Verbal descriptions modulate flavour perception and evaluation of wine: An exploratory study using Riesling and Muscat Bailey A wines

Takuya Onuma*

Kindai University, Faculty of Humanity-oriented Science and Engineering, Department of Management and Business, Kayanomori 11-6, Iizuka, Fukuoka, Japan

ABSTRACT

Significant efforts have been made in the wine industry to assist customers in making purchase decisions and enhance their drinking experience by providing verbal descriptions of the complex sensory characteristics of wine. However, the influence of such verbal descriptions on wine flavour perception and evaluation has not been examined. To fill this gap, this exploratory study performed an experiment using white Riesling wine from the 2021 German vintage and red Muscat Bailey A (MBA) wine from the 2019 Japanese vintage. Forty-six Japanese individuals (23 social drinkers and 23 wine experts) participated and were presented with two glasses of Riesling wine and two glasses of MBA wine with different verbal descriptions. One glass of Riesling/MBA wine was served with a description including its distinct aroma descriptor (i.e., a ‘petrol’ note for the Riesling wine and a ‘strawberry candy’ note for the MBA wine). For the other glass, the description did not include the descriptor. The participants were asked to taste the wines and rate their palatability, perceived sweetness intensities, acidity, astringency, alcohol sensation, body, and expensiveness on a 7-point scale. The results demonstrated that the social drinkers rated the Riesling wine with the ‘petrol’ descriptor as heavier-bodied and more expensive compared to the same wine without the descriptor. The wine experts were not as influenced as the social drinkers, but the experts rated the Riesling wine with the ‘petrol’ descriptor as less expensive. In the case of MBA wine with the ‘strawberry candy’ descriptor, the social drinkers rated it as more expensive, whereas the wine experts rated it as more palatable than the same wine without the descriptor. The results indicated that verbal descriptions of wine aroma indeed modulated the flavour perception and evaluation of wine, but the modulation effect varied between social drinkers and wine experts.

KEYWORDS: wine, perception, consumer, verbal description, Riesling, Muscat Bailey A
INTRODUCTION

1. Use of verbal descriptions in the wine industry

Wine is undoubtedly one of the most complex beverages in the human culture. The taste and aromatic profile of wine are influenced by a wide range of viticultural and winemaking factors, leading to its diverse and complex nature. Hence, choosing the right wine at the point of sale can be challenging for consumers. To assist consumers in making decisions, wine descriptions and details are provided on labels, menus, and online shops (Danner et al., 2017). Charters et al. (1999) revealed that among the back-label information of the wine bottle, consumers found simple descriptions of the tastes or aromas of the wines to be the most useful in helping them identify the wines. Similarly, elaborate taste descriptions on the back label were found to be one of the most highly valued information for consumers (Mueller et al., 2010).

In the hospitality section, significant efforts are made to help customers choose the right wine and enhance their drinking experiences. Association de la Sommellerie Internationale (ASI) demands sommeliers to “learn to add value to the administrative, operative, and experiential elements of any restaurant business” (ASI, 2022). It is essential for sommeliers and other restaurant personnel to be able to evaluate and communicate the complex sensory experiences of wine using proper wine vocabulary (Herdenstam et al., 2018). However, such communication is challenging because of the lack of clear agreement on how wine terminology is used to describe wines (Lehrer, 1975). To facilitate proper communication in the wine industry, lists of standardised wine aroma terminologies have been developed and encouraged to be used (Gawel et al., 2000; Noble et al., 1987; Noble et al., 1984; Pickering and Demiglio, 2008). Although the exact impact on customers’ flavour perception remains to be clarified, the proper use of these terminologies is believed to facilitate better communication and enhance the wine experience.

2. Impact of external information on wine-drinking experiences

Evidence from sensory and consumer sciences has consistently shown that extrinsic factors, such as information, can affect how we expect and perceive the flavour and quality of food and drinks (Piqueras-Fiszman and Spence, 2015; Seo, 2020). For instance, Lange et al. (2002) demonstrated the strong impact of the Champagne labelling on consumer evaluations and preferences. The results revealed that consumers could not discriminate Champagnes following blind tasting, yet significant differences in preferences for the products emerged when the labels were revealed, and the preferences observed reflected the hierarchy of the market. Masson et al. (2008) demonstrated that low-alcohol labelling has a negative impact on consumer perception. In their experiment, a reduced alcohol content claim on a low-alcohol wine label negatively affected both the expected and perceived quality of the wine. However, in a blind-tasting setting, the panels rated the quality of low-alcohol wine to be similar to that of regular wine.

Similarly, the price of wine (Plasmann et al., 2008), PDO information (Otheguy et al., 2023), and critics’ wine ratings (Siegrist and Cousin, 2009) have been demonstrated to affect consumers’ responses to and evaluations of wine. These results indicate that external information provided prior to or during tasting leads drinkers to develop particular expectations regarding the aroma/flavour and quality of wine, which modulate how wine is perceived and evaluated during tasting (Piqueras-Fiszman and Spence, 2015). Therefore, it is reasonable to assume that providing a verbal description of wine aroma characteristics modulates tasters’ expectations, perceptions, and evaluations of the wine.

