Novice's Guide to Wine*. Nashville, Tennessee: Turner Publishing Company.

Kreutzmann, S., Thybo, A. K., & Bredie, W. L. P. (2007). Training of a sensory panel and profiling of winter hardy and coloured carrot genotypes. *Food Quality and Preference, 18*(3), 482-489. doi: 10.1016/j.foodqual.2006.05.009

Labbe, D., Damevin, L., Vaccher, C., Morgenegg, C., & Martin, N. (2006). Modulation of perceived taste by olfaction in familiar and unfamiliar beverages. *Food Quality and Preference, 17*(7-8), 582-589. doi: 10.1016/j.foodqual.2006.04.006

Laing, D. G., Eddy, A., & Best, D. J. (1994). Perceptual characteristics of binary, trinary, and quaternary odor mixtures consisting of unpleasant constituents. *Physiology & Behavior, 56*(1), 81-93. doi: 10.1016/0031-9384(94)90264-x

Laing, D. G., & Francis, G. W. (1989). The capacity of humans to identify odors in mixtures. *Physiology & Behavior, 46*(5), 809-814. doi: 10.1016/0031-9384(89)90041-3

Laing, D. G., & Glemarec, A. (1992). Selective attention and the perceptual analysis of odor mixtures. *Physiology & Behavior, 52*(6), 1047-1053. doi: 10.1016/0031-9384(92)90458-e

Lange, C., Martin, C., Chabanet, C., Combris, P., & Issanchou, S. (2002). Impact of the information provided to consumers on their willingness to pay for Champagne: comparison with hedonic scores. *Food Quality and Preference, 13*(7-8), 597-608. doi: 10.1016/s0950-3293(02)00059-9

Langlois, J., Dacremont, C., Peyron, D., Valentin, D., & Dubois, D. (2011). Lexicon and types of discourse in wine expertise: The case of vin de garde. *Food Quality and Preference, 22*(6), 491-498. doi: 10.1016/j.foodqual.2010.10.008

LaTour, K. A., LaTour, M. S., & Feinstein, A. H. (2011). The Effects of Perceptual and Conceptual Training on Novice Wine Drinkers' Development. *Cornell Hospitality Quarterly, 52*(4), 445-457. doi: 10.1177/1938965511420695

Lawless, H. T. (1984). Flavor Description of White Wine by Expert and Nonexpert Wine Consumers. *Journal of Food Science, 49*(1), 120-123. doi: 10.1111/j.1365-2621.1984.tb13686.x

Lawless, H. T. (1999). Descriptive analysis of complex odors: reality, model or illusion? *Food Quality and Preference, 10*(4–5), 325-332. doi: 10.1016/S0950-3293(98)00052-4

Lawless, H. T., & Engen, T. (1977). Associations to odors - interference, mnemonics, and verbal labeling. *Journal of Experimental Psychology-*

Human Learning and Memory, 3(1), 52-59. doi: 10.1037//0278-7393.3.1.52

Lawless, H. T., & Heymann, H. (2010). *Sensory Evaluation of Food: Principles and Practices*: Springer.

Le Nez du Vin. (2014). Le Nez du Vin. Retrieved 27th August, 2014, from http://www.lenez.com/en/home/

Lehrer, A. (1975). Talking About Wine. *Language, 51*(4), 901-923. doi: 10.2307/412700

Lehrer, A. (1983). *Wine and Conversation*. Bloomingham, Indiana: Indiana University Press.

Lehrner, J. P. (1993). Gender differences in long-term odor recognition memory: Verbal versus sensory influences and the consistency of label use. *Chemical Senses, 18*(1), 17-26. doi: 10.1093/chemse/18.1.17

Lehrner, J. P., Gluck, J., & Laska, M. (1999). Odor identification, consistency of label use, olfactory threshold and their relationships to odor memory over the human lifespan. *Chemical Senses, 24*(3), 337-346. doi: 10.1093/chemse/24.3.337

Lelievre, 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*(3), 143-153. doi: 10.1007/s12078-009-9050-8

Lindsay, P. H., & Norman, D. A. (1977). *Human information processing: An introduction to psychology*. Waltham, Massachusetts: Academic Press.

