Comparison of sensory evaluation techniques for Hungarian wines

Zsuzsanna Guld | Diána Nyitrainé Sárda | Attila Gere | Anita Rác

1Department of Oenology, Szent István University, Budapest, Hungary
2Sensory Laboratory, Faculty of Food Science, Szent István University, Budapest, Hungary
3Plasma Chemistry Research Group, Research Centre for Natural Sciences, Budapest, Hungary

Correspondence
Attila Gere, Sensory Laboratory, Faculty of Food Science, Szent István University, Villányi út 29-43, H-1118 Budapest, Hungary.
Email: gere.attila@etk.szie.hu, gereattilaphd@gmail.com

Funding information
European Social Fund, Grant/Award Numbers: EFOP-3.6.3-VEKOP-16-2017-00005, VEKOP-2.3.3-15-2017-00022; European Union; Hungarian Academy of Sciences; Szent István University; National Research, Development and Innovation Office of Hungary, Grant/Award Numbers: KH_17 125608, K 119269

Abstract
The aim of this study was to compare different Hungarian Kadarka, Kékfrankos, and Cabernet franc wines produced and aged by the same methods and to compare two types of sensory analysis methods as well: the 100-point OIV system and quantitative descriptive analysis (QDA). Both tests were conducted by 12 assessors of the University of Pécs, Institute for Regional Development, Faculty of Horticulture and Oenology. This study provides conclusions about the use of sensory analysis methods, highlighting the advantages and disadvantages of QDA and the OIV system. Principal component analysis, analysis of variance (ANOVA), multiple factor analysis, and partial least squares discriminant analysis were used for the evaluation of the data. Our results showed that the sensory panel was able to discriminate the samples by both sensory methods; however, the information provided by them was significantly different. ANOVA clearly showed that the two methods have different sensitivity when comparing wines (commercial and produced wine samples) and QDA proved to be the more sensitive, as well as more detailed, method. Partial least squares discriminant analysis augmented the findings in the classification part of the different type of wine samples. In general, OIV is able to show the general quality of the wines, while QDA coupled with proper chemometric methods is able to describe why the given samples received good or bad OIV scores.

KEYWORDS
ANOVA, PLS-DA, sensometrics, sensory analysis, wine

Abbreviations: AOTL, – Aspect aspect other than limpidity; HPER, – Harmonious harmonious persistence; PIN, – Positive positive intensity Nose; PIT, – Positive positive intensity Taste; CabH, – produced Cabernet franc; CabR, – commercial Cabernet franc; KadH, – produced Kadarka; KadR, – commercial Kadarka; KekH, – produced Kékfrankos; KekR, – commercial Kékfrankos; OIV, – a method approved by International Organisation of Vine and Wine; QDA, – quantitative descriptive analysis (not to be confused with quadratic discriminant analysis).
1 | INTRODUCTION

1.1 | Factors influencing the sensory analysis of wines

The sensory analysis of wines can be divided into three main parts: the visual perceptions, the olfactory sensations, and the taste and mouth-feel sensations.\(^1\) Out of the mentioned three, olfactory and taste and mouth-feel sensations can be connected to specific chemical components of the wines. Usually, the interaction of sugars, acids, alcohols, and phenols can make a characteristic balance in the wines among their attributes. Regarding the olfactory sensations, the most important compounds are the acids, which cause characteristic odours, such as acetic acid being vinegary or formic acid having a strong pungent odour. On the other hand, acetaldehyde or hydrogen sulfide can be considered as off-odours. In the case of taste and mouth feel-sensations, there are more chemical factors affecting sensory quality. Sweetness of the dry wines for example comes from the presence of aromatic compounds combined with ethanol and glycerol. Ethanol can contribute to the balance in the sour taste; moreover, the sensory profile of wines can be altered by high alcohol levels, due to hotness and roughness.\(^2\) Phenolic compounds such as tannins are typical attributes of red wines,\(^3\) and they induce bitter and astringent sensations. Astringency has been identified as one of the most important sensory characteristics that define the quality and complexity of red wines.\(^4,5\) Wine astringency has been described as drying, rough, harsh, woody, and green.\(^5,6\) It can be emphasized that sensory analysis is the most direct method to evaluate wine astringency.\(^5,7\) Bitter taste intensity and astringency can be decreased due to natural micro-oxygenation during the period of aging,\(^8,9\) when the pigments are also stabilized.\(^10,11\) On the other hand, wine micro-oxygenation could improve the structure and fruitfulness of red wines.\(^12\)

1.2 | Sensory evaluation methods of wines

The most accepted and respected quality evaluations of wines are traditionally done by human wine experts. However, it must be mentioned that instrumental measurements are also available to assess the overall quality of wines. The basic principle of electronic tongues and noses is measuring the potential difference between a reference and a set of working electrodes. While electronic tongues require a liquid matrix, electronic noses are used in the headspace analysis of the samples. Although these devices give precise and repeatable results, their use is usually restricted to quality control, classification, and authentication issues.\(^13\)