Few studies have explored this possibility. Danner et al. (2017) examined the impact of back-label descriptions on consumer experience and demonstrated that the presentation of taste and aroma descriptions before tasting wine increased consumers’ wine liking and willingness to pay for the wine and elicited more intense positive emotions compared to blind tasting. However, it should be noted that their study revealed the impact of the presence (or absence) of verbal descriptions in general but not the impact of an individual descriptor. Appropriately using each descriptor in terminology is crucial because each descriptor represents a distinct sensory characteristic (Gawel et al., 2000; Noble et al., 1987; Noble et al., 1984; Pickering and Demiglio, 2008). Therefore, it would be beneficial to investigate further how verbal descriptions affect the drinking experience at the individual descriptor level. Moreover, few studies have investigated the effect of verbal descriptions on flavour perceptions, such as acidity, sweetness, tannins, and body perception of wine. As previously indicated, expectations shaped by external information can modulate the perception and evaluation of wine (Piqueras-Fiszman and Spence, 2015); hence, this possibility should be tested.

3. Research purpose

Considering the gap between the common practice in the wine industry of verbally describing aroma characteristics and the scientific understanding of their influences, the primary goal of this exploratory study was to determine whether providing verbal descriptions of wine aroma characteristics modulates flavour perception and evaluation of wine. To achieve this goal, we focused on two grape varieties, Riesling and Muscat Bailey A, and aroma descriptors that are distinctively associated with these varieties. Riesling is a type of Vitis vinifera grape that is widely planted for aromatic white winemaking. It produces a range of aromas, including stone fruits, flowers, and petrol/kerosene. 1,1,6-Trimethyl-1,2-dihydropnaphthalene (TDN) has been identified as a key component of the petrol/kerosene aroma of Riesling wines (Sacks et al., 2012; Tarasov et al., 2020; Winterhalter, 1991). TDN is believed to enhance the complexity of the wine bouquet when present in low to medium concentrations, but high levels of TDN can lead to unfavourable impressions because of the overpowering aroma of petrol/kerosene (Tarasov et al., 2020). In the context of wine appreciation, Riesling wines are often described as having a ‘petrol’ note; for instance, Tom Parker MW wrote about the grape variety...
that “the distinct lime and petrol character, mouth-watering acidity, harmony of sweet and savoury, light body and residual sugar (in that Mosel Spätlese) make for one of the—if not the—easiest grape varieties to distinguish when just getting into wine” (Parker, 2020). There has been much debate among winemakers about whether the term ‘petrol’ should be employed to characterise Riesling wine (Schmidt, 2013). Nevertheless, the term is strongly associated with the grape variety and is commonly used in wine communication.

Muscat Bailey A (MBA) is a hybrid grapevine cultivar bred by crossing Vitis labrusca (Bailey) with Vitis vinifera (Muscat Hamburg). Since its development in Japan in 1927, MBA has been cultivated and widely used for Japanese red wine production (Kobayashi et al., 2013). The variety is well known for producing lighter-bodied wine containing much greater concentrations of furanone, contributing to the variety’s distinct ‘strawberry’ note (Kobayashi et al., 2013; Sasaki et al., 2015a; Sasaki et al., 2015b). The variety’s notable sweet aroma is also described as a ‘candy’ note (Goto-Yamamoto, 2019; Robinson, 2016).

In our experiment, Japanese participants, including social drinkers and wine experts working in the wine industry, were recruited and asked to taste Riesling and MBA wines in a blind-tasting setting. Before and during the tasting, we provided them with a range of aroma descriptions of the wine, some of which contained a distinct (target) descriptor of the grape variety (i.e., ‘petrol’ for Riesling and ‘strawberry candy’ for MBA), whereas others did not. By comparing participants’ responses to the wine presented with or without the target descriptors, we examined the impact of the descriptors on the perception and evaluation of the wine. Based on previous research demonstrating that external information modulates the perception and evaluation of wine (Danner et al., 2017; Lange et al., 2002; Masson et al., 2008; Otheguy et al., 2023; Plassmann et al., 2008; Siegrist and Cousin, 2009), we hypothesised that giving the wine’s distinct aroma descriptor would modulate the flavour perception and evaluation of the wine. Considering the difference in tasing skills between social drinkers and wine experts (Croijmans et al., 2021; Parr et al., 2002), we also hypothesised that the impact of verbal descriptions would vary between social drinkers and wine experts.

**MATERIALS AND METHODS**

1. **Participants**

Forty-six Japanese individuals were recruited for this study (Table 1). Among them, following the criteria of previous studies (Croijmans et al., 2021; Parr et al., 2003; Parr et al., 2002), 23 participants were defined as wine experts who worked as certified sommeliers, certified wine advisers, professional wine retailers, or employees of wine restaurants and wine glass companies. The remaining 23 participants were defined as social drinkers. The participants were recruited through social media networks and word-of-mouth advertising. The sample size was determined based on previous studies investigating group differences (i.e., wine experts vs social drinkers) in wine odour recognition memory (Croijmans et al., 2021) and susceptibility to wine colouring (Parr et al., 2003), which employed around 25 participants in each participant group.

All participants were paid a ¥1,000 shopping voucher. Verbal and written explanations of the experiment were provided to the participants, and written informed consent signed by the participants was obtained before the experiment. After the research, the participants were debriefed on the true purpose of the investigation and provided their written informed consent. This study was conducted in accordance with the Declaration of Helsinki for Research involving Human Subjects and was approved by the Ethics Committee of Humanity-Oriented Science and Engineering, Kindai University.

| TABLE 1. Characteristics and wine consumption habits of the participants. |
|-----------------------------|-----------------------------|-----------------------------|
|                             | Social drinkers (N = 23)    | Wine experts (N = 23)       |
| Gender (N)                  | Women                       | Men                         |
| Mean age (min–max)          | 36.2 [20–61]                | 35.8 [22–60]                |
| Wine drinking frequency [N] | Once a week                  | 3                           |
|                             | Several times a month       | 9                           |
|                             | At least once a month       | 11                          |

2. **Wines**

Two wines (one white and one red) were selected based on informal tasting sessions and discussions with two wine experts, both professional sommeliers certified by the Japan Sommelier Association and with more than 10 years of experience in the wine business. The white wine was solely made from Riesling grapes from the 2021 German vintage with 11.5 % abv, and the red wine was solely made from MBA grapes from the 2019 Japanese vintage with 10.0 % abv. The Riesling and MBA wines were judged by the wine experts as having a notable ‘petrol’ note (Riesling) or ‘strawberry candy’ note (MBA).

