

Article

EVALUATING UNCERTAINTY IN SENSORY ANALYSIS. A CASE STUDY OF THE PANEL OF TASTERS OF THE DÃO REGIONAL WINE COMMISSION

AVALIAÇÃO DE INCERTEZA EM ANÁLISE SENSORIAL. CASO DE ESTUDO DO PAINEL DE PROVADORES DA COMISSÃO VITIVINÍCOLA REGIONAL DO DÃO

Manuel Maria Pinto

Dão Regional Wine Commission. Solar do Vinho do Dão, Rua Aristides Sousa Mendes, Apartado 10, 3501-908 Viseu. Portugal.

Tel: +351 935 576 384, email: mmpinto@gmail.com

(Received 02.12.2020. Accepted 19.03.2021)

SUMMARY

Metrological reliability in the sensory evaluation of products is of fundamental importance in ensuring the comparability of results. The issue of uncertainties regarding measured data in such contexts has led to new scientific approaches in the field of sensory analysis, which have become fundamental to studies that seek to compare results of sensory measurement and analysis, including studies regarding wine. The new normative framework, ISO/IEC 17025:2017, which replaces the previous version ISO/IEC 17025, published in 2005, draws attention to the concept of measurement uncertainty, stating that it should be considered by laboratories when assessing conformity and quantifying risks arising in decision-making processes. The present work examines the evaluation of uncertainty and its importance for the interpretation of results from sensory analyses, examining sensory analysis tests related to the certification of the Designation of Origin "Dão", the Designation of Origin "Lafões" and the Geographical Indication "Terras do Dão", carried out by a panel composed of 20 tasters of the Dão Regional Wine Commission, trained and accredited by the standard NP ISO/IEC 17025:2018. The results achieved led to a conclusion that the panel presented different values of uncertainty regarding the tests to determine sensory attributes. The results and respective values of uncertainty made it possible to assess the performance of the panel of tasters and to establish a level of risk (false accept, false reject and statistical assumptions) for decision rules used when issuing statements of conformity. In this study, a new methodology was presented to determine an estimate of uncertainty to be applied in the interpretation of results related to sensory analysis tests of wines for the certification of Designation of Controlled Origin "Dão", Designation of Origin "Lafões" and Geographical Indication "Terras do Dão". This progress led to an evolution in the day-to-day work of the Dão Regional Wine Commission tasting chamber enabling greater solidity in the opinions issued, in line with the requirements of the reference standard NP ISO/IEC 17025: 2018.

RESUMO

A incerteza associada a medições na avaliação sensorial de produtos é fundamental para a comparabilidade de resultados. O cálculo da incerteza associada a dados composicionais tem, assim, merecido uma abordagem transversal em metrologia científica no domínio analítico sensorial, revelando-se fundamental em estudos que impliquem a comparabilidade de resultados de medições, nomeadamente em estudos específicos que têm por objeto a análise sensorial de vinhos. O novo referencial normativo, ISO/IEC 17025:2017, que substituiu a anterior versão ISO/IEC 17025 publicada em 2005, chama a atenção para o conceito de incerteza de medição, referindo que os laboratórios o devem utilizar na avaliação da conformidade e na quantificação do risco associado a decisões. No presente trabalho, efetuou-se o estudo da estimativa da incerteza e da sua importância para a interpretação de resultados de análise sensorial, tendo como objeto os ensaios de análise sensorial inerentes à certificação da Denominação de Origem "Dão", da Denominação de Origem "Lafões" e Indicação Geográfica "Terras do Dão", efetuados por um painel constituído por 20 provadores da Comissão Vitivinícola Regional do Dão, treinado e acreditado pela norma NP ISO/IEC 17025:2018. Os resultados obtidos permitiram concluir que o painel apresenta diferentes valores de incerteza associados aos ensaios que permitiram a determinação dos atributos sensoriais. Os resultados e os respetivos valores de incerteza permitiram avaliar o desempenho do painel de provadores e estabelecer o nível de risco (falsa aceitação e falsa rejeição e pressupostos estatísticos) que se pretende ser associado à regra de decisão utilizada na emissão de declarações de conformidade. Neste estudo, foi apresentada uma nova metodologia para determinar uma estimativa de incerteza a ser aplicada na interpretação de resultados relativos a ensaios de análise sensorial de vinhos para a certificação de Denominação de Origem "Dão", Denominação de Origem "Lafões" e Geográfica Indicação "Terras do Dão". Esta evolução conduziu a uma evolução no dia-a-dia da câmara de prova da Comissão Regional dos Vinhos do Dão, permitindo uma maior solidez nos pareceres emitidos, em linha com os requisitos da norma de referência NP ISO/IEC 17025: 2018.

Keywords: evaluation of uncertainty, statement of conformity, sensory analysis, ANOVA.

Palavras-chave: estimativa de incerteza, declaração de conformidade, análise sensorial; ANOVA.