Within human sensory evaluation, we usually define three groups of sensory assessors: (a) consumers, (b) trained individuals, and (c) experts.\(^14\) It is important to note that these groups of assessors should be handled on a horizontal plane instead of a vertical one meaning that none of the groups is better or worse simply because they have different tasks to complete. Consumer assessors are used to rate the wine samples based on their preferences regardless of the quality of the given wine.\(^15\) A wine suffering from serious faults might be preferred over a high quality one simply because of the way consumers think. On the other hand, trained assessors need to complete a set of training tasks to prepare themselves to the actual sensory analysis. They are aware of their taste sensitivity, discrimination ability, and repeatability as well; hence, they act like a well-calibrated instrument. Trained panelists work in sensory panels consisting of 10 to 12 trained individuals in order to exclude as many disturbing factors as possible (such as illness and day-offs). Trained panels have scheduled trainings besides their usual work in order to keep their senses as sharp as possible. The third group consists of the experts, who are some of the most known wine experts of a given region or field. They train themselves regularly using different techniques to develop and maintain sharp taste and smell senses.\(^16\) Furthermore, they are highly motivated to become (and stay) recognized experts of the field. Experts are necessary in the field of wine sensory evaluation, but they cannot be used as everyday assessors. They work in national or international wine competitions, fairs, and other events, where they need to use their skills to distinguish samples with small differences. Experts might also be involved in the regular training of the wine panels.

The different groups of assessors require different sensory methods to complete their primary aims. In order to define the quality of a wine, usually trained panels are used; hence, the two most known methods will be discussed in detail. Sensory standards suggest quantitative descriptive analysis (QDA) to map the overall quality of the samples.\(^17\) QDA is a descriptive sensory method, during which the assessors define a set of sensory attributes describing the samples. This list contains attributes related to appearance, odour, taste, and mouthfeel, regardless of being positive or negative (eg, wine faults). The typical QDA output is the so-called spiderweb plot, which presents the average values of
each sample. Results of QDA are suitable inputs for accepted statistical methods in food research, such as analysis of variance (ANOVA) or principal component analysis (PCA).18

On the other hand, QDA is a time demanding and hardly customizable method for quality ratings (such as competitions); hence, the International Organization of Vine and Wine (OIV) 100-point method is the most widely applied sensory technique to rate wines.19 The OIV method uses four predefined sensory categories, which are applicable to all types of wines and thus are able to differentiate high and low quality. Although both methods are well known and widely applied, there is a lack of scientific literature dealing with the scientific comparison of QDA and OIV and to describe the advantages/disadvantages and similarities/differences between the two methods. Thus, the aim of the present study was to develop a QDA reference system for the analysis of six wines (from three different varieties) and conduct a multiway comparison between the produced and the commercial reference wines based on the two types of sensory analysis, with robust and specific chemometric tools.

## 2 | MATERIALS AND METHODS

### 2.1 | Wine varieties

Three characteristic wine varieties (Cabernet franc, Kékfrankos, and Kadarka) of the Szekszárd wine region in Southwest Hungary were produced. Cabernet franc, Kékfrankos, and Kadarka are red varieties of *Vitis vinifera*, which were selected based on their potentials to produce high-quality wines with oak barrel aging. These wine varieties are the most prominent, top-seller, and well-known ones in the Szekszárd wine region (Hungary). Moreover, they have sufficiently different characters in the sensory features to make a broad scope analysis based on their sensory attributes. Cabernet franc is a major grape variety even worldwide, Kékfrankos has higher polyphenol concentration (compared with other grape varieties), and Kadarka is also a major component of Hungarian wine production.20,21 The wines have not received any treatment other than racking and basic sulphuring up to 20 to 25 mg/L free sulfur concentration.

Harvest dates were different for the three varieties: early October for Cabernet franc, mid-October for Kadarka, and late October for Kékfrankos. For each variety, healthy raw material was processed. Alcoholic fermentation was carried out under controlled conditions for 21 days, on 18 to 24°C temperature, using UVAFERM BDX (Danstar Ferment AG, Zug, Switzerland) variety-specific yeast in the case of Kékfrankos. For Cabernet franc and Kadarka, alcoholic fermentation was carried out under controlled conditions for 14 days, on 25 to 28°C temperature, using UVAFERM BDX (Danstar Ferment AG, Zug, Switzerland) variety-specific yeast. Following fermentation, the different varieties (Kadarka, Kékfrankos and Cabernet franc) were stored in separate oak and barrique barrels for 24-month aging.

The three types of produced wines were compared with commercially available wines from the same varieties according to the sensory attributes. The commercially available wines met the following criteria: (a) high similarity to the produced samples, (b) easy availability, (c) high production amounts, and (d) reliable producer.

### 2.2 | Barrel types

For all three varieties, 500-L oak barrels and 225-L barrique barrels, produced from sissile oak (*Quercus petraea*), were applied for the 24-month aging process. We have used a first-load oak barrel for Kadarka, and 5-year, fourth-load barrels for Kékfrankos and Cabernet franc. The barrique barrels were first-load for Kadarka and third-load for Kékfrankos and Cabernet franc.

### 2.3 | Sensory evaluation

Sensory evaluations were conducted at University of Pécs, Faculty of Cultural Sciences, Education and Regional Development, Institute for Regional Development. Twelve members of the trained panel were sophomores (aged between 21 and 22, seven males, five females) of the University of Pécs, Faculty of Cultural Sciences, Education and Regional Development, Institute for Regional Development. These individuals received a training involving at least 120 hours of sensory analysis, which covered methodological, technical, and throughout tasting sessions. Tests were conducted using two replicates on two consecutive weeks to ensure data reliability. Assessors used two sensory methods: (a) the 100-point OIV system,19 which is a widely accepted sensory method in Hungary to determine the quality of wines, and (b) QDA,17,22,23 which is an internationally applied standard method to describe food samples. The testing room was
clean, light, odourless, and temperature-controlled. Samples (0.2 L/sample) were presented in standard, clear wine glasses at a usual consumption temperature.