Livermore, A., & Laing, D. G. (1996). Influence of training and experience on the perception of multicomponent odor mixtures. *Journal of*

Experimental Psychology-Human Perception and Performance, 22(2), 267-277. doi: 10.1037/0096-1523.22.2.267

Livermore, A., & Laing, D. G. (1998). The influence of odor type on the discrimination and identification of odorants in multicomponent odor mixtures. *Physiology & Behavior, 65*(2), 311-320. doi: 10.1016/s0031-9384(98)00168-1

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's Congruence Coefficient as a Meaningful Index of Factor Similarity. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences, 2*(2), 57-64. doi: 10.1027/1614-2241.2.2.57

Lyman, B. J., & McDaniel, M. A. (1986). Effects of encoding strategy on long-term memory for odors. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 38*(4), 753-765.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82*(4), 276-298.

Marino-Sanchez, F. S., Alobid, I., Centellas, S., Alberca, C., Guilemany, J. M., Canals, J. M., . . . Mullol, J. (2010). Smell training increases cognitive smell skills of wine tasters compared to the general healthy population. The WINECAT Study. *Rhinology, 48*(3), 273-276. doi: 10.4193/Rhino09.206

Marti, M. P., Mestres, M., Sala, C., Busto, O., & Guasch, J. (2003). Solid-phase microextraction and gas chromatography olfactometry analysis of successively diluted samples. A new approach of the aroma extract dilution analysis applied to the characterization of wine aroma. *Journal* *of Agricultural and Food Chemistry, 51*(27), 7861-7865. doi: 10.1021/jf0345604

McDaniel, M. A., & Fisher, R. P. (1991). Tests and test feedback as learning sources. *Contemporary Educational Psychology, 16*(2), 192-201. doi: 10.1016/0361-476X(91)90037-L

Meilgaard, M. C., Carr, B. T., & Civille, G. V. (2006). *Sensory Evaluation Techniques, Fourth Edition*. United Kingdom: Taylor & Francis.

Meilgaard, M. C., Dalgliesh, C. E., & Clapperton, J. F. (1979). Beer Flavour Terminology. *Journal of the Institute of Brewing, 85*(1), 38-42. doi: 10.1002/j.2050-0416.1979.tb06826.x

Meilgaard, M. C., Reid, D. S., & Wyborski, K. A. (1982). Reference standards for beer flavour terminology system. *Journal of the American Society of Brewing Chemists, 40*(0119).

Melcher, J. M., & Schooler, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language, 35*(2), 231-245. doi: 10.1006/jmla.1996.0013

Millisecond Software LLC. (2011). Inquisit 3.0. Seattle, WA.

Mitchell, C., & Hall, G. (2014). Can theories of animal discrimination explain perceptual learning in humans? *Psychological Bulletin, 140*(1), 283-307. doi: 10.1037/a0032765

Morrot, G., Brochet, F., & Dubourdieu, D. (2001). The color of odors. *Brain and Language, 79*(2), 309-320. doi: 10.1006/brln.2001.2493

Næs, T., Brockhoff, P., & Tomic, O. (2011). *Statistics for Sensory and Consumer Science*. New Jersey: Wiley.

Ng, M., Chaya, C., & Hort, J. (2013). Beyond liking: Comparing the measurement of emotional response using EsSense Profile and consumer defined check-all-that-apply methodologies. *Food Quality and Preference, 28*(1), 193-205. doi: 10.1016/j.foodqual.2012.08.012

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*(2), 143-146.

Noble, A. C., Arnold, R. A., Masuda, B. M., Pecore, S. D., Schmidt, J. O., & Stern, P. M. (1984). Progress towards a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture, 35*(2), 107-109.

O'Neill, L., Nicklaus, S., & Sauvageot, F. O. (2003). A matching task as a potential technique for descriptive profile validation. *Food Quality and Preference, 14*(7), 539-547. doi: 10.1016/s0950-3293(02)00108-8

Olsson, M. J. (1994). An interaction model for odor quality and intensity. *Perception & Psychophysics, 55*(4), 363-372. doi: 10.3758/BF03205294

Olsson, M. J. (1999). Implicit Testing of Odor Memory: Instances of Positive and Negative Repetition Priming. *Chemical Senses, 24*(3), 347-350. doi: 10.1093/chemse/24.3.347

Owen, D. H., & Machamer, P. K. (1979). Bias-free improvement in wine discrimination. *Perception, 8*(2), 199-209.