The bottles were opened at the beginning of each experimental day, checked for obvious cork taint, recapped by antioxidant closures, and chilled until service time. The bottles of Riesling and MBA wines were stored separately in temperature-controlled wine coolers so that the serving temperatures of the samples were approximately 11 °C and 13 °C, respectively. For each participant, two 50 ml samples of Riesling wine and two 50 ml samples of MBA wine were prepared in ISO standard wine-tasting glasses, resulting in four samples per participant.
3. Experimental design
This experiment involved two samples of Riesling wine and another two samples of MBA wine, meaning that the exact same sample was served in two different settings. Each time, the participants were provided with different verbal descriptions. For Riesling wine, one sample was described with general aroma characteristics (“This white wine has aromas of citrus fruit, peach, and flower.”), while the other sample was described with an additional descriptor ‘petrol’ (“This white wine has aromas of citrus fruit, peach, flower, and petrol.”). In this experiment, the former description was considered the control condition, whereas the latter was considered the target condition, which incorporated the distinct aroma character ‘petrol’ of the grape variety. For the MBA wine, the target condition description incorporated the target descriptor ‘strawberry candy’ (“This red wine has aromas of red fruit, blackberry, strawberry candy, and earth.”), whereas the control condition description did not (“This red wine has aromas of red fruit, blackberry, and earth.”). No other information about the wine, such as grape variety, country of origin, vintage, or price, was provided to the participants.

To ensure that the participants could not easily notice the exact same wine being presented twice, the Riesling and MBA wines were presented alternately. Therefore, half of the participants received samples in the order of Riesling, MBA, Riesling, and MBA. The other half received samples in the order of MBA, Riesling, MBA, and Riesling. The order of control and target conditions was counterbalanced throughout the experiment.

4. Procedure
The experiment was performed on a booth-like structure established in a shared workspace located in Fukuoka, Japan. The room temperature of the booth was maintained at approximately 24 °C (average temperature = 24.77, SD = 0.98), and the booth was kept free of noise, ambient odours, and salient furnishings. The participants were instructed not to eat anything for at least one hour before participation and not to wear perfume or cosmetic products with strong odours.

The participants were tested individually at the booth. After arrival, the participants were told that they were going to taste and rate a number of samples of commercial wines, but at this moment, they were not informed of how many samples they were going to taste. The participants then reviewed all rating scales they were going to use during the tasting session, which were presented on a touch screen. If anything was unclear to them, participants could ask the experimenter about the scales. Once it had been confirmed that all meanings of the scales were clear to the participants, the tasting session began.

During the tasting session, time control, presentation of instructions, and data collection were performed using the online questionnaire platform Pavlovia Surveys (https://pavlovia.org/docs/surveys/overview), which was run on a touchscreen computer. While the participants followed the instructions automatically presented on the screen, the experimenter monitored the participants and prepared the samples behind the curtain. First, the participants were asked to rinse their mouths with a bottle of mineral water and wait for 90 seconds while the experimenter poured the first sample of wine into a glass. Subsequently, a verbal description of the first sample appeared on the touch screen and the sample was served by the experimenter. The contents of the description depended on the experimental conditions allocated to the sample (i.e., target vs control condition). The participants touched the start button on the screen and they were asked to taste the sample freely and rate on a 7-point scale how palatable the wine was (1 = not at all, 7 = very palatable), the intensities of sweetness, acidity, astringency, and alcohol sensation (1 = very weak, 7 = very strong), body (1 = very light, 7 = very heavy), and expensiveness (1 = very cheap, 7 = very expensive). These rating scales were presented on the screen, and the participants responded by touching them.

During the rating, the participants were allowed to smell or taste the sample as many times as they wanted. Once the rating was completed, the participants touched the next button on the screen and the sample was immediately removed by the experimenter. They were again asked to rinse their mouths with water and rest for 90 seconds while the experimenter poured the next sample into a glass. Subsequently, a verbal description of the second sample appeared on the touch screen and the sample was served by the experimenter. This procedure was repeated for all four samples. The experiment lasted for approximately 20 min, and the participants were debriefed afterwards.

5. Data analysis
All data obtained in this study were analysed using the statistical software R [Version 4.1.2; R Core Team (2020)]. In total, 1840 observations were obtained (46 participants × four wine samples × seven rating attributes). To obtain clear and valid results, a median plus or minus 2.5 times the Median Absolute Deviation (MAD) method was used for outlier detection as this is a moderately conservative procedure and is strongly recommended as a reasonable default choice (Leys et al., 2013). MADs for each subgroup of rating data among the four samples were calculated and used. As a result, for the Riesling wine, 16 observations in the target condition and five observations in the control condition were excluded as outliers. For the MBA wine, 15 observations were excluded in the target condition and 6 observations in the control conditions.

