# Representational and sensory cues as drivers of individual differences in expert quality assessment of red wines

André Caissie, Laurent Riquier, Gilles de Revel, Sophie Tempere

Version of Record: https://www.sciencedirect.com/science/article/pii/S0950329320303013
Manuscript_ddf9ada8ccbe0ff063f808060258d07f



Short title: Drivers of individual differences in red wine assessments

# Representational and Sensory cues as Drivers of Individual Differences in Expert Quality Assessment of Red Wines

André F. Caissie[1], Laurent Riquier[1], Gilles De Revel[1] & Sophie Tempere[1]

[1] Unité de recherche Oenologie, EA 4577, USC 1366 INRA, ISVV, Université de Bordeaux, Bordeaux INP, F33882 Villenave d'Ornon France

Address correspondence to:

André F. Caissie, at andre.caissie@gmail.com

Sophie Tempere, at sophie.tempere@u-bordeaux.fr

## Abstract

The aim of this study was to model decisional consensus in expert red wine tastings, using an integrated competency framework. Wine assessment responses on both technical and emotional scales were collated for two wine categories (Premium vs. Secondary) under several different sensory conditions: six global tastings (all senses involved), three unimodal tastings (visual, smell, and taste), and three bimodal tastings (visual-smell, visual-taste, and taste-smell). Psychological predictors also included vocabulary and vividness of mental imagery associated with the various senses involved, together with professional experience indicators (age and tasting frequency). Principal component analyses revealed a greater response consensus with unimodal vision cues compared to all other sensory conditions (at least equal to global conditions). On average, a greater consensus was observed among technical quality scale responses under all sensory conditions, compared to emotional scale responses. The quality responses were used to build a 4-factor prediction model: age, wine imagery, vocabulary, and smell consensus. The image responses were used to build a two-factor prediction model: visual words (semantic knowledge) and visual-smell consensus. This indicated that the quality decisional consensus was based on smell information (wine aroma), combined with longevity/knowledge. In contrast, the image decisional consensus was based on visual information (wine color), combined with visual knowledge (and smell as a subordinate factor). Taken together, our results revealed previously uncharted individual differences in wine tasting and decision-making, concomitant with similarly weighted predictions based on sensory and psychological factors.


**Key words:** wine tasting; quality decisions; perception; vocabulary; mental imagery

## Introduction

Expert wine tasters are often required to make quality decisions based on multiple features at different stages in winemaking (clarification, blending, marketing). Wine tasting involves precise sensory evaluations, in which the taster's eyes, nose and mouth assess the organoleptic properties of the wine. Assessments are also likely to depend on the wine taster's own representations and the wine tasting instructions. In view of its multisensory nature and the various representational contexts, understanding wine thus involves a considerable number of decisional parameters.

The weight given to each aspect of sensory information varies, as each sense reveals specific information about the wine. Visual inputs are very important for initial user-product interactions (Schifferstein & Spence, 2008). They also provide accurate information in the shortest time (Vision: Herz & Engen, 1996; Smell: Holley & MacLeod, 1977), influencing subsequent smell and taste analyses (Spence, 2010). Visual analysis reflects regional knowledge about the physical aspects and color of a good-quality wine (Valentin *et al.*, 2016). Smell plays a different role, as the affective dimension dominates olfactory cognition of evoked memories (Engen, 1988; Herz & Schooler, 2002). Taste may also elicit effective and affective responses, but in a secondary or complementary role compared to the other senses (Sáenz-Navajas et al., 2016a; Varela & Gambaro, 2006). Moreover, the weight given to each sensory input may also vary among wine tasters.

Wine tasting is thus a core competency involving several sensory features and representations. In order to extract the most relevant information, expert wine tasters also rely on organized knowledge, based on the concept of a common professional foundation or strategy for assessing wines (Jackson, 2009; Peynaud, 1987). The minimum requirement for response consensus is a learning experience common to the individuals in a group (training, profession). Consensus further predicts that a given individual, A, will respond in a certain way to a specific stimulus, B, and that a group of individuals sharing the same experience as individual A will respond similarly to the same stimulus. Admittedly, complete decisional consensus remains theoretical, as wines exhibit varying sensory features, which, in turn, may elicit a variety of psychological responses. For example, wine quality decisions depend on the specific culture (i.e., geographical area), with experienced tasters only tending toward decisional consensus (Sáenz-Navajas et al., 2016a). It was, therefore, interesting to investigate the wine response factors involved in expert decision-making, to evaluate whether individual differences were mainly driven by sensory or psychological factors, or a combination of both.