Pagès, J. (2005). Collection and analysis of perceived product inter-distances using multiple factor analysis: Application to the study of 10 white wines from the Loire Valley. *Food Quality and Preference, 16*(7), 642-649. doi: 10.1016/j.foodqual.2005.01.006

Pangborn, R. M., Berg, H. W., & Hansen, B. (1963). The influence of color on discrimination of sweetness in dry table wine. *American Journal of Psychology, 76*, 492-495.

Paredes-Olay, C., Moreno-Fernandez, M. M., Rosas, J. M., & Ramos-Alvarez, M. M. (2010). ROC analysis in olive oil tasting: A Signal Detection Theory approach to tasting tasks. *Food Quality and Preference, 21*(5), 562-568. doi: 10.1016/j.foodqual.2010.03.003

Parr, W. V., Heatherbell, D., & White, K. G. (2002). Demystifying wine expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses, 27*(8), 747-755. doi: 10.1093/chemse/27.8.747

Parr, W. V., White, K. G., & Heatherbell, D. A. (2003). The nose knows: influence of colour on perception of wine aroma. *Journal of Wine Research, 14*(2/3), 79-101. doi: 10.1080/09571260410001677969

Parr, W. V., White, K. G., & Heatherbell, D. A. (2004). Exploring the nature of wine expertise: what underlies wine experts' olfactory recognition memory advantage? *Food Quality and Preference, 15*(5), 411-420. doi: 10.1016/j.foodqual.2003.07.002

Pavlov, I. P. (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. United Kingdom: Oxford University Press.

Perrin, L., Symoneaux, R., Maitre, I., Asselin, C., Jourjon, F., & Pages, J. (2008). Comparison of three sensory Napping (R) procedure: Case of methods for use with the ten wines from Loire valley. *Food Quality and Preference, 19*(1), 1-11. doi: 10.1016/j.foodqual.2007.06.005

Peynaud, E. (1987). *The Taste of Wine: The Art and Science of Wine Appreciation*. California: Wine Appreciation Guild.

Piggott, J. R., Simpson, S. J., & Williams, S. A. R. (1998). Sensory analysis. *International Journal of Food Science & Technology, 33*(1), 7-12. doi: 10.1046/j.1365-2621.1998.00154.x

Poucher, W. A. (1993). *Perfumes, Cosmetics and Soaps*: Springer Netherlands.

Prescott, J., Norris, L., Kunst, M., & Kim, S. (2005). Estimating a “consumer rejection threshold” for cork taint in white wine. *Food Quality and Preference, 16*(4), 345-349. doi: 10.1016/j.foodqual.2004.05.010

Rabin, M. D., & Cain, W. S. (1984). Odor recognition - familiarity, identifiability, and encoding consistency. *Journal of Experimental Psychology-Learning Memory and Cognition, 10*(2), 316-325. doi: 10.1037//0278-7393.10.2.316

Rabin, M. D., & Cain, W. S. (1986). Determinants of measured olfactory sensitivity. *Perception & Psychophysics, 39*(4), 281-286. doi: 10.3758/BF03204936

Richardson, J. T., & Zucco, G. M. (1989). Cognition and olfaction: a review. *Psychological Bulletin, 105*(3), 352-360.

Robinson, J. (2008). *How to Taste: A Guide to Enjoying Wine*. New York City, New York: Simon & Schuster.

Robinson, J., Harding, J., & Vouillamoz, J. (2012). *Wine grapes: A complete guide to 1,368 vine varieties, including their origins and flavours*. United Kingdom: Penguin.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long-term retention. *Psychol Sci, 17*(3), 249-255. doi: 10.1111/j.1467-9280.2006.01693.x

Ross, C. F., Weller, K. M., & Alldredge, J. R. (2012). Impact of Serving Temperature on Sensory Properties of Red Wine as Evaluated Using Projective Mapping by a Trained Panel. *Journal of Sensory Studies, 27*(6), 463-470. doi: 10.1111/joss.12011

Rouby, C., Pouliot, S., & Bensafi, M. (2009). Odor hedonics and their modulators. *Food Quality and Preference, 20*(8), 545-549. doi: 10.1016/j.foodqual.2009.05.004

Rouquier, S., Blancher, A., & Giorgi, D. (2000). The olfactory receptor gene repertoire in primates and mouse: Evidence for reduction of the functional fraction in primates. *Proceedings of the National Academy of Sciences, 97*(6), 2870-2874. doi: 10.1073/pnas.040580197

Royet, J., Plailly, J., Saive, A., Veyrac, A., & Delon-Martin, C. (2013). The impact of expertise in olfaction. *Frontiers in Psychology, 4*. doi: 10.3389/fpsyg.2013.00928

Rozin, P. (1982). "Taste-smell confusions" and the duality of the olfactory sense. *Perception & Psychophysics, 31*(4), 397-401.