Linear mixed models were adopted to treat multiple sources of random variation (Judd et al., 2017). The fixed factors in this study were the participant group (i.e., social drinker vs. wine expert), type of verbal description (i.e., target vs control condition), and their interactions. Although these three terms were specified as fixed factors to perform the modelling, our main interest was the effect of verbal description separately for each participant group and not a global effect or general group differences. Therefore, only paired comparisons between the target and control conditions within each participant group were performed and assessed in the following analyses.
Several random factors were specified together with the fixed factors. First, because the same participants repeated the tasting for multiple wine samples, the experimental design was the repeated-measures. Therefore, the ID of participants was set as a random intercept. Second, because the presentation order of the wines differed between participants, the effect of the presentation order on the response or evaluation of each wine was considered a random intercept. Linear mixed modelling was performed using the \textit{lmer} function in the \textit{lmerTest} package (Kuznetsova et al., 2017). The calculation of estimated marginal means and evaluation of the fixed factor (the verbal description) within each participant group were performed using the \textit{emmeans} function in the \textit{emmeans} package (Lenth, 2020).

To evaluate the estimated effect of the fixed factor, this study calculated the test statistic (t-value) and the probability (p-value) of obtaining the test statistics under the null hypothesis for the point estimates and also estimated the 95\% confidence interval (95 \% CI) for the effect size. Importantly, Snedecor and Cochran (1967) have cautioned against the conventional dichotomous decision procedure of null hypothesis testing, whereby researchers simply “reject” or “fail to reject” the null hypothesis based on a certain threshold of p-value (i.e., 0.05). This way of statistical testing, for example, concluding that the statistically “non-significant” results showed “no effect” or “no difference,” has also been criticised recently (Amrhein and Greenland, 2022; Amrhein et al., 2019). Therefore, following recent advice, this study calculated the counternull value, an effect size supported by exactly the same amount of evidence (p-value) as the null hypothesis (Rosenthal and Rubin, 1994).

The counternull value reminds us of the fact that the assertion that the obtained effect size is not significantly different from zero is no more true than the assertion that the effect size is not significantly different from the counternull value (Rosenthal and Rubin, 1994). These calculations and estimations were performed using the \textit{conf_dist} function in the \textit{pvaluefunctions} package in the R software (Infanger, 2021).

### RESULTS

#### 1. Ratings for the Riesling wine

Observed rating data for the Riesling wine are shown in Figure 1. Paired comparisons between the target and control conditions were performed within each participant group. The statistics obtained are shown in Table 2, and their p-value functions are shown in Figure 2. For the social drinkers, the most evident differences were observed for the body and expensiveness ratings. Compared to the control condition, the body rating for the target condition wine was 0.97 points higher (95 \% CI[0.40, 1.55], \( p = 0.001, \text{ counternull} = 1.94 \)), and the expensiveness rating was 0.58 points higher (95 \% CI[0.18, 0.97], \( p = 0.005, \text{ counternull} = 1.15 \)). In addition, the social drinkers rated the target condition wine as 0.48 points more palatable (95 \% CI[–0.09, 1.06], \( p = 0.095, \text{ counternull} = 0.97 \)), 0.48 points sweeter (95 \% CI[–0.14, 1.09], \( p = 0.124, \text{ counternull} = 0.96 \)), 0.40 points more astringent (95 \% CI[–0.15, 0.95], \( p = 0.151, \text{ counternull} = 0.80 \)). Although these differences were not “statistically significant” from the conventional dichotomous viewpoint (i.e., \( p > 0.05 \)), it should also be noted that, as their counternull values show, 0.97 points palatability enhancement, 0.96 points sweetness enhancement, and 0.80 points astringency enhancement were also supported by exactly the same amount of evidence as zero difference, respectively. The social drinkers also rated the target condition wine as 0.22 points less acidic (95 \% CI[–0.94, 0.51], \( p = 0.546, \text{ counternull} = –0.44 \)) and 0.21 points more alcoholic (95 \% CI[–0.15, 0.95], \( p = 0.406, \text{ counternull} = 0.42 \)); however, even though considering their counternull values (0.44 points acidity inhibition and 0.42 points alcohol enhancement), these differences were relatively small and negligible compared to the differences observed for the other rating items.

Contrary to the social drinkers, the wine experts rated the target condition wine as 0.36 points less expensive (95 \% CI[–0.74, 0.04], \( p = 0.073, \text{ counternull} = –0.71 \)) and

### Table 2. Results of the linear mixed modelling evaluating the fixed effect of the verbal description (the effect of the target condition against the control condition) for the ratings of the Riesling wine.