2.4 | Quantitative descriptive analysis

In total, 27 different sensory attributes were applied in QDA, namely: colour intensity, turbidity, glycerin, global odour intensity, redberry odour, brandied cherry odour, vanilla odour, dark chocolate odour, woody odour, spicy odour, green odour, acidity, astringency, global taste intensity, redberry flavour, ethanol flavour, vanilla flavour, dark chocolate flavour, woody flavour, spicy flavour, green flavour, bitter taste, sour taste, sweet taste, raw aftertaste, sour aftertaste, and off flavour intensity. Attributes were rated on a 100-point scale (Table 1). Sensory attributes were defined during the first session of the QDA. During QDA, reference materials were used to anchor the scales and to provide further aid to the assessors. The use of reference materials reduces the deviation between panelists; hence, more reliable data were obtained. The reference materials for the attributes are listed in Table 1.

2.5 | 100-point OIV sensory analysis

The second type of sensory assessment was done according to the OIV 332/A/2009 resolution, where assessors rated the samples based on four categories according to the following sensory attributes: visual category: limpidity (clarity), aspect other than limpidity; nose category: genuineness, positive intensity, quality; taste category: genuineness, positive intensity, harmonious persistence, quality; and harmony category: overall judgement. Scores of each sensory attribute were recorded, and an overall score was calculated as the sum of the individual attribute scores.

2.6 | Panel performance

Performance of the trained panel was evaluated using the PanelCheck ver. 1.4 software. PanelCheck provides all the necessary evaluations needed to assess the repeatability, discrimination ability of the individual assessors, and panel agreement.

One-way ANOVA is conducted on the data with assessors as factors. The resulting F-values and mean-square-error (MSE) values are used to assess the performance of the individual sensory panel members. Assessors having lower MSE values can reproduce their results better.

2.7 | Analysis of variance

ANOVA and factorial ANOVA methods are commonly used techniques for the comparison of the classes of samples based on different features (in our case, ie, the values of the sensory attributes). The ANOVA method with the use of post hoc tests—computed after the primer analysis—makes a pairwise comparison of the average values of the different groups of samples. In our case, the (a) sensory attributes (10 or 27 levels), (b) the type of the wine samples (two levels, commercially available or produced), and the (c) reproducibility (two levels) were used as factors (categorical variables) for the analysis. STATISTICA 13 (TIBCO Software Inc., Palo Alto, CA, USA) software was used for the ANOVA analyses.

2.8 | Principal component analysis

PCA is one of the most important methods in chemometrics for pattern recognition. It is an unsupervised technique, which means that we do not use the class memberships (categorical variables) in the analysis. PCA is a dimension-reducing method; ie, principal components are computed from the linear combination of the original variables with two constraints: orthogonality and normalization. PCA was used in the statistical evaluation of the dataset based on the sensory profiles of the wines, provided by the panelist.
| Attributes          | Lower Endpoint (0) | Higher Endpoint (100) | References                                                                 |
|---------------------|--------------------|-----------------------|-----------------------------------------------------------------------------|
| **Visual attributes** |                    |                       |                                                                             |
| Colour intensity    | light              | dark                  | Verbal description                                                          |
| Turbidity           | clear              | turbid                | Verbal description                                                          |
| Glycerin            | thin               | thick                 | Verbal description                                                          |
| **Odour attributes** |                    |                       |                                                                             |
| Global odour intensity | weak            | intense               | Verbal description                                                          |
| Redberry odour      | weak               | intense               | dried red berries                                                          |
| Brandied cherry odour | weak            | intense               | Cherry Queen dessert (Bonbonetti Choco Ltd. - Budapest, Hungary)            |
| Vanilla odour       | weak               | intense               | Bourbon vanillin sugar (Dr Oetker-Jánossomorja, Hungary)                    |
| Dark chocolate odour | weak             | intense               | Chocolate without sugar (Wawel, 70% cocoa content - Poland)                 |
| Woody odour         | weak               | intense               | Oak chips 5 g                                                               |
| Spicy odour         | weak               | intense               | 0.1 g milled black peppercorns                                              |
| Green odour         | weak               | intense               | Green paprika                                                              |
| **Mouthfeel**       |                    |                       |                                                                             |
| Acidity             | none               | strong                | Verbal description                                                          |
| Astringency         | weak               | intense               | Black tea (10 filters in 1 L of water for 20 minutes) (Lipton Tea Earl Grey Classic - Unilever, Hungary) |
| **Taste and flavour attributes** |                |                       |                                                                             |
| Global taste intensity | weak            | intense               | Verbal description                                                          |
| Redberry flavour    | weak               | intense               | Frozen red berries                                                          |
| Ethanol flavour     | weak               | intense               | Aqueous solution containing 30% ethanol (food grade)                        |
| Vanilla flavour     | weak               | intense               | Oak chips with vanilla toast (5 g) soaked in water                           |
| Dark chocolate flavour | weak           | intense               | Chocolate without sugar (Wawel 70% cocoa content - Poland)                 |
| Woody flavour       | weak               | intense               | Oak chips (5 g) soaked in water                                             |
| Spicy flavour       | weak               | intense               | 0.1 g black peppercorns                                                     |
| Green flavour       | weak               | intense               | Verbal description                                                          |
| Bitter taste        | weak               | intense               | Aqueous solution containing 0.06% caffeine                                  |
| Sour taste          | weak               | intense               | Aqueous solution containing 0.07% citric acid                              |
| Sweet taste         | weak               | intense               | Aqueous solution containing 1.6% sugar cane                                 |
| Raw aftertaste      | weak               | intense               | Verbal description                                                          |
| Sour aftertaste     | weak               | intense               | Verbal description                                                          |
| Off flavour intensity | weak             | intense               | Pickled cucumber                                                            |
In our case, the method was applied to the QDA and the 100-point OIV sensory data as well. Score and factor loading vectors were used together as a biplot to show the patterns of the different attributes and panelists. PCA was run with the XL-Stat software (Addinsoft, Long Island, NY, USA).