Some studies have reported inter-individual physiological and psychological differences, especially in olfaction (Tempere et al., 2011; Tempere et al., 2014a; Criado et al., 2019) and mouthfeel (e.g., astringency, Bajec and Pickering, 2008). For these senses, varying wine representations and descriptions make it difficult to obtain a strong consensus in wine tasting (Sauvageot, Urdapilleta, & Peyron, 2006; Parr et al., 2014; Loison et al., 2015; Muñoz-Gonzales et al., 2019).

Most studies have compared experts to novice tasters (Hughson & Boakes, 2001, 2002, 2009; Lawless, 1984; Melcher & Schooler, 1996; Parr *et al.*, 2011; Sáenz-Navajas *et al.*, 2015; Solomon, 1990). Expert wine tasters have better-organized knowledge (Hughson & Boakes, 2002, 2001) and better-defined category sets of vocabulary descriptors for wines compared to novices (Lehrer, 1975; Melcher & Schooler, 1996; Solomon, 1990). Their choice of words is also more abstract and technical, compared to an untrained matched panel (Gawel, 1997). Moreover, the search for specific sensory cues relates to prior experience or learned representations of wine quality hierarchies: each wine taster has wine models (or prototypes, Brochet & Dubourdieu, 2001) with which to compare actual tasting cues (common models within the same production area: Hopfer & Heymann, 2014; Torri *et al.*, 2013). These models have been shown to become increasingly precise following repeated wine-tasting experience (Schifferstein & Desmet, 2008). It is thus interesting to examine the issue of consistency among these representational models in a selected group of expert wine tasters from the same area.

The objective of this study was to investigate both the sensory and psychological determinants of decisional consensus among tasters during wine tastings and assess whether common representations emerged. Data were collected on individual responses to twenty commercial wine samples, during twelve wine tastings, under several different sensory conditions. A group of expert wine tasters scored each red wine using a holistic response scorecard. Indeed, as reported by Lawless (1995) and Sáenz-Navajas et al. (2016b), holistic approaches are better suited to considering multisensory integration processes and individual differences. Different forms of holistic questions were tested, representative of the multidimensional aspect of the quality concept: arousal, quality, certainty of quality, image, and hedonism. For prediction, the wines assessed were based on *a priori* quality distinctions (premium vs. secondary), considered readily available to the tasters under usual tasting conditions (i.e., all senses involved). This created a benchmark for evaluating the role of each of the sensory conditions (unimodal, bimodal) and various psychological processes that contribute to wine assessments.

The wine tasting schedule featured question criteria on technical and emotional scales, evaluating each sensory modality in isolation (unimodal) or in combination with other sensory modalities (bimodal or global). Data analysis also integrated psychological task scores related to verbal performance and vividness of imagery. Verbal descriptors were expected to reveal the regional 'corpus' used by the wine tasters to describe wine quality (as a specific domain of expertise: Rapp, 2014; Brochet & Dubourdieu, 2001). Imagery task performance was designed to reveal the influence of representational prototypes (Andrade, May, Deeprose, Baugh, & Ganis, 2013; Marks, 1973) on expert decisions (Croijmans, Speed, Arshamian, & Majid, 2020). Professional indicators and age completed the experience factors. The model predictions thus translated into four potential predictors of consensus under *a priori* quality contrasts (sensory, verbal, mental imagery, and experience factors) for an approach to wine assessment based on individual differences.

## Methods

### *2.1. Participants*

Sixty-two expert wine tasters (24 women, 38 men; mean age: 48.5 ± 9.5 years; French-speaking) from the Bordeaux area in France participated in this study. They were all professional wine tasters and/or had significant experience tasting the wines from the participating PDO area (average 3.14 ± 3.22 tastings per week). None of the wine tasters reported any particular health impairments or sensory loss and all provided written informed consent prior to the start of the study. They were treated according to the WMA Declaration of Helsinki guidelines for research (World Medical Association 2013, para. 26). Only 22 wine tasters completed all six sessions (mean age: 49.1 ± 9.4, 15 males, 7 females). Careful analysis was thus required to compensate for this methodological issue.

### *2.2. Wine selection*

Twenty commercial red wines were selected for the purposes of this study. They included 14 premium and 6 secondary wines, according to the current quality hierarchy in the participating PDO area (2014 vintage). The wines were representative of the commercial wine population concerned, originated from the winegrowing area of interest, and were matched by soil type (limestone-clay, sandy-clay, gravelly-sand), grape variety, and technical approach. Decisions on inclusion from among an initial set of 40 wines also followed an informed set of screening tests involving sensory selection, as well as physical and chemical analyses. A panel of five judges tasted the wines individually, under blind conditions, decided whether they detected any obvious faults, and noted their decision. These tastings were coupled with GC-MS analysis with stir-bar extraction to detect olfactory faults (Franc, David and De Revel, 2009). Wines that received negative decisions from three or more judges were excluded (25% of the wines) and only those with no significant sensory faults were included. Then, the 20 wine samples included were subjected to the customary physical and chemical analyses: CieLab©, *L**a*b* Color Space CIE, 1976 for color composition; Foss © for ethanol levels, acidity, and pH (Table 1).