Russell, A. M. T., & Boakes, R. A. (2011). Identification of confusable odours including wines: Appropriate labels enhance performance. *Food*

Quality and Preference, 22(3), 8-8. doi: 10.1016/j.foodqual.2010.11.007

Sage, A. (2002, 14 January). Cheeky little test exposes wine 'experts' as weak and flat, *Times*.

Schuster, M. (2009). *Essential Winetasting*. London, UK: Mitchell Beazley.

Shankar, M. U., Levitan, C. A., Prescott, J., & Spence, C. (2009). The Influence of Color and Label Information on Flavor Perception. *Chemosensory Perception, 2*(2), 53-58. doi: 10.1007/s12078-009-9046-4

Siebert, T. E., Wood, C., Elsey, G. M., & Pollnitz, A. P. (2008). Determination of rotundone, the pepper aroma impact compound, in grapes and wine. *Journal of Agricultural and Food Chemistry, 56*(10), 3745-3748. doi: 10.1021/jf800184t

Sigma-Aldrich. (2012). Aldrich Handbook - a Catalog of Fine Chemicals and Laboratory Equipment. In Sigma-Aldrich (Ed.). Miamisburg, Ohio.

Small, D. M., & Prescott, J. (2005). Odor/taste integration and the perception of flavor. *Exp Brain Res, 166*(3-4), 345-357. doi: 10.1007/s00221-005-2376-9

Solomon, G. E. A. (1988). Great expectorations: The psychology of novice and expert wine talk. *Unpublished Doctoral Dissertation, Harvard University*.

Solomon, G. E. A. (1990). Psychology of Novice and Expert Wine Talk. *American Journal of Psychology, 103*(4), 495-517. doi: 10.2307/1423321

Solomon, G. E. A. (1997). Conceptual change and wine expertise. *Journal of the Learning Sciences, 6*(1), 41-60. doi: 10.1207/s15327809jls0601_3

Spence, C. (2010). The multisensory perception of flavour. *Psychologist, 23*(9), 720-723.

Stevenson, R. J., Case, T. I., & Boakes, R. A. (2005). Implicit and explicit tests of odor memory reveal different outcomes following interference. *Learning and Motivation, 36*(4), 353-373. doi: 10.1016/j.lmot.2005.02.002

Stevenson, R. J., & Wilson, D. A. (2007a). Odour perception: An object-recognition approach. *Perception, 36*(12), 1821-1833. doi: 10.1068/p5563

Stevenson, R. J., & Wilson, D. A. (2007b). Odour perception: an object-recognition approach. *Perception, 36*(12), 1821-1833.

Stone, H., Bleibaum, R., & Thomas, H. A. (2012). *Sensory Evaluation Practices*. New York: Elsevier Science.

Sulmont-Rosse, C., Issanchou, S., & Koster, E. P. (2005). Odor naming methodology: Correct identification with multiple-choice versus repeatable identification in a free task. *Chemical Senses, 30*(1), 23-27. doi: 10.1093/chemse/bjh252

Symonds, M., & Hall, G. (1995). Perceptual learning in flavor aversion conditioning: Roles of stimulus comparison and latent inhibition of common stimulus elements. *Learning and Motivation, 26*(2), 203-219. doi: 10.1016/0023-9690(95)90005-5

Tempere, S., Cuzange, E., Malak, J., Bougeant, J. C., de Revel, G., & Sicard, G. (2011). The Training Level of Experts Influences their Detection

Thresholds for Key Wine Compounds. *Chemosensory Perception, 4*(3), 99-115. doi: 10.1007/s12078-011-9090-8

Tominaga, T., Masneuf, I., & Dubourdieu, D. (1995). A S-cysteine conjugate, precursor of aroma of white sauvignon. *Journal International des Sciences de la Vigne et du Vin, 29*, 227-232.