22

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

INTRODUCTION

Only in recent times scientists have succeeded in developing sensory testing as a formalized, structured, and codified methodology, trying to develop new sensory methods and refining the existing ones (Meilgaard *et al.*, 2016). The main objective of a sensory analysis is to reach valid and reliable tests that provide data based on which sound decisions can be made. As stated by the Nordic Committee on Food Analysis (2013), the importance of measurement uncertainty has gained increased acceptance within most fields of metrology, and there is a vast amount of literature on the subject within other areas, outside the scope of the sensory analysis. In this specific field, the literature is scarce, becoming a rarely touched subject among sensory scientists.

Given the inherent subjectivity and variability of human evaluators and to make measurements of food analysis more objective, trained panels were developed since the 1940's and the field of sensory evaluation emerged and clearly progressed.

The trends in the evolution of sensory evaluation applied to quality control in the last 20 years have been notorious with the focus on three main areas: (i) on the establishment of a standard programme of sensory evaluation and in-house training, together with the instrumental analyses, aimed at assuring product quality; (ii) on food quality certification, requiring a method of panel performance evaluation; (iii) on parameters which could contribute to differentiate products according to the quality grade as well as on the modelization of quality standards (Nguyen *et al.*, 2014).

The issue of uncertainties regarding measured data have led to new scientific approaches in the field of sensory analysis, which have become fundamental to studies that seek to compare results of sensory measurement and analysis, including studies regarding wine.

Some relevant articles related to the theme of quality control in sensory analysis and whose indicators can be used to estimate uncertainty have been addressed by several authors (Lea *et al.*, 1995; Rossi, 2001; McEwan *et al.*, 2002; Kermit and Almlı, 2005; Latreille *et al.*, 2006; Gawel and Godden, 2008; Luciano and Næs, 2009; Etaio *et al.*, 2010a,b; Tomic *et al.*, 2010; Watson and Petrie, 2010; Tomic *et al.*, 2013; Nguyen *et al.*, 2014; Vilela *et al.*, 2015).

With the recent publication of NP EN ISO/IEC 17025:2018, the knowledge of the uncertainty associated with results has become particularly relevant, as this uncertainty falls under the "additional requirements" relating to decision rules, taking into account risks underlying the issuing of statements of conformity in test reports, where the

state of conformity or not, with a given specification, standard or requirement is established. This standard states that "*When a statement of conformity to a specification or standard is provided, the laboratory shall document the decision rule employed, taking into account the level of risk (such as false accept and false reject and statistical assumptions) associated with the decision rule employed, and apply the decision rule*". Knowledge of the value of an associated measurement uncertainty (that is, its level of risk) is therefore crucial to determine compliance and to apply the decision rules established.

As in any analytical determination, the results obtained in sensory analysis can be reliable to a greater or lesser degree, meaning that the magnitude of uncertainty must be assessed (Bower, 2013).

When analysing this issue, the Nordic Committee on Food Analysis (2013) warns that it is important to make the scientific community aware of the fact that data may include errors, and this may have consequences for conclusions drawn. It should be noted that the word "error", with its scientific meaning, is not synonymous with defect or failure. In order to dispel this ambiguity, the word "residual" is often used instead of "error".

Two types of error associated with the statistical tests used are described as follows (Taylor, 2013): (i) Type I error, which occurs when a statistically significant test result suggests a relationship that does not actually exist (false positive); (ii) Type II error, which occurs when a statistically significant test result does not suggest a relationship, but a relationship exists (false negative). Thus, with same level of error being inherent in all measurement, an acceptable level of associated uncertainty should be established.

Sensory analysis assessed by a trained or untrained panel of tasters constitutes a human measurement. While trained tasters cannot be considered machines, a trained sensory panel, composed of several tasters, can be regarded as a specialised instrument for sensory measurement. Such a panel should be able to provide valid and reliable sensory measurements (Meiselman, 2013).

Determining measurement uncertainty increases the reliability of results, facilitates comparability of results from inter-laboratory studies and enables users to make a reasoned decision regarding any differences between the result achieved and the reference value obtained through an inter-laboratory study (Pinto and Barros, 2015).

According to COFRAC – Comité français d'accréditation (2018), in its approach to sensory analysis, the confidence intervals, the standard error

or the power of a test can be used as indicators to measure uncertainty.

The approach used in the present work to calculate uncertainty follows the recommendations from the Nordic Committee on Food Analysis (2013), which uses the analysis of variance (ANOVA) output to estimate the measurement uncertainties. ANOVA procedures represent a powerful approach to the analysis and understanding of sensory analysis data (Lea *et al.*, 1997), based on comparison of variances and using replicates in a balanced sensory design, with replicates applied to observations made at random under a set of equal conditions and with all sources of random variation involved.

The ANOVA is a method of data analysis commonly used to test the significance of the effect of different factors and to estimate the components of variance, that is the amount of variability in observations that can be ascribed to the various sources of variability under analysis. These components of variance are particularly useful in designing sampling plans to monitor several properties of a product being tested and in establishing quality control procedures (Snee, 1974).

The main objective of any measurement is not to obtain the true value, but to associate it with an acceptable uncertainty for the measure in question. Thus, the measurement result is only complete if it is associated with the uncertainty of the estimate.