**Table 1.** Physical and chemical analyses for all wines (by wine number, type, color composition *L**a*b*, and Foss analyses for ethanol levels, total acidity, and pH)[1-3].

| Wine # | Wine type | L*(D65) (color) | a*(D65) (color) | b*(D65) (color) | Ethanol level | Total Acidity | pH |
|---|---|---|---|---|---|---|---|
| 1 | Premium | 8.6 | 38.4 | 14.7 | 14.42 | 3.36 | 3.60 |
| 2 | Secondary | 19.6 | 48.4 | 32.7 | 13.23 | 2.85 | 3.71 |
| 3 | Premium | 12.9 | 42.4 | 21.9 | 13.41 | 2.98 | 3.62 |
| 4 | Premium | 10.7 | 40.7 | 18.2 | 13.47 | 3.15 | 3.79 |
| 5 | Secondary | 14.8 | 45.5 | 25.2 | 13.28 | 3.30 | 3.61 |
| 6 | Premium | 22.0 | 51.2 | 36.6 | 13.34 | 3.13 | 3.61 |
| 7 | Premium | 20.4 | 51.1 | 34.4 | 13.29 | 3.17 | 3.47 |
| 8 | Secondary | 13.1 | 42.5 | 22.2 | 13.13 | 2.95 | 3.66 |
| 9 | Premium | 14.8 | 45.3 | 25.3 | 13.55 | 3.25 | 3.57 |
| 10 | Premium | 13.7 | 44.8 | 23.5 | 13.47 | 3.29 | 3.63 |
| 11 | Premium | 10.2 | 40.8 | 17.4 | 13.79 | 3.03 | 3.57 |
| 12 | Premium | 16.7 | 47.4 | 28.4 | 13.35 | 3.35 | 3.71 |
| 13 | Premium | 11.0 | 41.3 | 18.7 | 14.61 | 3.30 | 3.52 |
| 14 | Secondary | 16.1 | 46.9 | 27.3 | 12.75 | 3.13 | 3.58 |
| 15 | Premium | 16.2 | 47.2 | 27.4 | 13.63 | 3.29 | 3.52 |
| 16 | Secondary | 16.7 | 47.5 | 28.5 | 13.23 | 2.98 | 3.59 |
| 17 | Premium | 9.4 | 39.4 | 15.9 | 14.34 | 3.72 | 3.70 |
| 18 | Premium | 15.8 | 46.2 | 26.7 | 12.50 | 3.21 | 3.58 |
| 19 | Secondary | 15.3 | 45.9 | 26.0 | 13.74 | 3.10 | 3.60 |
| 20 | Premium | 9.8 | 40.3 | 16.7 | 13.80 | 3.49 | 3.50 |

[1] Repeatability: Expressed by the standard deviation between the measures taken in the same moment on multiple consecutive determinations. [2] *L**a*b* standard deviations calculated at 0.1%. [3] pH: 0.02 pH units; Ethanol: 0.04% vol; Total acidity: 0.014 g/L.

### *2.3. Tasks and Procedure*

The wine tasters participated in six wine-tasting sessions over a span of 5 months. Due to their different time-tables and the tasting room capacity, the tasters were divided into two separate groups for each session. At the start of the first session, the wine tasters were informed about the study context and protocol. They completed a biographical information sheet, as well as a written informed consent sheet. The sessions were conducted in a tasting-room equipped with individual booths. Each session was comprised of two consecutive wine tastings (unconstrained, then constrained), with intervening mental imagery, vocabulary, semantic knowledge, and word frequency tasks (see Table 2 and description below).