2.9 | Multiple factor analysis

Multiple factor analysis (MFA) is used to study data sets in which the elements are described by a set of variables, structured in groups. The number of variables may differ in each group, but the nature of the variables should be the same within one group. MFA may be considered as a general factor analysis, and in our case, it is based on PCA because our variables (both QDA and OIV) are quantitative. MFA first conducts two separate PCAs on the given two data sets and stores the value of the first two eigenvalues of both PCAs. The stored eigenvalues then are used in the second step as weights, when a weighted PCA is computed on all columns of the input data table (eg, on both data sets at the same time). As a result, MFA is able to visualize not only the scores and loadings but also the impact of data tables. MFA was run using the XL-Stat software (Addinsoft, Long Island, NY, USA).

2.10 | Partial least squares discriminant analysis

Partial least squares discriminant analysis (PLS-DA) is a supervised pattern recognition method, which is the classification form of PLS regression. Here, class memberships are defined in the Y variable. Thus, the Y vector is not continuous like in PLS regression. The basic idea of PLS-DA is closely connected to PCA: the X and Y matrices are decomposed to score (T, U) and loading vectors (and error matrices) with the maximization of covariance. The nonlinear iterative partial least squares algorithm is a frequently used technique for the estimation of modelling parameters. The number of latent variables (PLS components) determines the complexity of the models, and a proper validation protocol (such as cross-validation) is needed for the preparation of robust and reliable models.

PLS-DA was used to compare the produced and commercial wines of each variety based on their 100-point OIV sensory data and QDA, separately. The most important variables are visualized with the variable importance projection (VIP) plot, where the higher the value, the better the variable. Moreover, the performance of the classifications was presented on receiver operating characteristic (ROC) curves, where sensitivity is plotted against 1-specificity at varying score thresholds. AUC value (area under the ROC curve) was used as performance parameter for the comparison of the models. PLS-DA was run using the XL-Stat software (Addinsoft, Long Island, NY, USA).

3 | RESULTS AND DISCUSSION

3.1 | Panel performance analysis

Performance of the sensory panel was monitored by calculating panel performance metrics using PanelCheck software in order to get a closer picture of the performance of the individuals and the panel, too. F-plots showed acceptable discrimination ability since the majority of the assessors provided F-values exceeding the 5% significance level. Generally, assessors achieved higher F-values in the case of appearance attributes. In the case of taste attributes, F-values proved to be acceptable.

3.2 | Comparison of the Kadarka wine samples

Chemometric analysis of the Kadarka samples was carried out on the dataset of the 100-point OIV sensory analysis, and that of QDA as well. In the first case, the data matrix contained 10 sensory attributes, which were stacked into one variable. In this way, the sensory attributes could be used as a factor (categorical variable) in the ANOVA evaluation. Twelve panelists were used, and the evaluation was carried out in two replicates; thus, in total, 24 rows were assigned to each attribute. In the second case, the same fusion process was carried out based on the original 27 attributes. Thus, in both cases, the matrix contained one dependent variable, which contained the sensory attribute values and three factors, namely, (a) samples (two groups—KadH/KadR), (b) attributes (10 or 27 attributes), and (c) reproducibility...
The result of the factorial ANOVA analysis on the 100-point OIV sensory dataset showed that there was no clear difference between the produced and the commercial Kadarka samples, based on the single attribute values. While the produced sample received higher average ratings for nose quality and taste quality, the difference was not statistically significant ($\alpha=0.05$). On the other hand, in the case of quantitative descriptive analysis (QDA) with 27 attributes, the samples were significantly different. In this case, the produced Kadarka sample was clearly different from the commercial wine. Thus, QDA was more sensitive in this case for the comparison.

The sensory attributes were also checked together with the samples as factors for both datasets. The result showed that in the case of 100-point OIV sensory test, the combination of these two factors was not significant. In the single attribute analysis (one-way ANOVA), only one attribute was significantly different amongst the samples: the limpidity. For the QDA evaluation, the combination of the two mentioned factors was clearly significant. The one-way ANOVA showed nine attributes out of 27 were significant alone, as well: colour intensity, turbidity, redberry odour, woody odour, spicy odour, astringency, ethanol flavour, dark chocolate flavour, and woody flavour. The results of the evaluation can be seen in Figures 1A and 1B. The difference can be explained with the use of the first-load oak barrel (produced from two varieties of Hungarian sissile oak) in the case of the produced Kadarka, which is connected to the high intensity of redberrys, the woody and spicy odours, and the dark chocolate and woody flavours.