MATERIALS AND METHODS

Panel of expert assessors

Each tasting session counted seven professional tasters. A group of 20 professional tasters were previously selected, and at least seven tasters had to be present in any session. The large group included four women and 16 men aged between 24 and 70 years. The 24 tasting sessions considered for this study took place at different times, over several weeks. The tasting panel was accredited by the NP EN ISO/IEC 17025:2018, an international standard to determine the competence of analytical laboratories.

Since tasters perceive tastes and odours differently, their concepts of wine quality may diverge considerably. To remedy this situation, it was considered in this study that the existing recommendations should be considered (Jackson, 2014), so the number of tasters should be large enough to buffer individual idiosyncrasies, or the tasters should be trained and selected rigorously for the particular skills required.

For an internal sensory panel, the assessors should have an appropriate level of sensory acuity. Normally, to comply with this requirement, screening tests may be used based on the ability to recognize basic tastes and accurately compare model

solutions with varying levels of, for example, sucrose, sodium chloride or citric acid. (Kemp *et al.*, 2018).

High attention has been paid to this, and the recruitment, selection, training and monitoring of candidates intended to become sensory assessors of Dão Regional Wine Commission followed the guidelines established in the ISO 8586:2012.

According to ISO 11132:2012, performance in the context of a descriptive panel and individual panellists comprises of the ability to detect, identify, and measure an attribute, use attributes in a similar way to other panels and within tasters, discriminate between stimuli, use a scale properly, repeat their own results and reproduce results from other panels and other tasters. Special attention was given to taster's qualification based on four important qualities: their ability to differentiate the products (discrimination) consistently (repeatable) and consensually (in agreement) applicable on all the sensory attributes and to ensuring detection and recognition of typical wine taints, defects and off-notes.

Panel members were pre-trained in the use of the scales, and reference samples were used for each of the anchor points as recommended in ISO 8586:2012. The pre-training involved familiarization with the use of the adopted scale using wines that were submitted to this study. The practice of routine tasting in a business environment was also a requirement giving them time to learn how to use the scale.

Samples

The wines under analysis corresponded to all the wine typologies within the scope of accreditation of the Dão Regional Wine Commission: wine with Designation of Origin "Dão"; wine with Designation of Origin "Lafões"; wine with Geographical Indication "Terras do Dão"; sparkling wine with Designation of Origin "Dão" (Comissão Vitivinícola Regional do Dão, 2019). Of these, 97.0% corresponded to wines with Designation of Origin "Dão", 2.20% of Geographical Indication "Terras do Dão", and 0.70% of Designation of Origin "Lafões". In general, 65.9% were red wines, 29.6% were white wines, and 4.4% were rosé wines. A total of 127 wines were considered in this study.

Procedure

This study was integrated into the usual work dynamics, and for the sake of representativeness, all tasters assessed all wines within the scope of the reference standard NP ISO/IEC 17025:2018, accreditation and taking into account the different colour categories (white, red, and rosé). All tasters were able to contribute to this study without changing the usual work dynamics.

The 24 tasting sessions took place in a tasting room which generally follows the guidelines defined in the ISO 8589:2007 standard, on different days, usually one session per week and always in the morning. In the scope of this study, one tasting session consists of several tasting series according to the wine category/colour (white, red, and rosé) and with pre-established rules according to the internal methodology adopted by the Dão Regional Wine Commission "RI-08 Sensory Analysis"(Comissão Vitivinícola Regional do Dão, 2020)

All sessions were analysed by a minimum of seven tasters and included 24 wines per tasting, of which five or six corresponded to replicated wines (served in duplicate), ensuring the inclusion in each series of samples of at least one replica. Replication was found to be a good way to assess descriptive analysis results to evaluate tasters for consistency and agreement with other tasters, and it is essential for all sensory analytical evaluation because of the nature and the limited number of subjects in each test. The use of replication provides insight into individual response behaviour and is crucial to assess taster reliability. In this study, the definition of replica follows the approach made in EA-4/09:G2017 that corresponds to "evaluate a sample more than once"; "a replicated observation is defined as a new independent measurement taken under the same set of conditions as the original one" (Lea *et al.*, 1997).

Bearing in mind that there can be numerous sources of variability, a deep effort was made to control, minimize, and/or measure them. Particular attention has been paid to product coding, test environment, serving order, and design and analysis, in addition to tasters and products.

All factors that can influence sensory analysis were properly controlled to minimise their possible effects on the results, allowing each taster to have a similar sensory experience, under the same environmental conditions.

The study focused on nine sensory attributes: appearance – colour intensity; appearance - limpidity; aroma - complexity; aroma - intensity; flavour - complexity; flavour - balance; flavour - bitterness; flavour - body; flavour - after taste.

All tasters called for the tasting session were asked to evaluate the same wine samples, which were different from each other, presenting different sensory properties, and were provided to tasters in a random sequence in a balanced sensory design (Rogers, 2017) submitted to simultaneous assessment. Tasting glasses used had the characteristics defined in ISO 3591:1977, and were

marked at the base with random three-digit codes to identify the sample. In the balanced sensory design adopted, an equal number of replicates of each sample was presented to each taster; each sample was submitted an equal number of times; each sample was submitted in combination with all the samples an equal number of times during the session; and the order of the samples was randomly assigned to each taster.