**Table 2.** Task schedule over six consecutive tasting sessions (n = number of tasters). The same 14 premium and 6 secondary red wines were used for all tastings.

| Sessions (1 to 6) | 1. Unconstrained tastings | 2. Psychological tasks | 3. Constrained tastings | 4. Word frequency tasks |
|---|---|---|---|---|
| **1** (n=55) | Tasting: 20 wines, all senses Clear wine glasses | Vividness of Visual Imagery Questionnaire (VVIQ) | Tasting: 20 wines, visual unimodal, clear wine glasses, nose clips | Visual descriptors |
| **2** (n=54) | Tasting: 20 wines, all senses Clear wine glasses | Vividness of Olfaction Imagery Questionnaire (VOIQ) | Tasting: 20 wines, smell unimodal, black wine glasses | Olfactory descriptors |
| **3** (n=52) | Tasting: 20 wines, all senses Clear wine glasses | Vividness of Gustatory Imagery Questionnaire (VGIQ) | Tasting: 20 wines, taste unimodal, black wine glasses, nose clips | Gustatory Descriptors |
| **4** (n=47) | Tasting: 20 wines, all senses Clear wine glasses | Vividness of Somesthetic Imagery Questionnaire (VSIQ) | Tasting: 20 wines, visual-smell multimodal, clear wine glasses | Somesthetic descriptors |
| **5** (n=43) | Tasting: 20 wines, all senses Clear wine glasses | Vividness of Wine Imagery Questionnaire (VWIQ I) | Tasting: 20 wines, visual-taste, multimodal, clear wine glasses, nose clips | Global descriptors I (all senses) |
| **6** (n=42) | Tasting: 20 wines, all senses Clear wine glasses | Mill-Hill Vocabulary scale + VWIQ II | Tasting: 20 wines, taste-smell, multimodal, black wine glasses | Global descriptors II (all senses) |

*2.3.1. Unconstrained wine tastings*

Wine tasters started each session by tasting and rating the wines using all sensory cues (i.e., vision, smell, and taste: unconstrained or global wine tasting). The wine tasters served themselves 25mL portions of wine at room temperature, in certified wine glasses (ISO 3591:1977), and in random order. Tasting started immediately after individual servings, according to five unstructured, continuous-response scales: arousal, quality, certainty (of quality), image, and hedonism.

Arousal was defined as the strength of the initial sensory response (low vs. high arousal) to the wine tasted. Arousal was a bi-polarized item, as high arousal marks could indicate either intense negative (unpleasant) or positive (pleasant) sensory feedback, while low arousal indicated weak sensory feedback. Quality was assessed by whether the wine tasted exhibited all the characteristics of a premium wine from the area of interest, based on the wine tasters' own representations of premium wine quality (Ballester *et al.*, 2008: adapted for quality: poor example vs. good example). Wine tasters further rated the certainty of their quality response (uncertain vs. certain). Evaluations of the wine's potential to evoke mental images (low image vs. high image) and pleasure/hedonism (dislike vs. like) were also collated. Image was also a bi-polarized item: high image marks could indicate either negative (unattractive) or positive (attractive) imagination feedback.

Response data (or holistic ratings) corresponded to the measured distance in centimeters (between 0 and 11.3 cm) from the leftmost anchor of the gradient lines to the response markings given by the wine tasters. Unconstrained tastings were generally completed in under thirty minutes. Following the unconstrained tasting, instructions were given for the start of the following task (psychological tasks, see Table 2), which required ten minutes to complete (see below for a description of these tasks).

*2.3.2. Constrained wine tastings*

Constrained tasting procedures were very similar to the unconstrained tastings described above. However, they required particular sensory controls, as well as modified equipment and instructions. Choice of wine glasses (clear vs. opaque black) and the use of approved nose clips were the main variations in the setup (Table 2). Under unimodal smell and taste, as well as bimodal smell-taste conditions, wine tasters used opaque black wine glasses that occluded the wines' visual features (Oberfeld *et al.* 2009). During unimodal visual and taste, as well as bimodal visual-taste conditions, they wore approved nose clips (Fim Medical, Villeurbane, France) to block ortho- and retro-nasal perception (CE, FDA approved, ISO 13485; Sáenz-Navajas *et al.*, 2012). In-mouth tasting was not allowed under unimodal visual and smell or bimodal visual-smell conditions. Constrained tastings were generally completed in under thirty minutes. Following the constrained tasting, instructions were given for the start of the following task (word frequency tasks, see Table 2).

*2.3.3. Vividness of imagery tasks*

Mental imagery tasks were administered between wine tastings (Table 2), including traditional visual and olfactory vividness of imagery tasks (VVIQ: Marks, 1973; VOIQ: Gilbert, Crouch & Kemp, 1998, VVIQ French adaptation, Denis, 1979), as well as newly created tasks assessing wine tasters' self-reported vividness of imagery for gustation (or taste, VGIQ), somesthesia (or body sense, VSIQ), and wine (VWIQ). Standard instructions were given to the wine tasters for each task. They first read stimulus descriptions involving sense/image cues in-context or in familiar situations, as itemized in each of the measurements. Then, using a 5-point Likert response scale, they rated the vividness, or degree of realism, with which their mental representations depicted these specific stimulus descriptions. Imagery scores represent the average responses given by the wine tasters for each of the five measurements.