### 3.3 Comparison of the Cabernet franc samples

The Cabernet franc samples were evaluated in the same way as the Kadarka samples. The two types of the datasets were analysed separately here as well. This wine variety showed a clear difference between the produced and commercial samples in both cases (100-point OIV sensory test and QDA) during factorial ANOVA. The difference was

![Figure 1](image_url)
statistically significant at the $\alpha=0.05$ level for the wine samples (CabH and CabR); the produced Cabernet franc received significantly higher average intensity values in both tests. The combination of the type of the wine samples and the sensory attributes as factors was also statistically significant in the 100-point OIV sensory test and in QDA, as well. The result of the ANOVA evaluation can be seen in Figures 2A and 2B.

In the evaluation of the sensory attributes one by one with one-way ANOVA method, the results showed that many attributes can be connected to the difference between the two Cabernet franc samples. In the 100-point OIV test, only limpidity had no significant effect between the two wines, and in QDA, there were 13 sensory attributes out of 27 with significant differences (redberry odour, brandied cherry odour, vanilla odour, dark chocolate odour, woody odour, astringency, global taste intensity, redberry flavour, ethanol flavour, vanilla flavour, dark chocolate flavour, woody flavour, and off flavour intensity). Based on QDA, the visual attributes were not very different between the two wines. On the other hand, the biggest differences were definitely in the odour attributes. This difference can be explained with the filtration and aging of the wines: unfiltered wines with oak aging (such as the produced Cabernet franc sample) have more intense odour than the filtered and bottled ones.

3.4 | Comparison of the Kékfrankos samples

For the Kékfrankos samples, the evaluation process was the same as in the previous two cases. The produced and the commercial wine samples were significantly different in the 100-point OIV test and QDA based on the ANOVA analysis with the factor of samples ($\alpha=0.05$). However, in this case, the produced sample was worse than the commercial one. While the 100-point OIV test showed the smaller—but still significant—difference, in the case of the more sensitive QDA, it was higher difference. This part of the evaluation can be seen in Figures 3A and 3B.
Combination of the attributes and samples factors resulted significant differences in both datasets. The sensory attributes were evaluated one by one as well with one-way ANOVA. The results can be seen in Figures 4A and 4B for both wine samples and dataset.

Only one attribute was significant in the case of the 100-point OIV sensory test, and three sensory attributes were significant in the case of QDA ($\alpha=0.05$). In QDA (Figure 4B), it can clearly be seen that the produced samples obtained lower intensity values in some odour attributes (e.g., global odour intensity and spicy odour or green odour) than the commercial sample.

### 3.5 Comparison of produced and commercial samples

The 100-point OIV sensory test and the QDA analysis were also compared for each variety with PLS-DA analysis. The classification was based on the “type” of the wines, thus the produced and commercial samples. Cross validation with...
jack-knifing algorithm was used, and the attributes were compared with VIP algorithm. The final validated classification models for each variety can be seen in Figure 5 for the 100-point OIV test and in Figure 6 for the QDA analysis. Every model has a ROC curve plot, which shows the performance of the models with the AUC values.

**Figure 5** The results of partial least squares discriminant analysis for each wine varieties in the case of 100-point OIV sensory dataset. A – D: Kadarka samples, B – E: Kékfrankos samples, C – F: Cabernet franc samples. D, E, F subplots are assigned to the variable importance (VIP scores).

**Figure 6** Results of partial least squares discriminant analysis for each wine varieties in the case of quantitative descriptive analysis dataset. A-D: Kadarka samples; B-E: Kékfrankos samples; C-F: Cabernet franc samples. D, E, and F subplots are assigned to the variable importance (variable importance projection scores).
As it can be seen from the ROC curves presented by Figures 5 and 6, QDA method resulted better AUC values in two cases, but in the case of Cabernet samples, the AUC values were highly similar, almost identical. This result was supported by the ANOVA presented in the paper. In the case of Cabernet samples, the OIV method gave the highest number of significant attributes among all OIV evaluations. Moreover, in the variable importance plots (especially in the case of OIV evaluation), Cabernet samples have the most amount of attributes with high VIP scores. The top ranked attributes were almost identical to the results of ANOVA analysis. PLS-DA of QDA results gave similar AUC values, all over 0.8, meaning that the method provides much better input data for the differentiation of the samples in similar varieties of wines (e.g., the differences are described better using QDA) compared with OIV.

### 3.6 Reproducibility tests

All in all, two replicates of the sensory tests were executed on the six different wine samples. Thus, the reproducibility of the results was also checked with the ANOVA evaluations. Based on the results, the two replicates had no significant effect on the two different tests (100-point OIV sensory test and QDA) (α=0.05). Here, the six wine samples were used in the same dataset for the ANOVA; thus, it could give us a more illustrative figure about the wines and replicates. Figures 7A and 7B show the results of the factorial ANOVA evaluation with the use of reproducibility and samples as factors. The two replicates were in good agreement in each case.