To achieve sensorial results, a linear scale for Quantitative Descriptive Analysis (QDA®) was used. The sensory description was obtained by rating a set of products according to the list of sensory attributes listed above. There was no instrumental support in the evaluation of sensory attributes or access to other laboratory analytical information, so the evaluation was strictly performed by the taster's sensory apparatus.

This is an approach that allows the intensities of all important sensory attributes to be specified. According to the literature (Stone *et al.*, 1974), in this situation, the use of a QDA® scaling method similar to an interval scale is recommended, since analysis of variance is the standard statistical technique for comparing products in descriptive analyses.

An individual result capture system (CVRD_Mod.025.02:14-03-2019) developed specifically for this purpose was used. Each taster was instructed to make a mark along a straight line to indicate the intensity of the attribute or their degree of preference, with the score increasing from left to right (minimum 0, maximum 10). A total line length of 10 cm was established. A mark was made by a taster sliding their finger over a bar on touch screen monitors. The scale used to evaluate the sensory attributes is exemplified in Figure 1.

The boundaries of the line were designated as "not very intense" vs. "very intense" or "unbalanced" vs. "balanced". Lines for different attributes were positioned sequentially. A small pilot test was done to eliminate potential problems with the use of the adopted scale. An algorithm available in the computer system for capturing results allowed the conversion of the relative position of the mark on the bar into a number representing the distance to the end measured in centimetres.

It is desirable to adopt a way of acquiring results that allows for greater discrimination since this will imply a lower number of evaluations and, consequently, a lower number of analysis/judgements and a reduction of the corresponding time and cost (Park *et al.*, 2007).

ASPETO/ASPECT

Colour Intensity

Intensidade de Cor

Inadequada Adequada

Inappropriate. Appropriate.

...

Clarity

Limpidez

Muito Turvo Limpido/Brilhante

Very cloudy. Clean / Bright.

...

AROMA/AROMA

Aromatic complexity

Complexidade Aromática

Simples (1 Descritor) Muito Complexo (≥4 Descritores)

Simple (1 descriptor). Very complex (> 4 descriptors).

...

Aromatic Intensity

Intensidade Aromática

Pouco Intenso Muito Intenso

Not Very Intense. Very Intense.

...

FLAVOR/FLAVOUR

Complexity

Complexidade Aromática

Simples (1 Descritor) Muito Complexo (≥4 Descritores)

Simple (1 descriptor). Very complex (> 4 descriptors).

...

Balance (Acidity / Sweetness)

Equilíbrio (Acidez/Doçura)

Desequilibrado Equilibrado

Unbalanced. Balanced.

...

Astringency / Bitterness

Adstringência/Amargor

Desequilibrado Equilibrado

Unbalanced. Balanced.

...

Body / Structure

Corpo/Estrutura

Simples Encorpado

Simple. Full-bodied.

...

Persistence

Fim de Boca

Curto Prolongado (≥4 segundos)

Short. Extended (> 4 seconds).

Figure 1. Example of wine tasting sheet.

Statistical treatment

Data analysis was performed using Microsoft Excel (Microsoft, 2020) and XLSTAT statistical tools (Addinsoft, 2019).

Repeated trials design and the analysis were established in order to measure sources of variability and separate (partition) them from the main effects before determining whether the variables of interest were significant.

The reason why this option was made based on the work of some reference authors, namely that consider ANOVA the option method to understand the influence of experimental factors (a factor of an experiment is a controlled independent variable; a variable whose levels are set by the experimenter) on a quantitative dependent variable (the dependent variables are variables to characterize or predict from other independent variables) (Lê and Worch, 2015; Stone *et al.*, 2020). The dependent variables were considered the sensory attributes; the experimental factors, also called independent variables, were considered the factors associated with the product effect, the taster effect, the session effect, and eventually all their first-order interactions. Notion of interaction is a concept of utmost importance in statistics, as important as it is hard to apprehend and, consequently, hard to interpret practically (Lê and Worch, 2015). These authors stressed that the knowledge about interaction is important in understanding the nature of disagreements between individual subjects and a panel.

In this study, the guidelines of the Nordic Committee on Food Analysis (2013) were followed and the assessment of measures of uncertainty was calculated using a two-factor ANOVA with replication for each of the sensory attributes across 24 tasting sessions.

Different ANOVA models must be used depending on the design of the sensory panel. If replicate evaluations are performed in the same session, all samples are evaluated multiple times in the same session, then a two-way ANOVA model, with taster and sample as the effects, should be used (E3000–18).

RESULTS AND DISCUSSION

The data set considered in this study result from a total of 24 tasting sessions carried out in different time periods and consists of 127 wines, evaluated twice by a total of 20 tasters in nine attributes.

The data was validated to verify whether missing or incorrect data was detected and to ensure that the conclusions are consistent with the original database. This step involves looking at the whole data set to ascertain if the results were gathered correctly (for example, no sample presentation issues), are robust

enough to use, relevant to the project (meet the action standard requirements) and that the data set is complete (no missing data).

In each of the 24 tasting sessions, a minimum of five wine replicas were included. A two-factor ANOVA with replication was performed for the nine sensory attributes resulting in a total of 260 ANOVA outputs. It is not possible in a study of this nature to show the complete data; for example, the calculation of scores for the attribute "appearance – colour Intensity" in tasting session nr. 6 is given in Table I.