Tominaga, T., Murat, M. L., & Dubourdieu, D. (1998). Development of a method for analyzing the volatile thiols involved in the characteristic aroma of wines made from Vitis vinifera L. cv. Sauvignon Blanc. *Journal of Agricultural and Food Chemistry, 46*, 1044-1048.

Tominaga, T., Peyrot des Gachon, C., & Dubourdieu, D. (1998). A new type of flavor precursors in Vitis vinifera L. cv. Sauvignon Blanc: S-Cysteine Conjugates. *Journal of Agricultural and Food Chemistry, 46*, 5215-5219.

University of Adelaide. (2014). Syllabus for the Bachelor of Viticulture and Oenology, 2015. 25 August, 2014, from http://www.adelaide.edu.au/degree-finder/bvito_bvitoenol.html

Valentin, D., Chollet, S., & Abdi, H. (2003). Les mots du vin : experts et novices diffèrent-ils quand ils décrivent des vins ? *Corpus, 2*.

Verfaillie, K. (2000). Perceiving human locomotion: priming effects in direction discrimination. *Brain Cogn, 44*(2), 192-213. doi: 10.1006/brcg.2000.1228

Walk, H. A., & Johns, E. E. (1984). Interference and facilitation in short-term memory for odors. *Perception & Psychophysics, 36*(6), 508-514. doi: 10.3758/BF03207510

Walk, R. O. (1966). Perceptual learning and the discrimination of wines. *Psychonomic Science, 5*(2), 57-58. doi: 10.3758/BF03328278

Wansink, B., van Ittersum, K., & Painter, J. E. (2005). How descriptive food names bias sensory perceptions in restaurants. *Food Quality and Preference, 16*(5), 393-400. doi: 10.1016/j.foodqual.2004.06.005

Ward, L. M. (2013). The thalamus: gateway to the mind. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*(6), 609-622. doi: 10.1002/wcs.1256

Whalley, K. (2013). Neural coding: Timing is key in the olfactory system. *Nature Reviews Neuroscience, 14*(7), 458-458. doi: 10.1038/nrn3532

Wilson, D. A., & Stevenson, R. J. (2006). *Learning to Smell: Olfactory Perception from Neurobiology to Behavior*: Johns Hopkins University Press.

Wine Awakenings Inc. (2014). Wine Awakenings. Retrieved 27th August, 2014, from http://www.wineawakenings.com

Wolters, C. J., & Allchurch, E. M. (1994). Effect of training procedure on the performance of descriptive panels. *Food Quality and Preference, 5*(3), 203-214. doi: 10.1016/0950-3293(94)90036-1

Wood, C., Siebert, T. E., Parker, M., Capone, D. L., Elsey, G. M., Pollnitz, A. P., . . . Herderich, M. J. (2008). From wine to pepper: Rotundone, an obscure sesquiterpene, is a potent spicy aroma compound. *Journal of Agricultural and Food Chemistry, 56*(10), 3738-3744. doi: 10.1021/jf800183k

Yeomans, M. R., Chambers, L., Blumenthal, H., & Blake, A. (2008). The role of expectancy in sensory and hedonic evaluation: The case of smoked salmon ice-cream. *Food Quality and Preference, 19*(6), 565-573. doi: 10.1016/j.foodqual.2008.02.009

Zampini, M., Sanabria, D., Phillips, N., & Spence, C. (2007). The multisensory perception of flavor: Assessing the influence of color cues on flavor discrimination responses. *Food Quality and Preference, 18*(7), 975-984. doi: 10.1016/j.foodqual.2007.04.001

Zellner, D. A., McGarry, A., Mattern-McClory, R., & Abreu, D. (2008). Masculinity/femininity of fine fragrances affects color-odor correspondences: A case for cognitions influencing cross-modal correspondences. *Chemical Senses, 33*(2), 211-222. doi: 10.1093/chemse/bjm081

Zucco, G. M., Carassai, A., Baroni, M. R., & Stevenson, R. J. (2011). Labeling, identification, and recognition of wine-relevant odorants in expert sommeliers, intermediates, and untrained wine drinkers. *Perception, 40*(5), 598-607. doi: 10.1068/p6972