<table>
<thead>
<tr>
<th>Group</th>
<th>Attribute</th>
<th>Estimate</th>
<th>95 % CI</th>
<th>SE</th>
<th>t-value</th>
<th>p-value</th>
<th>Counternull</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social drinkers</td>
<td>Palatability</td>
<td>0.485</td>
<td>[–0.089, 1.058]</td>
<td>0.284</td>
<td>1.706</td>
<td>0.0953</td>
<td>0.969</td>
</tr>
<tr>
<td></td>
<td>Sweetness</td>
<td>0.478</td>
<td>[–0.136, 1.093]</td>
<td>0.305</td>
<td>1.569</td>
<td>0.1239</td>
<td>0.956</td>
</tr>
<tr>
<td></td>
<td>Acidity</td>
<td>–0.219</td>
<td>[–0.944, 0.506]</td>
<td>0.360</td>
<td>–0.609</td>
<td>0.5458</td>
<td>–0.438</td>
</tr>
<tr>
<td>Expensiveness</td>
<td>Astringency</td>
<td>0.398</td>
<td>[–0.150, 0.946]</td>
<td>0.272</td>
<td>1.464</td>
<td>0.1505</td>
<td>0.796</td>
</tr>
<tr>
<td></td>
<td>Alcohol</td>
<td>0.212</td>
<td>[–0.297, 0.721]</td>
<td>0.253</td>
<td>0.840</td>
<td>0.4058</td>
<td>0.424</td>
</tr>
<tr>
<td></td>
<td>Body</td>
<td>0.972</td>
<td>[0.395, 1.550]</td>
<td>0.286</td>
<td>3.396</td>
<td>0.0014</td>
<td>1.944</td>
</tr>
<tr>
<td>Wine experts</td>
<td>Palatability</td>
<td>–0.035</td>
<td>[–0.596, 0.527]</td>
<td>0.278</td>
<td>–0.125</td>
<td>0.9013</td>
<td>–0.069</td>
</tr>
<tr>
<td></td>
<td>Sweetness</td>
<td>0.174</td>
<td>[–0.437, 0.784]</td>
<td>0.303</td>
<td>0.575</td>
<td>0.5685</td>
<td>0.348</td>
</tr>
<tr>
<td></td>
<td>Acidity</td>
<td>–0.217</td>
<td>[–0.928, 0.493]</td>
<td>0.352</td>
<td>–0.618</td>
<td>0.5401</td>
<td>–0.435</td>
</tr>
<tr>
<td></td>
<td>Astringency</td>
<td>–0.041</td>
<td>[–0.569, 0.487]</td>
<td>0.261</td>
<td>–0.156</td>
<td>0.8771</td>
<td>–0.081</td>
</tr>
<tr>
<td></td>
<td>Alcohol</td>
<td>–0.345</td>
<td>[–0.851, 0.161]</td>
<td>0.251</td>
<td>–1.375</td>
<td>0.1763</td>
<td>–0.690</td>
</tr>
<tr>
<td></td>
<td>Body</td>
<td>–0.091</td>
<td>[–0.666, 0.485]</td>
<td>0.285</td>
<td>–0.318</td>
<td>0.7522</td>
<td>–0.181</td>
</tr>
<tr>
<td></td>
<td>Expensiveness</td>
<td>–0.355</td>
<td>[–0.745, 0.035]</td>
<td>0.193</td>
<td>–1.839</td>
<td>0.0729</td>
<td>–0.711</td>
</tr>
</tbody>
</table>
FIGURE 1. Boxplots of ratings for the Riesling wine.
White points represent the mean of ratings.

FIGURE 2. P-value functions of the estimated mean difference in the rating of Riesling wine between the target and control conditions.
The peak of each curve represents a point estimate of the mean difference. The vertical lines indicate null values (true difference = 0).
Circles represent the counternull values. The dashed horizontal line shows the conventional “significance” threshold, $p = 0.05$, and the area above the line shows the 95% CI of the estimated mean difference.
FIGURE 3. Boxplots of ratings for the MBA wine.
White points represent the mean of ratings.

FIGURE 4. $P$-value functions of the estimated mean difference in the rating of MBA wine between the target and control conditions.
The peak of each curve represents a point estimate of the mean difference. The vertical lines indicate null values (true difference = 0). Circles represent the counternull values. The dashed horizontal line shows the conventional “significance” threshold, $p = 0.05$, and the area above the line shows the 95 % CI of the estimated mean difference.
0.34 points less alcoholic (95 % CI [–0.85, 0.16], \( p = 0.176 \), counternull = –0.69). The difference in the body ratings was almost negligible (–0.09 points, 95 % CI [–0.67, 0.48], \( p = 0.752 \), counternull = –0.18). Likewise, the differences in palatability (–0.04 points, 95 % CI [–0.60, 0.53], \( p = 0.901 \), counternull = –0.07), sweetness (0.17 points, 95 % CI [–0.44, 0.78], \( p = 0.569 \), counternull = 0.35), acidity (–0.22 points, 95 % CI [–0.93, 0.49], \( p = 0.540 \), counternull = –0.44), and astringency (–0.04 points, 95 % CI [–0.57, 0.49], \( p = 0.877 \), counternull = –0.08) were relatively small and almost negligible even when considering their counternull values.

### 2. Ratings for the MBA wine

Observed rating data for the MBA wine are shown in Figure 3. Paired comparisons between the target and control conditions were performed within each participant group. The statistics obtained are shown in Table 3, and their p-value functions are shown in Figure 4. The social drinkers rated the target condition wine as 0.50 points more expensive (95 % CI [–0.85, 0.16], \( p = 0.176 \), counternull = 1.01), 0.36 points heavier-bodied (95 % CI [–0.31, 1.03], \( p = 0.281 \), counternull = 0.72), 0.32 points more alcoholic (95 % CI [–0.25, 0.89], \( p = 0.263 \), counternull = 0.64), 0.29 points more palatable (95 % CI [–0.25, 0.84], \( p = 0.281 \), counternull = 0.59), and 0.25 points more astringent (95 % CI [–0.44, 0.94], \( p = 0.467 \), counternull = 0.50). The differences in sweetness (–0.09 points, 95 % CI [–0.71, 0.52], \( p = 0.762 \), counternull = 0.19) and alcohol (–0.05 points, 95 % CI [–0.64, 0.53], \( p = 0.863 \), counternull = –0.10) were relatively small and almost negligible even when considering their counternull values.