The data presented in Table I are the results of a 6 x 7 cross classification design with duplicate observations obtained at each of the 6 x 7 = 42 treatment combinations for a total of 84 observations. The two-way ANOVA output for these data is shown in Table II. These values were obtained in accordance with the recommended procedure 27 of the Nordic Committee on Food Analysis (2013).

For each of the sources of variation, the expected mean squares (E(MS)) were calculated to establish which F-tests should be performed to examine the hypotheses on the sources of variation, with (E(MS)) indicating the appropriate error term for each variation source (Snee, 1974).

The expected mean of squares, (E(MS)), are used to determine the appropriate error terms for the source of variation in an ANOVA model (Sit, 1995).

The Nordic Committee on Food Analysis recommendations were followed in this procedure. Therefore, standard deviations and variances for the different sources of variation were thus estimated starting with the Error (bottom row of Table III) and replacing all the values σ^2 in column (E(MS)) of Table III with their estimates. The variance, which represents the extent to which tasters repeat themselves, was then estimated by MS (Error), equal to 0.044 or 0.2098 in the form of standard deviation. It should be highlighted that this standard deviation, as well as all the other conclusions derived from Table II, applies to all tasters. To draw similar conclusions based on a single taster, the ANOVA must be carried out for each taster separately.

The interest of the present study lays in the effect of the Replicates*Taster interactions. In the line corresponding to Replicates*Taster, σ^2 was replaced by the value found (0.044) to obtain the estimate 0.112 for MS (Replicates*Taster), or 0.335 in the form of standard deviation by applying the square root of 0.112. The values for MS (Replicates) and MS (Taster) can be estimated in a similar way.

After 24 tasting sessions, each with a minimum of five replicates, the values of uncertainty (expressed as standard deviation) were obtained for each of the sensory attributes listed in Table III.

Table I

Scores obtained in tasting session nr. 6 for the sensory attribute "appearance - colour intensity" (values in cm)

	Taster 1	Taster 2	Taster 3	Taster 4	Taster 5	Taster 6	Taster 7
Replicate 1a	8.3	7.5	8.3	7.3	8.4	8.6	8.6
Replicate 1b	8.3	8	8.4	7.5	8.2	8.2	8.8
Replicate 2a	8	7	7.8	6.4	8.2	7.5	8.8
Replicate 2b	7.3	7.3	7.2	6.4	7.8	7.5	8.7
Replicate 3a	7.7	7.7	7.8	6.6	7.4	7.3	8.8
Replicate 3b	7.5	7.3	8.4	6.8	7.5	7.6	8.9
Replicate 4a	8.1	8	8.7	8	9	9.9	9
Replicate 4b	8.4	8.1	8.2	7.8	8.7	9.6	8.8
Replicate 5a	7.7	8.2	8.2	7.1	8.2	8.6	8.8
Replicate 5b	8	7.7	8.2	7.5	8.5	8.5	8.7
Replicate 6a	8.6	7.7	8.2	7.9	8.5	9.8	8.8
Replicate 6b	8.6	7.6	8	7.7	8.6	9.7	8.9

Replicate 1 = red wine from Designation of Origin "Dão"; Replicate 2 = red wine from Geographical Indication "Terras do Dão"; Replicate 3 = red wine from Designation of Origin "Dão"; Replicate 4 = red wine from Designation of Origin "Dão"; Replicate 5 = red wine from Designation of Origin "Dão"; Replicate 6 = red wine from Designation of Origin "Dão".

Table II

ANOVA results for the sensory attribute "appearance – colour intensity"

Source of variation	SS	df	MS	F	p-value	F crit.	(E(MS))
Replicates	12.097	5	2.419	54.927	2.429E-17	2.437	$\sigma^2 + 2 * \sigma^2$ (Replicates*Taster) + 14 * S (Replicates)
Taster	19.645	6	3.274	74.332	9.124E-21	2.323	$\sigma^2 + 2 * \sigma^2$ (Replicates*Taster) + 12 * σ^2 (Taster)
Replicates*Taster	8.0478	30	0.268	6.090	7.568E-08	1.730	$\sigma^2 + 2 * \sigma^2$ (Replicates*Taster)
Error	1.85	42	0.044				σ^2
Total	41.64	83					

SS = sum of squares; df = degrees of freedom; MS = average sum of squares (= SS/df); F = MS for any source (wine, panel or taster x wine) divided by the appropriate error term; p-value = the probability/error of obtaining an F value which is at least as large as the value actually observed if there was no source of variation ('null hypothesis'); F crit. = a point on the statistical distribution under the null hypothesis which defines a set of values that allows for the rejection of the null hypothesis. This set is called the critical or rejection region; (E(MS)) = expected value expressed in terms of variance components. For details about general rules on calculating the expected mean of squares and for information on unbalanced or more complex models, Montgomery (1991) and Neter *et al.* (1985) should be consulted.

As the sessions included replicate samples, it became possible to perform a two-way ANOVA in which the data of the taster, the replicate samples (served in duplicate), and the effect of interaction were evaluated.