### 3.7 Comparison of the two sensory tests

The ANOVA results are summarized for the three wine varieties and for the two types of sensory tests in Table 2 to make an informative conclusion about the use of the different sensory tests.

Comparison of the two sensory methodologies was also carried out by means of PCA. During PCA, the obtained sensory data were centred and scaled; hence, the effect of different scales in the case of OIV was eliminated. Figure 8A presents the PCA biplot of the OIV sensory evaluation, while Figure 8B presents the one of the QDA. PCA biplots are used for visualization purposes; hence, this way loadings (sensory attributes) and scores (samples) are plotted at the same time. Samples close to each other are considered similar, while the closer an attribute to a sample is, the more intense the attribute in the sample is. OIV analysis shows less discrimination among the samples; only CabR is located

---

**FIGURE 7** A and B, Results of factorial analysis of variance with the use of reproducibility and samples as factors

---

**TABLE 2** Detailed summary of the findings based on the wine varieties and sensory tests

|                      | 100-Point OIV Test | QDA          |
|----------------------|--------------------|--------------|
|                      | Kadarka            | Cabernet franc | Kékfrankos | Kadarka | Cabernet franc | Kékfrankos |
| Samples              | No                 | Yes           | Yes        | Yes     | Yes           | Yes         |
| Sensory attributes*samples | No       | Yes           | No         | Yes     | Yes           | Yes         |
| No. of significant sensory attributes | 1        | 9             | 1          | 9       | 13            | 3           |

Abbreviation: QDA: quantitative descriptive analysis.
farther from the others. All other samples are close to each other, meaning higher similarity based on the attributes evaluated. Regarding the attributes, limpidity and odour intensity proved to be the most different; all the other attributes explained the same information, which is shown by the loading vectors pointing to the same direction with the same extent (length of the vector) close to each other. Figure 8A shows that the produced samples were rated as more intense odour, longer harmonious persistence, and better taste quality. The only exception is KekH, which received lower scores.

On the other hand, PCA based on the QDA data shows higher discrimination among the samples as the six samples are placed far from each other and have higher scores values. The differentiation of CabR is well explained; assessors gave high (intense) scores to sour aftertaste, green odour and flavour, off flavour intensity, and raw aftertaste. Kadarka samples proved to be the fruitiest with intense red berry odour and flavour and spicy flavour. The produced CabH and KadH samples were rated as woodier, astringent with intense vanilla odour and flavour describing the aromas coming from barrique ageing.

MFA provides an excellent way to evaluate differently structured data sets (with different dimensions). It conducts a separate PCA on each data set and is able to present their amalgamated results on the same loading and scores plots. Figure 9A presents the loading plot of the two sensory data sets. The OIV attributes correlated well with the red berry odour and flavour and spicy flavour attributes, while the odour intensity attribute of the OIV test correlated well with the odour attributes describing the barrique aging. The loadings also present that fault attributes from QDA (off flavour, sour flavour, etc.) show negative correlation with the OIV attributes suggesting that the OIV system tends to rate positive aspects rather than the negative ones.

On the MFA scores plot (Figure 9B), each sample is described by three points, one corresponding to the OIV and one to the QDA evaluation, while the MFA result is located between the two; hence, MFA is able to visualize the average position of the samples taking into account both PCA results. Figure 9B clearly shows the superiority of QDA in terms of sample discrimination. For each sample, QDA results are farther from each other. The only exception is CabR, which is better discriminated by the OIV system than by QDA (still, it is well discriminated by QDA, too). Comparing Figures 9A and 9B, the main reason of the differentiation of the CabR sample can be explained easily. The OIV attributes (loadings) show a completely different direction, meaning that the sample obtained low scores. Since OIV penalizes samples (a perfect sample obtains the maximum score), the assessors found multiple faults in the wine. On the other hand, QDA is able to identify the reason of these faults, namely, the intense off flavour, sour aftertaste, and sour taste. Kékfrankos samples proved to be similar by both methods, while Kadarka samples were better discriminated by QDA. Figure 9A shows that QDA offers a more complex evaluation and several well and noncorrelated attributes were used. OIV, on the contrary, focuses on the positive attributes, and its vectors
are close to each other, showing less discriminating power. It can also be seen that raspberry odour and flavour, as well as spicy flavour from QDA, correlated highly with the odour and taste attributes of OIV, meaning that assessors focused on these attributes during the OIV evaluation.

4 | DISCUSSION

Although sensory evaluation is conducted by human panelists, it is critically important to check the performance of the panel members in order to provide “calibrated” results, similar with instrumental measurements, such as colorimetric measures. However, it should be considered that human panelists will always produce higher deviations compared with instrumental measurements. In order to handle the deviation (or disagreement), sensory practices should be conducted, where the weak points of the panelists can be identified and proper, personalized tasks are created for each assessor to improve the sensory abilities. Our results show that a wine panel consisting of young individuals, whose taste, and smell abilities are not declined35 and who have spent hundreds of hours by practicing may provide deviating data. Moreover, the results are highlighting the importance of detailed, statistically supported panel performance evaluations during wine tasting sessions.