The wine experts rated the target condition wine as 0.46 points more palatable (95 % CI [–0.07, 0.99], \( p = 0.086 \), counternull = 0.92). Although relatively small, they also rated the target condition wine as 0.28 points less astringent (95 % CI [–0.95, 0.39], \( p = 0.404 \), counternull = –0.56), 0.26 points more sweet (95 % CI [–0.32, 0.85], \( p = 0.373 \), counternull = 0.52), and 0.25 points less alcoholic (95 % CI [–0.82, 0.32], \( p = 0.386 \), counternull = –0.50). The differences in acidity (0.04 points, 95 % CI [–0.50, 0.58], \( p = 0.872 \), counternull = 0.09), body (–0.07 points, 95 % CI [–0.72, 0.59], \( p = 0.832 \), counternull = –0.14), and expensiveness (0.02 points, 95 % CI [–0.54, 0.58], \( p = 0.934 \), counternull = 0.05) were almost negligible even when considering their counternull values.

### DISCUSSION

#### 1. Overview

Significant efforts have been made in the wine industry to assist customers in making purchase decisions and enhance their drinking experience by providing verbal descriptions of the complex sensory characteristics of wine (ASI, 2022; Gawel et al., 2000; Herdenstam et al., 2018; Noble et al., 1984; Pickering and Demiglio, 2008). Nevertheless, the impact of such verbal descriptions on the drinking experience remains unclear. To address this gap, our exploratory study investigated whether providing verbal descriptions of wine aroma characteristics modulates flavour perception and wine evaluation. We recruited both social drinkers and wine experts and asked them to taste Riesling and MBA wines, accompanied by various descriptions of their aromas. The results revealed that, as hypothesised, the verbal descriptions presented prior to tasting had an impact on the flavour (sensory) perception and hedonic evaluation of wine, and this influence varied between social drinkers and wine experts. These results are important because 1) they demonstrate for the first time that a single aroma descriptor, not just a general description (Danner et al., 2017), affects the wine-drinking experience, and 2) most previous studies have reported the effects of extrinsic information only on hedonic responses (i.e., preference or liking ratings) on wine (Danner et al., 2017; Lange et al., 2002; Masson et al., 2008; Plassmann et al., 2008; Siegrist and Cousin, 2009).

#### 2. The impact of the ‘petrol’ descriptor on the perceptions and evaluations of Riesling wine

When the Riesling wine was described as having a ‘petrol’ note, which is commonly associated with the grape variety (Parker, 2020; Ross et al., 2014; Sacks et al., 2012; Schmidt, 2013; Tarasov et al., 2020; Winterhalter, 1991), the social drinkers perceived the wine as having a heavier

### TABLE 3. Results of the linear mixed modelling evaluating the fixed effect of the verbal description (the effect of the target condition against the control condition) for the ratings of the MBA wine.

<table>
<thead>
<tr>
<th>Group</th>
<th>Attribute</th>
<th>Estimate</th>
<th>95 % CI</th>
<th>SE</th>
<th>t-value</th>
<th>p-value</th>
<th>Counter-null</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social drinkers</td>
<td>Palatability</td>
<td>0.394</td>
<td>[–0.249, 0.836]</td>
<td>0.269</td>
<td>1.092</td>
<td>0.2809</td>
<td>0.587</td>
</tr>
<tr>
<td></td>
<td>Sweetness</td>
<td>–0.093</td>
<td>[–0.710, 0.523]</td>
<td>0.306</td>
<td>–0.305</td>
<td>0.7617</td>
<td>–0.187</td>
</tr>
<tr>
<td></td>
<td>Acidity</td>
<td>0.320</td>
<td>[–0.249, 0.890]</td>
<td>0.282</td>
<td>1.134</td>
<td>0.2628</td>
<td>0.641</td>
</tr>
<tr>
<td></td>
<td>Astringency</td>
<td>0.252</td>
<td>[–0.440, 0.943]</td>
<td>0.343</td>
<td>0.734</td>
<td>0.4668</td>
<td>0.504</td>
</tr>
<tr>
<td></td>
<td>Alcohol</td>
<td>–0.050</td>
<td>[–0.635, 0.534]</td>
<td>0.290</td>
<td>–0.173</td>
<td>0.8632</td>
<td>–0.101</td>
</tr>
<tr>
<td></td>
<td>Body</td>
<td>0.362</td>
<td>[–0.307, 1.031]</td>
<td>0.332</td>
<td>1.092</td>
<td>0.2811</td>
<td>0.724</td>
</tr>
<tr>
<td></td>
<td>Expensiveness</td>
<td>0.505</td>
<td>[–0.067, 1.076]</td>
<td>0.284</td>
<td>1.779</td>
<td>0.0823</td>
<td>1.009</td>
</tr>
<tr>
<td>Wine experts</td>
<td>Palatability</td>
<td>0.459</td>
<td>[–0.069, 0.986]</td>
<td>0.261</td>
<td>1.755</td>
<td>0.0864</td>
<td>0.917</td>
</tr>
<tr>
<td></td>
<td>Sweetness</td>
<td>0.261</td>
<td>[–0.324, 0.846]</td>
<td>0.290</td>
<td>0.900</td>
<td>0.3734</td>
<td>0.522</td>
</tr>
<tr>
<td></td>
<td>Acidity</td>
<td>0.043</td>
<td>[–0.496, 0.583]</td>
<td>0.267</td>
<td>0.163</td>
<td>0.8715</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>Astringency</td>
<td>–0.280</td>
<td>[–0.949, 0.390]</td>
<td>0.332</td>
<td>–0.844</td>
<td>0.4037</td>
<td>–0.560</td>
</tr>
<tr>
<td></td>
<td>Alcohol</td>
<td>–0.248</td>
<td>[–0.818, 0.322]</td>
<td>0.283</td>
<td>–0.877</td>
<td>0.3855</td>
<td>–0.496</td>
</tr>
<tr>
<td></td>
<td>Body</td>
<td>–0.069</td>
<td>[–0.725, 0.586]</td>
<td>0.325</td>
<td>–0.214</td>
<td>0.8316</td>
<td>–0.139</td>
</tr>
<tr>
<td></td>
<td>Expensiveness</td>
<td>0.023</td>
<td>[–0.536, 0.583]</td>
<td>0.277</td>
<td>0.083</td>
<td>0.9339</td>
<td>0.046</td>
</tr>
</tbody>
</table>
body and being more expensive. It was shown that consumers’ understanding of the wine body is not based on the textural concept; rather, it reflects the holistic flavour perception and overall magnitude of intensity (Niimi et al., 2017). Hence, the term ‘petrol’ might have intensified social drinkers’ flavour perception, and that perceptual change might have resulted in enhanced body perception. Additionally, when the ‘petrol’ descriptor was present, the wine was rated as being slightly more palatable and sweeter. These results suggest that the use of the descriptor ‘petrol’ for Riesling wine led to the perception of an enhanced body and sweetness, and these perceptual changes might have resulted in greater acceptance (i.e., palatability) and higher quality assessment (i.e., expensiveness). These results are also in line with previous findings that presenting proper and elaborate aroma descriptions increases consumers’ wine liking and willingness to pay for the wine (Danner et al., 2017).