In general, the more observations (replicates) are available the more precise conclusions can be (Lea *et al.*, 1997). In practice, however, there may be economic and other constraints that impose limits on experimental design; that is, there is always a compromise between the ideal situation from a statistical point of view and what is practically feasible.

The results of the uncertainty are shown in Table III (expressed in standard deviation) for the sensory attributes studied over the various sessions held.

Variances are additive, but standard deviations are not (Walker, 2018). This means that the variance of the sum of two independent (uncorrelated) random variables is simply the sum of the variances of each of the variables. This is important for many statistical analyses. The units of variance are the square of the original units, which is difficult to interpret. The units of a standard deviation are the same as the original variable and, therefore, are much easier to interpret.

Table III

Values of uncertainty (expressed in standard deviation) for the sensory attributes studied over the various tasting sessions held

Session	Data	App. - -Col. Int.	App. - Limp	Aro. - Comp.	Aro. - Int	Flav. - Comp.	Flav. - Bal.	Flav. - Bit.	Flav. - Bod.	Flav. - AT
1	02/05/2019	0.22	0.19	0.47	0.46	0.40	0.45	0.54	0.37	0.39
2	01/08/2019	0.25	0.11	0.40	0.42	0.31	0.38	0.44	0.42	0.50
3	19/09/2019	0.19	0.28	0.48	0.48	0.46	0.48	0.52	0.50	0.56
4	26/09/2019	0.38	0.29	0.50	0.51	0.43	0.41	0.43	0.44	0.41
5	28/09/2019	0.23	0.18	0.53	0.56	0.48	0.53	0.48	0.54	0.57
6	18/10/2019	0.33	0.25	0.40	0.35	0.39	0.39	0.39	0.40	0.44
7	14/11/2019	0.35	0.28	0.28	0.28	0.21	0.27	0.19	0.27	0.25
8	21/11/2019	0.19	0.30	0.56	0.52	0.43	0.41	0.48	0.41	0.41
9	19/12/2019	0.31	0.22	0.24	0.23	0.18	0.27	0.17	0.24	0.22
10	27/12/2019	0.45	0.24	0.41	0.45	0.51	0.60	0.65	0.60	0.64
11	06/02/2020	0.22	0.09	0.32	0.33	0.36	0.36	0.40	0.41	0.37
12	27/02/2020	0.37	0.17	0.40	0.44	0.46	0.46	0.48	0.47	0.54
13	05/03/2020	0.38	0.16	0.43	0.44	0.37	0.46	0.37	0.37	0.40
14	06/03/2020	0.21	0.13	0.40	0.31	0.48	0.39	0.33	0.36	0.33
15	19/03/2020	0.28	0.27	0.29	0.32	0.31	0.33	0.29	0.32	0.29
16	08/04/2020	0.37	0.16	0.38	0.34	0.41	0.43	0.50	0.48	0.54
17	09/07/2020	0.31	0.15	0.55	0.58	0.50	0.60	0.57	0.54	0.59
18	17/07/2020	0.41	0.09	0.14	0.17	0.34	0.34	0.43	0.36	0.37
21	11/09/2020	0.22	0.25	0.65	0.65	0.62	0.60	0.59	0.61	0.62
19	01/10/2020	0.30	0.10	0.49	0.44	0.43	0.47	0.42	0.47	0.52
23	12/11/2020	0.25	0.24	0.61	0.59	0.41	0.39	0.41	0.37	0.40
20	19/11/2020	0.35	0.18	0.46	0.49	0.51	0.52	0.54	0.54	0.53
22	11/12/2020	0.29	0.10	0.30	0.30	0.31	0.27	0.30	0.30	0.28
24	30/12/2020	0.40	0.06	0.28	0.24	0.33	0.37	0.43	0.33	0.35

The highest values of uncertainty are shaded in red. **App. - Col. Int.** = appearance – colour Intensity; **App. - Limp.** = appearance - limpidity; **Aro. - Comp.** = aroma - complexity; **Aro. - Int.** = aroma - intensity; **Flav. - Comp.** = flavour - complexity; **Flav.- Bal.** = flavour - balanced; **Flav.Bit.** = flavour - bitterness; **Flav. - Bod.** = flavour - body; **Flav.- AT** = flavour – after taste/persistence.

By observing the data related to the uncertainty values found in the different tasting sessions and for the different attributes, it can be inferred that the lowest values of uncertainty obtained were those related to the attributes of the appearance (colour intensity and limpidity), which was expectable as it is a more objective parameter. For the remaining sensory attributes (aroma - complexity; aroma - intensity; flavour - complexity; flavour - balanced; flavour - bitterness; flavour - body; flavour - after taste), the uncertainty values found are higher when compared to those of the appearance, and very similar to each other, with minor differences between them.

The highest values of uncertainty are highlighted in Table III and an attempt was made to obtain an explanation for this fact. The scores attributed by the tasters in the replicas of these tasting sessions and the sensory attributes with the highest values of

uncertainty were examined. It was found that the divergence between the tasters in the appreciation of these sensory attributes was greater, which obviously led to higher values of uncertainty.

Individual variation in odorant perception has long been known. Variation results not only from the ability to detect, identify, and measure the intensity of odours, but also from emotional response to them (Jackson, 2014).