The VGIQ, VSIQ, and VWIQ imagery tasks were created with a similar response format to that of the VVIQ and VOIQ. Inspiration was also drawn from other validated (multisensory) imagery tasks (e.g., Plymouth Sensory Imagery Questionnaire, see Andrade *et al.*, 2013). Four independent judges, involved in food, wine, and psychological sciences, elaborated the final versions of the tasks. The VGIQ (20 items) assessed the vividness of taste images, with stimulus descriptions integrating taste qualities: acid, bitter, salty, sweet, and umami. The VSIQ (26 items) assessed the vividness of somesthetic images about qualities related to touch senses: temperature, pressure, position, movement, mouthfeel. The VWIQ (33 items) assessed the vividness of wine images with stimulus descriptions used in wine production and tasting: visual, smell, taste, mouthfeel/somesthetic. Careful consideration was given to creating domain-specific items to assess these image cues (The VGIQ, VSIQ, and VWIQ Questionnaires are included in the Supplementary Materials S1). The main purpose was to build a multifaceted descriptor of mental imagery, with components covering distinct aspects of wine tasting, tapping into the wine tasters' tasting representations.

*2.3.4. Vocabulary tasks*

Verbal intelligence was assessed, using the Mill-Hill Vocabulary scale (French adaptation: Deltour, 1993; Raven, Styles, & Raven, 1998), which specifically challenged the wine tasters on the accuracy of their definitions and synonyms (34 items). A separate word-frequency task evaluated the wine tasters' lexical capacities, as well as word consensus, specifically ascribed to premium and secondary wines from the PDO area. The aim was to collect the words that wine tasters used to identify the various sensory aspects related to premium and secondary wines (visual, smell, taste, and somesthetic features). Word frequency tasks were devised for all wine tasting sessions. In free-form tasks, the wine tasters were instructed to write word descriptors they thought adequately described the visual (Session 1), olfactory (Session 2), gustatory (Session 3), and synesthetic (Session 4) features of premium and secondary wines in two boxes. In sessions 5 and 6, wine tasters were instructed to write all the possible words (for all sense cues) to describe the sensory features of premium and secondary wines. The word lists required ten minutes to complete. The main purpose for their inclusion was to build a multifaceted descriptor vocabulary, with components covering distinct aspects of wine tastings, tapping into wine tasters semantic and lexical representations. This provided input for word relatedness analyses for each sense, through word frequencies, co-occurrences (double words), and in-group consensus (Rapp, 2014; Brochet & Dubourdieu, 2001: word relatedness and lexical analysis discussions).

## Data Analyses

The data analyses were programmed using ©*R programming with ©RStudio*. Wine rating responses for two *a priori* defined wine categories (Premium vs. Secondary) were collated in six consecutive sessions (S1 to S6) and converted into percentage ratings. Outlier reduction (±3 SDs) on the total available responses from each tasting session, within subjects, wines, and conditions, eliminated less than 1% outliers (data cells) per scale. Overall, a low propensity for outliers was observed and the response distributions proved acceptable for skewness (- .70 > 0, slightly negatively distributed) and kurtosis indices (within limits ±2: Trochim & Donnelly, 2008; Gravetter, Wallnau & Forzano, 2018). Response distributions were also considered normal (or approx. normal), as mean and median distribution ratios were between .90 and 1 (across scales and sensory conditions).

### *3.1. Expert agreement or consensus on all scales and in all sensory conditions*

Principal component analyses (PCA) were conducted on individual responses on all scales, under all sensory conditions (unconstrained S1 to S6, constrained unimodal and bimodal), to evaluate the level of consensus (or agreement) among wine tasters. In this way, the construct validity of the various response scales was evaluated within the sensory conditions. Although the term 'construct' may not fit exactly in the context of wine tastings, the 20 cued wine stimuli, repeated in all the tastings, were treated as test items. In these analyses, wine tasters translated their own wine ratings into an ordinal quality 'hierarchy' (unbeknownst to them, due to the 'blind' nature of the wine tastings).

### *3.2. Consensus prediction analyses using MLRs.*

Global consensus criteria were subjected to predictive analyses, using age, experience, sensory consensus, vocabulary, and mental imagery as predictors. The response scale criteria were chosen on the basis of preliminary descriptive analyses. Predictive modeling focused on individual factor scores from global quality and image scale ratings, as they were considered non-redundant or independent ($rs \leq .45$, $p < .05$). All other scales were considered for inclusion, although Arousal responses were considered redundant to Image responses, and Hedonism responses were considered redundant to Quality responses ($rs \geq .61$, $ps < .001$). The certainty (of quality) scale also demonstrated the lowest response consensus and poorest relationships to other response scales ($rs \leq .32$, $ps < .001$). Therefore, only the quality (technical) and image (emotional/psychological) scales were deemed suitable for these analyses.