Interestingly, the use of the descriptor ‘petrol’ led to a slight increase in its astringency rating among social drinkers. Unlike red wine, which is typically produced with the intended extraction of tannins, Riesling wine is generally not astringent, and the wine used in the present study did not have a noticeable astringent quality (note that both social drinkers and wine experts rated its astringency as the least strong property compared to other properties, see Table 1). However, when the ‘petrol’ descriptor was present, the social drinkers may have mistakenly attributed their trigeminal sensation caused by acidity (Viana, 2011) to astringent sensation, possibly because of their impression of chemical products. Nevertheless, this change in astringency did not seem to have an averse effect as the palatability rating did not decrease but rather enhanced in the target condition. Taken together, these findings suggest that verbally describing Riesling wine using the descriptor ‘petrol’ can positively enhance the experience of social drinkers.

By contrast, the wine experts were less influenced by the presence of the descriptor. The lower susceptibility of wine experts to verbal influence may be partly attributed to their well-established perceptual skills to taste and evaluate wines without relying on verbal memory or odour-naming skills (Croijmans et al., 2021; Parr et al., 2002). Although ratings of sensory properties remained unaffected, their quality assessment was affected to some extent; they rated the wine as less expensive when the descriptor ‘petrol’ was present. The decrease in expensiveness rating may have been partly caused by the wine experts considering the ‘petrol’ character in young Riesling wine to be a sensory defect (Robinson, 2006). However, it should be noted that the palatability of the wine was not affected by the descriptor ‘petrol,’ suggesting that, at least in this case, the verbal influence on wine experts is primarily cognitive (i.e., quality assessment) rather than sensory or hedonic.

3. The impact of the ‘strawberry candy’ descriptor on the perceptions and evaluations of MBA wine

When MBA wine was described as having a ‘strawberry candy’ note, which is considered the grape variety’s characteristic aroma (Goto-Yamamoto, 2019; Kobayashi et al., 2013; Robinson, 2016; Sasaki et al., 2015a; Sasaki et al., 2015b), the social drinkers rated the wine as more expensive. This result was not expected as we thought the term ‘candy’ would lead social drinkers to see the wine more approachable and less serious. This unexpected result can be explained by the possibility that the addition of the descriptor ‘strawberry candy’ made the social drinkers perceive the wine as more complex, having a wide range of aroma characteristics, resulting in an increase in the expensiveness rating. They also rated the wine as slightly more acidic, probably through the association of the term ‘strawberry’ with its fresh, sharp acidity. The result also showed that the presence of the descriptor slightly enhanced the palatability of the wine, indicating that verbally describing MBA wine using the descriptor ‘strawberry candy’ can positively enhance the experience of social drinkers.

Intriguingly, the descriptor ‘strawberry candy’ also slightly enhanced the astringency and body perceptions of social drinkers. This result may be explained by the expectation disconfirmation and contrast effect (Cardello and Sawyer, 1992; Piqueras-Fiszman and Spence, 2015). Following the experiment, participants were debriefed regarding the true purpose of the study, which examined the influence of aroma descriptors. One notable and frequently observed response among social drinkers was that they were unable to confidently identify the ‘strawberry candy’ note in the wine (e.g., “Although I had been informed that the red wine had a ‘strawberry candy’ note, I was unable to find that note or character in the wine.”). Conversely, many wine experts reported that once they saw the descriptor, they became confident that the wine was made from MBA grapes. While these responses were not officially recorded, this suggests that the target descriptor helped the experts form accurate expectations about the taste/flavour profile of MBA wine but not the social drinkers. If this is the case, the social drinkers may have set inaccurate expectations regarding the wine before tasting (e.g., “This candy-like wine must be sweet and light-bodied, without any astringency.”). Once they tasted the wine, their expectations were violated, as the MBA wine used was a completely dry style with no noticeable residual sugar. Disconfirmation of expectation is known to highlight and exaggerate the gap between expectation and actual perception, that is, the contrast effect (Cardello and Sawyer, 1992; Piqueras-Fiszman and Spence, 2015). Hence, in this study, a slight enhancement of astringency and body perception might have been observed in the social drinkers. To clarify this point, future studies should measure the degree of expectation disconfirmation and compare it between social drinkers and wine experts.