An explanation for the inconsistency of values is been under study and, although training improves the consistency, it is unlikely to eliminate genetic-based idiosyncrasies (Lawless, 1984).

CONCLUSIONS

The new normative framework, ISO/IEC 17025:2017, which replaces the previous version

ISO/IEC 17025, published in 2005, draws attention to the concept of calculation uncertainty, stating that it should be considered by laboratories when assessing conformity and quantifying risks arising in decision-making processes.

Prior to this study, it was not possible to characterise the level of risk associated with each of the sensory attributes or to assess the overall risk when seeking to assess the conformity of a product with the specifications of a given wine.

In this study, it was possible to determine an estimate of uncertainty, which can be applied in the interpretation of results relating to sensory analysis tests of wines for the certification of Designation of Controlled Origin "Dão", Designation of Origin "Lafões" and Geographical Indication "Terras do Dão".

The results allowed drawing the conclusion that the panel presented different values of uncertainty associated with the various sensory attributes. The uncertainty data obtained is a tool to be used in maintaining the panel's activity in decision-making. It will also allow the performance of the panel of tasters of the Dão Regional Wine Commission to be characterised. Such results are essentially useful in assessing the level of risk associated with the decision rule, improving the reliability of statements of conformity to be issued.

The approach described in procedure 27 of the Nordic Committee on Food Analysis (2013) has proved to be appropriate, simple to carry out, and sufficiently comprehensive to encompass all sources of uncertainty.

Through the study here presented, it was demonstrated that it is possible to fully comply with the requirements of the standard NP EN ISO/IEC 17025:2018 regarding the presentation and estimation of uncertainty in sensory testing.

ACKNOWLEDGEMENTS

The author thanks the Dão Regional Wine Commission for its collaboration in the provision of data essential for this study and for its financial support for the translation of the article.

REFERENCES

Addinsoft, 2019. XLSTAT. 2019.2.2.

Bower J.A. 2013. Statistical methods for food science: Introductory procedures for the food practitioner. 334 p. Wiley-Blackwell, Chichester.

COFRAC, 2018. Guide technique d'accréditation - Analyses sensorielles et tests consommateurs. Available at: <https://tools.cofrac.fr/documentation/LAB-GTA-33> (accessed on 20.12.2021)

Comissão Vitivinícola Regional do Dão, 2019. Anexo técnico de acreditação N° L0230-1. Instituto Português de Acreditação. Available at:

[http://www.cvrdao.pt/images/documentos/ATE_L0230A1\(2020-11-10\).pdf](http://www.cvrdao.pt/images/documentos/ATE_L0230A1(2020-11-10).pdf) (accessed on 10.12.2020)

Comissão Vitivinícola Regional do Dão, 2020. RI08 – Análise sensorial. Available at: <http://www.cvrdao.pt/images/documentos/RI08-An%C3%A1lise%20Sensorial-02-2020.pdf> (accessed on 28.08.2020)

E3000–18. Standard guide for measuring and tracking performance of assessors on a descriptive sensory panel., 13 pp. ASTM International, West Conshohocken.

EA-4/09:G2017. Accreditation for sensory testing laboratories. Available at: <https://european-accreditation.org/wp-content/uploads/2018/10/ea-4-09-g-rev02-february-2017.pdf> (accessed on 15.11.2020).

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

Etaio I., Albisu M., Ojeda M., Gil P., Salmerón J., Elortondo F.P., 2010b. Sensory quality control for food certification: A case study on wine. Panel training and qualification, method validation and monitoring. *Food Control*, **21**, 542-548.

Gawel R., Godden P., 2008. Evaluation of the consistency of wine quality assessments from expert wine tasters. *Aust. J. Grape and Wine Res.*, **14**, 1-8.

ISO 3591:1977. Sensory analysis - Apparatus - Wine-tasting glass. 3 p. International Organization for Standardization, Geneva.

ISO 8586:2012. Sensory analysis - General guidelines for the selection, training and monitoring of selected assessors and expert sensory assessors. 28 p. International Organization for Standardization, Geneva.

ISO 8589:2007. Sensory analysis - General guidance for the design of test rooms. 16 p. International Organization for Standardization, Geneva.

ISO 11132:2012. Sensory analysis - Methodology - Guidelines for monitoring the performance of a quantitative sensory panel. 23 p. International Organization for Standardization, Geneva.

ISO/IEC 17025:2017. General requirements for the competence of testing and calibration laboratories. 30 p. International Organization for Standardization, Geneva.

Jackson R.S., 2014. Wine assessment and sensory analysis. In: *Wine science: Principles and applications*. 865 - 866. Jackson R.S. (ed.), Elsevier, Cambridge.

Kemp S.E., Hort J., Hollowood T., 2018. Descriptive analysis in sensory evaluation. 724 p. Wiley-Blackwell, Hoboken.

Kermit M., Almlí V., 2005. Assessing the performance of a sensory panel- panelist monitoring and tracking. *J. Chemom.*, **19**, 154-161.