Consensus factor scores were associated with each individual, for each unimodal and bimodal sensory condition PCA (quality and image scales). These observations were calculated as supplementary (illustrative) variables, positioned in the global PCA space. They were projected onto the principal components (Dim1 and Dim2) of the active observations (criteria, unconstrained sensory condition PCAs) to obtain common references in the factor analyses on both the predictive and active levels[1].

The wine tasters included in the predictive analyses completed all 6 sessions, to ensure that individual factor scores (loading on Dimension 1) were obtained for six similar conditions, in both quality and image analyses. Fair reliabilities of .53 and .51 were obtained for quality and image factor scores, respectively. As the analyses covered all of the six unconstrained wine tasting sessions, the six factor scores stacked in the criteria data column (data table) with 132 observational cases. Correspondingly, at discrete temporal intervals (1 month), the criteria were controlled for the progression (and dynamics) ascribed to the longitudinal nature of the study. Unbeknownst to them, each wine taster ranked the wines and repeatedly generated his/her own individual wine rankings during each unconstrained wine tasting. Each individual ranking was given a factor score, as a reduced indicator for the clarity of wine representations, for direct comparison with the average group ranking (under quality or image, for the current purposes). Thus, each individual's wine ranking (hierarchy) was projected onto Dim1 (Cos2), as the distance between 'own clarity' of wine representations and the average group clarity of wine representations. Factor scores or individual projections were between -1 very poor and +1 excellent, thus making it possible to rank each individual in terms of his/her level of belongingness to the group consensus on the wines (quality and image scales: proximity to +1 indicated strong consensus belongingness). These factor scores (Cos2) reflected the importance of a principal component (Dim1) for a given observation (wine taster, wine ratings) on a vector of original variables (wine rating consensus on quality and image; see Abdi & Williams, 2010, for further details on PCAs). All predictors were based on 6 repetitions (132 observational cases), each stacked in a column on the data matrix, to predict the global consensus criteria (S1 to S6) (active consensus, criteria: cglobal).

---

[1] Admittedly, supplementary references showed no difference from own-reference factor scores in the model analyses. This procedure controlled the sensory space referential for the predictive modeling.

*3.3. Sensory rating and consensus predictors*

Factor score variates were obtained for quality and image consensus in visual, smell, taste, visual-smell, visual-taste, and smell-taste ratings. These observations were calculated as supplementary (illustrative) variables, positioned in the global PCA space. They projected onto the principal components (Dim1 and Dim2) of the active observations (criteria, unconstrained sensory condition PCAs). Initially, each variate was submitted to a prediction model (in a stand-alone set of regression analyses to active consensus) to verify potential sensory consensus predictions in the regression models. The list of trialed predictors on both quality and image scales (with descriptive statistics) are included in Appendix A1 (cf. Tables 4 and 5 in the Results section, for the predictors retained in the analyses).

*3.4. Psychological predictors and psychometric assessments*

As already described, the psychological predictors were related to mental imagery and word descriptors for the different wine categories (Premium vs. Secondary). Further biographical information was related to age, gender, education, work occupation, and wine-tasting experience (number of wine tastings/week).

It was critical to meet basic measurement requirements for the new imagery and vocabulary tasks. Test data assessing the level of adequacy of these psychological measurements are reported below. Each type of psychological predictor was further subjected to initial prediction models (in a stand-alone set of regression analyses to active consensus). The aim of this trial approach was to obtain an estimate of the potential level of prediction for each individual predictor. This approach was generalized under the definitions given (cf. Appendix A2, list of predictors and descriptive statistics). Only those predictors with significant prediction levels were retained in the final models.

The imagery questionnaires all showed acceptable internal consistency correlations (α coefficients > .70) and (fair-to) excellent (> .40) item-total correlations or individual discrimination coefficients (VVIQ: .77(α), .37(disc.); VOIQ: .85(α), .50(disc.); VGIQ: .90 (α), .51(disc.); VSIQ: .88(α), .46(disc.); VWIQ: .92 (α), .47(disc.)). They correlated well together (avg. $r$=.53, $ps$ < .01) and quite poorly with vocabulary tasks situated outside their elected measurement concept, word frequency tasks (avg. $r$=.02, $ps$ > .01), and the Mill-Hill vocabulary test (avg. $r$=.04, $ps$ > .01). This suggested good convergent validity of the imagery domain with very good discriminant validity for all imagery measurements (vs. vocabulary).