For the wine experts, a relatively more evident influence was found for palatability rating; the wine experts rated the wine as more palatable when the target descriptor was present. They also rated the wine as slightly sweeter and less astringent. As mentioned above, the experts often reported that they were more confident that the red wine was made from MBA grape when the wine was described as having a
strawberry candy’ note. This may have helped the experts form accurate taste/flavour expectations for the MBA wine (i.e., sweet-smelling aroma and fewer tannins) and biased their perception to confirm that style, leading to greater acceptance of the wine (Piqueras-Fiszman and Spence, 2015). Although this explanation in terms of expectations should be carefully assessed in future research, the results indicate that describing MBA wine with its distinct ‘strawberry candy’ aroma has a positive impact on both wine experts and social drinkers. Again, this is in line with previous findings that proper aroma descriptions increase acceptance of wine (Danner et al., 2017).

4. Limitations and future directions

We found that presenting the target descriptors for the Riesling and MBA wines had a positive impact on the perception and evaluation of social drinkers. This further strengthens the previous notion that clearly describing the distinct aroma characteristics of wine is an effective and desirable way to enhance the wine experience (Danner et al., 2017). However, further evidence is still needed to strengthen that conclusion. For instance, the target condition in this study always had one additional descriptor than the control condition, which leaves the possibility that the wines were perceived as more complex because they had a greater number of descriptors and, thus, were perceived as more expensive and preferable. Therefore, future studies should align the number of descriptors between conditions and investigate the effects of the presence of target descriptors. Specifically, it is beneficial to investigate the effect of adding mismatched descriptors on flavour perception and evaluation of the wine (e.g., ‘petrol’ note for MBA wine and ‘strawberry candy’ note for Riesling wine). By comparing the influence between the case where the matched descriptor is added, as in this study and the case where a mismatched descriptor is added, we can further understand the underlying psychological mechanism of the verbal influence.

This study has some limitations. First, the sample size of this study (23 participants in each group) was relatively small, even though it was determined by following previous studies employing two different participant groups, around 25 individuals in each of social drinker and wine expert groups (Croijmans et al., 2021; Parr et al., 2003). We believe that this sample size did not severely devalue our findings, but future research employing larger sample sizes would be desirable to obtain clearer results. Second, chemical analysis of wine samples was not performed in the present study. By incorporating such analysis, researchers could elucidate the interactions between the chemical composition of wine and the manipulation of verbal descriptions. Third, unlike previous studies that measured the concentration of TDN in Riesling wine and investigated its detection threshold (Ross et al., 2014; Tarasov et al., 2020), the present study did not measure the concentration of aroma compounds contributing to the target note (i.e., TDN in Riesling and furaneol in MBA wines). By measuring the concentration of the target aroma compound, the influence of the verbal description of the target note can be evaluated more effectively. More specifically, it enables us to evaluate and control the degree of conformity (or discrepancy) between the given description and the actual flavour profile of the wine. Fourth, because the influence of verbal description on wine perception may rely on expectation formation, it would have been useful to measure participants’ expectations before tasting. Measuring participants’ knowledge and understanding of grape varieties and their typical aromatic styles would also be useful. These measures would allow researchers to speculate on how and why the target descriptors influence the perception and evaluation of the way they did. All these limitations should be carefully addressed in the future to broaden our understanding of the influence of verbal descriptions on wine perception.

The results indicate that the impact of verbal description varies between social drinkers and wine experts. It is also plausible that verbal descriptions could have varying influences on people from different cultural backgrounds. In fact, Ross et al. (2014) demonstrated that the consumer rejection threshold of TDN in Riesling wine was different between consumers from New Zealand and the USA. Therefore, to broaden our understanding of the verbal influences on wine perception, our data from Japanese social drinkers and wine experts should be compared with those from different countries/regions in future cross-cultural studies.

The present study did not employ the conventional “significance” testing because researchers have cautioned that the use of P-values in the conventional, dichotomous way may cause misinterpretation of results and inaccurate conclusions (Amrhein and Greenland, 2022; Amrhein et al., 2019; Snedecor and Cochran, 1967). It is important to recognise that a statistically non-significant result does not necessarily confirm the null hypothesis (no difference between groups or no treatment effect); similarly, a statistically significant result does not necessarily support the other hypotheses (Fisher, 1935). This is particularly crucial in the medical field, where false claims of “no difference” can conceal unintended drug effects or potential health risks (Amrhein et al., 2019). Although the present study is not in the medical field, we believe this non-dichotomous approach has certain advantages for our exploratory research because it offers more nuanced interpretations regarding the impact of verbal descriptions on wine perception, thereby motivating further investigation in the future.

CONCLUSION

This exploratory study revealed that providing verbal descriptions of wine prior to tasting can influence the flavour perception and evaluation of the wine, but the effect was different for social drinkers and wine experts. Social drinkers rated the Riesling wine with the ‘petrol’ descriptor as heavier-bodied and more expensive than the same wine without the descriptor. Wine experts were not so influenced as the social drinkers, but the experts rated the Riesling wine with the ‘petrol’ descriptor as less expensive. In the case of MBA wine with the ‘strawberry candy’ descriptor, the social
drinkers rated it as more expensive, whereas the wine experts rated it as more palatable than the same wine without the descriptor. Further elaborate research is required to broaden our understanding of the mechanism of verbal influence and to establish more effective verbal tools that can improve wine communication and enhance drinkers’ wine experience.

ACKNOWLEDGEMENTS

This study was supported by the SAPPORO Bioscience Foundation.

REFERENCES