Regarding the other objective of our study (sensory method comparison), PCA and PLS discriminant analysis of the two methods showed that OIV has less discriminative power compared with the QDA method. PCA loadings of the OIV analysis show that OIV attributes have high correlation with overall judgement. It should be noted that including overall judgement in sensory tests might influence all the other attributes since after a higher hedonic rating one might underestimate weaknesses and overestimate strengths. QDA is able to provide a more detailed description of the samples, listing positive and negative attributes at the same time. However, this advantage makes the evaluation process longer and requires more time from the panel leader. In order to reduce standard deviation, reference samples or materials are used. Assessors compare each sample to the reference stimulus, which makes the evaluation process longer. The higher number of sensory attributes also increases the evaluation time; hence, choosing the proper number of attributes is also critical to avoid fatigue. The last and the most important characteristic of QDA is the set of attributes. While OIV is a general method and can be applied to several types of wines, QDA is a highly specialized method, which lists attributes present in any of the analysed samples in the sample set. This means that a different list of descriptors may be needed to evaluate red, white, or rosé wines, which renders their comparison difficult or even impossible. However, a wine-specialized set of descriptors and reference materials would solve this issue and would enable practitioners to use QDA as an everyday tool.
In general, we can conclude that professional wine tasters prefer to taste red wines with fine astringency, wooden substances like vanilla and dark chocolate, and last but not least, fruity intensity. The transfer of substances can be allowed from the wood to the wine during at least one year of aging. Our 2-year aged wines contained fine tannins, fruity and spicy nature, as we presented through our results.

5 | CONCLUSIONS

In this study, two types of sensory analyses (100-point OIV and QDA) were used and compared in the evaluation of three red wine varieties. References for the QDA analysis were also determined in the case of 27 sensory attributes. The results showed that the QDA method is more sensitive to differences between the samples; thus, it has a bigger discriminatory power. Factorial ANOVA, PCA, and PLS-DA have highlighted this difference unequivocally for each wine. In the future, QDA analysis with a wine-specialized set of descriptors and reference materials could be an appropriate and valuable replacement of the general OIV method. It has to be mentioned that there are several wine-related sets of descriptors with reference materials, but none of them has been defined as “standard.” In general, OIV is able to show only the general quality of the wines, while QDA coupled with proper chemometric methods give a detailed and statistically supported description of the wines.

ACKNOWLEDGEMENT

The authors thank Dávid Bajusz for proofreading the manuscript. The authors are also indebted to Pál Mészáros, Ernő Módos, and Péter Vida for providing the wine samples. The work of A. R. was supported by the National Research, Development and Innovation Office of Hungary under grant numbers K 119269 and KH_17 125608. The work of Zs.G. was supported by Szent István University, Doctoral School of Food Science. Attila Gere thanks the support of the Premium Postdoctoral Program of the Hungarian Academy of Sciences and VEKOP-2.3.3-15-2017-00022. The project is supported by the European Union and co-funded by the European Social Fund (grant agreement no. EFOP-3.6.3-VEKOP-16-2017-00005).

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

ORCID

Zsuzsanna Guld https://orcid.org/0000-0001-7530-7806
Attila Gere https://orcid.org/0000-0003-3075-1561
Anita Rácz https://orcid.org/0000-0001-8271-9841

REFERENCES

1. Jackson RS. Wine Tasting: A Professional Handbook. London, United Kingdom: Elsevier Ltd; 2002.
2. King ES, Dunn RL, Heymann H. The influence of alcohol on the sensory perception of red wines. Food Qual Prefer. 2013;28(1):235-243. https://doi.org/10.1016/j.foodqual.2012.08.013
3. Lesschaeve I, Noble AC. Polyphenols: factors influencing their sensory properties and their effects on food and beverage preferences. Am J Clin Nutr. 2005;81(1):330S-335S. https://doi.org/10.1093/ajcn/81.1.330S
4. Peynaud E. The Taste of Wine: The Art and Science of Wine Appreciation. London: Macdonald Orbis; 1987.
5. Cheynier V, Sarni-Manchado P. Wine taste and mouthfeel. In: Reynolds A, ed. Managing Wine Quality. Cambridge: Woodhead Publishing Limited; 2010:29-72.
6. Lawless HT, Corrigan CJ, Lee CB. Interactions of astringent substances. Chem Senses. 1994;19(2):141-154.
7. Ma W, Guo A, Zhang Y, Wang H, Liu Y, Li H. A review on astringency and bitterness perception of tannins in wine. Trends Food Sci Technol. 2014;40(1):6-19. https://doi.org/10.1016/j.tifs.2014.08.001
8. Cejudo-Bastante MJ, Castro-Vázquez L, Hermosín-Gutiérrez I, Pérez-Coello MS. Combined effects of prefermentative skin maceration and oxygen addition of must on color-related phenolics, volatile composition, and sensory characteristics of Airén white wine. J Agric Food Chem. 2011;59(22):12171-12182. https://doi.org/10.1021/jf202679y
9. Gambuti A, Rinaldi A, Ugliano M, Moio L. Evolution of phenolic compounds and astringency during aging of red wine: effect of oxygen exposure before and after bottling. J Agric Food Chem. 2013;61(8):1618-1627. https://doi.org/10.1021/jf302822b
10. Atanasova V, Fulcrand H, Cheynier V, Moutounet M. Effect of oxygenation on polyphenol changes occurring in the course of wine-making. Anal Chim Acta. 2002;458(1):15-27. https://doi.org/10.1016/S0003-2670(01)01617-8
11. Cano-López M, Pardo-Minguez F, Schmauch G, et al. Effect of micro-oxygenation on color and anthocyanin-related compounds of wines with different phenolic contents. *J Agric Food Chem*. 2008;56(14):5932-5941. https://doi.org/10.1021/jf8006147