Similarly, fair/good correlations were observed between the word frequency indicators (avg. $r$=.45, $ps$ < .01). A weaker relationship was observed with the Mill-Hill test (avg. $r$=.18, $ps$ > .05, SD=.15), with acceptable internal consistency (.70 (α)) and lower individual discrimination (.25, perfectly scored items 1, 12, 13, 14, 18, 20, and 21 excluded, Mean percent score = 81.4 ± 3.1%).

Word frequencies were calculated for both wine categories and all sessions (S1 to S6: sum of premium and secondary words in each session). Specific words were evoked for the particular sensory instances, while word descriptor similarities were also observed within and among all the senses (i.e., visual, smell, taste, somesthetic, all senses, see Table 3 below). The word lists submitted by each individual were screened and the entries with the least good fit (< 1%) were discarded. Predictors were thus sensory-specific word frequencies in all sessions (visual, smell, taste, somesthetic, all senses, and all senses repeated) and total word frequency (sum of all words) for premium and secondary wines, respectively.

The proportion of descriptors for premium wines was also calculated (visual, smell, taste, somesthetic, all senses, all senses repeated, and all conditions from S1 to S6 together) from the premium word descriptor ratio at each session. The number of words used to describe both premium and secondary wines (as a redundancy indicator) in each session and sensory-specific domain was also calculated (word doubles for visual, smell, taste, somesthetic, all senses, and all senses repeated). These predicted the verbal representations in the active quality and image wine rating consensus.

Word frequency, like the relative sum of words, was limited, for an understanding of the nature of the words used (or shared) in wine descriptions. For instance, on average, fewer words were reported for Secondary wines (about 40%, see Table 3 below), whereas the level of agreement more accurately informed each category and sensory condition. Within sensory conditions (visual, smell, taste) and wine categories (premium, secondary), the total sum of occurrences of each word in the group was calculated and divided by the total sum of words for each wine category (e.g., group premium word total). Each word was thus associated with a group probability (or percentage) of occurrence. Each probability was matched to the words reported by each individual wine taster. These probabilities were summed, and divided by the number of words reported intra-individually. Thus, individual word consensus indicates the level of agreement reached or probability of agreement for every word captured and reported by an individual. The newly calculated consensus variables thus controlled for word frequencies (visual, smell, taste, somesthetic, all senses, all senses repeated, and all conditions from S1 to S6 together: Premium, Secondary, Total[2]).

All psychological predictors were normally distributed (mean/median ratios between .95 and 1.05). The VWIQ test-retest reliability (wine tasters repeated the VWIQ on two occasions, one month apart) was good (r=.74, $p < .01$). Word frequency test-retest reliability, with the tasters repeating the same word frequency task in sessions 5 and 6, was fair/good (one month interval: r=.59, $p < .01$). Considering these results, the vocabulary tasks were fit for inclusion in the analyses (vocabulary domain). Since the new imagery tasks (VGIQ, VSIQ, and VWIQ) performed similarly to (if not better than) existing tasks (VVIQ and VOIQ), they were also considered fit for inclusion in the analyses (imagery domain).

Finally, several tests were conducted to evaluate the distributional appropriateness of all predictors (and criteria). This was crucial in the preliminary modeling steps, in view of the fact that few wine tasters were included in the predictive analysis model. The data were cleaned for outliers at ±3 SD, generally relatively few in number (overall <1%), which were replaced by within-variable means. After outlier reductions, the data were centered and each predictor was controlled in interactions with session numbers (1 to 6) and wine taster identities (id. Numbers 1 to 22). Centering the variables and controlling for these interaction terms minimized the effects due to stacked repetitions for predictors in the data table.

## Results

### *4.1. Expert agreement or consensus on all scales and under all sensory conditions (all participants)*

Fig. 1 shows bi-plots of the variables (wines) and individuals (wine tasters) tested (see Table 2 for the number of participants included in each sensory condition). Global results are in the left panels, unimodal in the middle panels, and bimodal in the right panels. Bearing in mind that PCAs are not a significance test of difference, levels of disparities for individual wines, tasters, and sensory conditions were still included. In general, secondary wines tended to project to the leftmost of the PCA quadrants, whereas premium wines projected to the rightmost of the quadrants, loosely matching the *a priori* quality rule. This was clearly visible for global (e.g., Global S1, Global S6) and visual conditions, but not as systematic for the smell, taste, and smell-taste conditions. This revealed that wine distinctions were not

[2] Totals were calculated for all variables, collapsing over wine categories premium and secondary.

systematic within wine categories and that some wines tended to deviate from their projected or predicted inclusion category (e.g., Premium wines in the leftmost of the quadrants). Individual projections on the first dimension (Dim1) also suggested much poorer consensus in conditions involving smell and/or taste (≤ 27%), with a considerable inter-individual spread (see arrow displays that show bi-directionality within groups). In contrast, these were much improved when all senses were involved (≥ 42.5%) and even better (to-similar) in instances where vision was involved (Visual=53%, Visual-Taste=38.8%, and Visual-Smell=44.3%). Individuals thus achieved greater consensus in rating the wines when visual cues were available, for quality-control or decision-making.