Latreille J., Mauger E., Ambroisine L., Tenenhaus M., Vincent M., Navarro S., Guinot C., 2006. Measurement of the reliability of sensory panel performances. *Food Qual. Prefer.*, **17**, 369-375.

Lawless H.T., 1984. Flavor description of white wine by expert and nonexpert wine consumers. *J. Food Sci.*, **49**, 120-123.

Lê S., Worch T., 2015. Analyzing sensory data with R. 374 p. CRC Press, Boca Raton.

Lea P., Næs T., Rødbotten M., 1997. Analysis of variance for sensory data. 102 p. Wiley, Chichester.

Lea P., Rødbotten M., Næs T., 1995. Measuring validity in sensory analysis. *Food Qual. Prefer.*, **6**, 21-326.

- Luciano G., Næs T., 2009. Interpreting sensory data by combining principal component analysis. *Food Qual. Prefer.*, **20**, 167-175.
- McEwan J.A., Hunter E.A., Gemert L.J., Lea P., 2002. Proficiency testing for sensory profile panels: measuring panel performance. *Food Qual. Prefer.*, **13**, 181-190.
- Meilgaard M.C., Civille G.V., Carr B.T., 2016. Sensory evaluation techniques. 630 p. CRC Press, Boca Raton.
- Meiselman H., 2013. The future in sensory/consumer research: Evolving to a better science. *Food Qual. Prefer.*, **27**, 208-214.
- Microsoft, 2020. Microsoft Excel for Microsoft 365. 16.0.13628.20234 64 bits.
- Montgomery D., 1991. Design and analysis of experiments. 684 p. John Wiley and Sons, Hoboken.
- Neter J., Wasserman W., Kutner M.H., 1985. Applied linear statistical models. 1396 p. Irwin Inc., New York.
- Nguyen D.Q., Le T.M., Nguyen D.H., 2014. Role of sensory evaluation in quality control: A textual point of view. In: *HCMUT*, Ho Chi Minh city.
- Nordic Committee on Food Analysis (2013). NMKL Procedure No. 27: Measurement uncertainty in sensory analysis, 1-24.
- NP ISO/IEC 17025:2018. General requirements for the competence and calibration laboratories. 37 p. International Organization for Standardization, Geneva.
- Park J.Y., O'Mahony M., Kim, K.O., 2007. 'Different-stimulus' scaling errors; effects of scale length. *Food Qual. Prefer.*, **18**, 362-368.
- Per Lea - Nofima Mat, Ås, Norway. Measurement uncertainty in sensory analysis. European Sensory Network. Available at: <https://esn-network.com/research/conferences/pangborn-2009/lea-nofima/> (accessed on 21.07.2020)
- Pinto M., Barros P., 2015. Ensaio de aptidão sensorial em química enológica. In: *Química enológica - métodos analíticos. Avanços recentes no controlo da qualidade de vinhos e de outros produtos*. 497-505. Curvelo Garcia A.S., Barros P. (ed.). Publindústria, Porto.
- Rogers L., 2017. Sensory panel management: A practical handbook for recruitment, training and performance. 376 p. Elsevier Science, Cambridge.
- Rossi F., 2001. Assessing sensory panelist performance using repeatability and reproducibility measures. *Food Qual. Prefer.*, **12**, 467-479.
- Sit V., 1995. Analyzing ANOVA designs biometrics information handbook no 5. Available at: <https://www.for.gov.bc.ca/hfd/pubs/Docs/Wp/WP07.pdf> (accessed on 18.11.2020).
- Snee R.D., 1974. Computation and use of expected mean squares in analysis of variance. *J. Qual. Technol.*, **6**, 128-137.
- Stone H., Bleibaum R.N., Thomas H.A., 2020. Descriptive analysis. In: *Sensory evaluation practices*. 265. Stone H., Bleibaum R.N., Thomas H.A., Academic Press, New York.
- Stone H., Sidel J., Oliver S., Woolsey A., Singleton R.C., 1974. Sensory evaluation by quantitative descriptive analysis. In: *Descriptive sensory Analysis in practice*. 23-34. Gacula M.C. (ed), Wiley-Blackwell, Scottsdale.
- Taylor C.S., 2013. Validity and validation (understanding statistics). 206 p. Oxford University Press, New York.
- Tomic O., Forde C., Delahunty C., Næs T., 2013. Performance indices in descriptive sensory analysis – A complimentary screening tool for assessor and panel performance. *Food Qual. Prefer.*, **28**, 122-133.
- Tomic O., Luciano G., Nilsen A., Hyldig G., Lorensen K., Næs T., 2010. Analysing sensory panel performance in a proficiency test using PanelCheck software. *Eur. Food Res. Technol.*, **230**, 497-511.
- Vilela A., Monteiro B., Correia E., 2015. Sensory profile of Port wines: categorical principal component analysis, an approach for sensory data treatment. *Ciência Tec. Vitiv.*, **30**, 1-8.
- Walker J.A., 2018. Elements of Statistical Modeling for Experimental Biology. Available at: https://www.middleprofessor.com/files/applied-biostatistics_bookdown/_book/ (accessed on 20.01.2021)
- Watson P.F., Petrie A., 2010. Method agreement analysis: A review of correct methodology. *Theriogenology*, **73**, 1167-1179.