12. Schmidtke LA, Clark AC, Scollary GR. Micro-oxygenation of red wine: techniques, applications, and outcomes. *Crit Rev Food Sci Nutr*. 2011;51(2):115-131. https://doi.org/10.1080/10408390903434548

13. Soós J, Várvolgyi E, Dénes LD, Kovács Z, Felföldi J, Magyar I. Application of electronic tongue to discriminate white wines originated from different regions of Hungary. *Acta Aliment*. 2014;43:132-139. https://doi.org/10.1556/AAlim.43.2014.Suppl.19

14. ISO 6658:2005, Sensory analysis—methodology—general guidance.

15. Lawless HT, Heimann H. *Sensory Evaluation of Food*. 2nd ed. New York, NY: Springer Verlag; 2010. https://doi.org/10.1007/978-1-4419-6488-5.

16. Savela-Huovinen U, Muukkonen H, Toom A. Sensory expert assessor’s learning practices at workplace: competencies and contexts in sensory evaluation. *J Sens Stud*. 2018;33(2):1-10. https://doi.org/10.1111/joss.12315

17. ISO 13299. Sensory analysis—methodology, General guidance for establishing a sensory profile. 2003.

18. Bernhardt B, Sipos L, Kókai Z, et al. Comparison of different *Ocimum basilicum* L. gene bank accessions analyzed by GC-MS and sensory profile. *Ind Crop Prod*. 2015;67:498-508. https://doi.org/10.1016/j.indcrop.2015.01.013

19. OIV 332A. OIV Standard for International Wine and Spirituous Beverages of Vitivinicultural Origin Competitions. 2009.

20. Nikfardjam MSP, Márk L, Avar P, Figler M, Ohmacht R. Polyphenols, anthocyanins, and trans-resveratrol in red wines from the Hungarian Villány region. *Food Chem*. 2006;98(3):453-462. https://doi.org/10.1016/j.foodchem.2005.06.014

21. Guld ZS, Rácz A, Timá H, Kállay M, Nyitrainé SD. Effects of aging in oak barrels on the trans-resveratrol and anthocyanin concentration of red wines from Hungary. *Acta Aliment*. 2019;48(3):349-357. https://doi.org/10.1556/066.2019.0004

22. ISO 11035. Sensory analysis—identification and selection of descriptors for establishing a sensory profile by a multidimensional approach. 1994.

23. Stone H, Sidel JL. *Sensory Evaluation Practices*. 2nd ed. London: Elsevier; 2004. https://doi.org/10.1016/B978-0-12-374181-3.00017-X.

24. ISO 8589. Sensory analysis—General guidance for the design of test rooms. 2007.

25. Jackson RS. *Wines: Wine Tasting A Professional Handbook*. Elsevier A. London, UK; 2002. https://doi.org/10.1016/B978-0-12-384947-2.00757-1.

26. Sherman E, Harbertson JF, Greenwood DR, Villas-Bôas SG, Fiehn O, Heymann H. Reference samples guide variable selection for correlation of wine sensory and volatile profiling data. *Food Chem*. 2018;267:344-354. https://doi.org/10.1016/j.foodchem.2017.10.073

27. Næs T, Brockhoff PB, Tomic O. *Statistics for Sensory and Consumer Science*. 1st Ed. Chichester, West Sussex, United Kingdom: John Wiley & Sons, Ltd; 2010. https://doi.org/10.1002/9780470669181.

28. Rácz A, Andrić F, Bajusz D, Héberger K. Binary similarity measures for fingerprint analysis of qualitative metabolomic profiles. *Metabolomics*. 2018;14(3):1-9. https://doi.org/10.1007/s11306-018-1327-y

29. Wold S, Esbensen K, Geladi P. Principal component analysis. *Chemom Intel Lab Syst*. 1987;2(1-3):37-52.

30. Rácz A, Héberger K, Rajkó R, Elek J. Classification of Hungarian medieval silver coins using X-ray fluorescent spectroscopy and multivariate data analysis. *Herit Sci*. 2013;1(1):1-9. https://doi.org/10.1186/2050-7445-1-2

31. Addinsoft. XLSTAT statistical and data analysis solution. 2019.

32. Abdi H, Williams LJ. Principal component analysis. *Wiley Interdiscip Rev Comput Stat*. 2010;2(4):433-459. https://doi.org/10.1002/wics.101

33. Pages J. *Multiple Factor Analysis by Example Using R*. 1st ed. New York, NY: Chapman and Hall/CRC; 2005. https://doi.org/10.1201/b17700.

34. Barker M, Rayens W. Partial least squares for discrimination. *J Chemometr*. 2003;17(3):166-173. https://doi.org/10.1002/cem.785

35. Doty RL. Age-Related Deficits in Taste and Smell. *Otolaryngol Clin North Am*. 2018;51(4):815-825. https://doi.org/10.1016/j.otc.2018.03.014

How to cite this article: Guld Z, Nyitrainé Sárdy D, Gere A, Rácz A. Comparison of sensory evaluation techniques for Hungarian wines. *Journal of Chemometrics*. 2020;34:e3219. https://doi.org/10.1002/cem.3219