(Insert Fig. 1 here)

The PCA results for the various sensory conditions are reported in Table 3. A quick inspection of Table 3 convincingly shows better consensus, as a rule, in quality-rating scales (vs. all other scales). However, correlating factor scores between scales suggest a similar relation pattern on all scales. Similar sensory condition effects were found between quality, image, and hedonism scales to those for arousal and certainty scales. Taste (then smell and smell-taste) consistently gave the lowest consensus percentages on PCA Dim1. As a rule, global conditions consistently produced better consensus, followed by bimodal conditions, and, lastly, unimodal conditions.

**Table 3.** Principal Component Analysis (PCAs) results (percentage explained variance) for every response scale and sensory condition (with correlational analysis on factor scores Dim1 between response scales)

| **Conditions** | **Quality** | | **Arousal** | | **Image** | | **Hedonism** | | **Certainty** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Dim1 (%) | Dim2 (%) | Dim1 (%) | Dim2 (%) | Dim1 (%) | Dim2 (%) | Dim1 (%) | Dim2 (%) | Dim1 (%) | Dim2 (%) |
| Global S1 | 46.5 | 7.8 | 28.1 | 10 | 24.3 | 11 | 37.3 | 8.2 | 18 | 11.3 |
| Global S2 | 46.6 | 7.2 | 24.1 | 9.5 | 26.3 | 9.8 | 36.3 | 8 | 16.2 | 12.7 |
| Global S3 | 43.5 | 8.6 | 26.8 | 10.7 | 25.3 | 12.8 | 34.2 | 9 | 21.5 | 10.9 |
| Global S4 | 42.5 | 8.4 | 27.9 | 9.2 | 28 | 11.9 | 36.3 | 9.2 | 23.6 | 11.2 |
| Global S5 | 44.3 | 8.6 | 32.3 | 11.2 | 29.8 | 11.4 | 38.1 | 9.9 | 18.8 | 12.3 |
| Global S6 | 44 | 9.7 | 28.8 | 11.8 | 28.7 | 11.5 | 33.6 | 11.2 | 23.4 | 11.3 |
| Visual | 52.9 | 7.3 | 23.3 | 12 | 26.8 | 11.5 | 42.5 | 8 | 17.6 | 13.6 |
| Smell | 24.1 | 8.9 | 13.6 | 10.7 | 14 | 11.9 | 24.7 | 9.1 | 13.3 | 10.7 |
| Taste | 16 | 11.2 | 14.4 | 12.8 | 12.9 | 10.4 | 14.3 | 10.8 | 14 | 11.3 |
| Visual-Smell | 44.3 | 8.6 | 25.1 | 12.1 | 25.3 | 10.4 | 39.6 | 10.4 | 25.1 | 10.1 |
| Visual-Taste | 38.8 | 10.6 | 29.27 | 10.63 | 25.7 | 12.7 | 28.9 | 10.8 | 21.3 | 12.8 |
| Smell-Taste | 27 | 11 | 21.34 | 12.69 | 15.7 | 12.5 | 25.1 | 11 | 15.7 | 10.6 |
| **Global avg.** | 44.6 | 8.4 | 28 | 10.4 | 27.1 | 11.4 | 35.9 | 9.25 | 20.25 | 11.6 |
| **Uni. avg.** | 31 | 9.1 | 17.1 | 11.8 | 17.9 | 11.3 | 27.2 | 9.3 | 14.9 | 11.8 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Bim. avg.** | 36.7 | 10.1 | 25.23 | 11.8 | 22.2 | 11.8 | 31.2 | 10.73 | 20.7 | 11.2 |

| **Correlations – Factor scores (Dim1) between response scales** | | | | | |
|---|---|---|---|---|---|
| r-Qual. (*ps*<.05) | 1 | .76 | .90 | .96 | .57 |
| r-Arous. (*ps*<.05) | | 1 | .90 | .77 | .70 |
| r-Image. (*ps*<.05) | | | 1 | .85 | .72 |
| r-Hedo. (*ps*<.05) | | | | 1 | .55 |
| r-Cert. (*ps*<.05) | | | | | 1 |

*4.2. Consensus prediction with mental imagery and vocabulary predictors (22 participants)*

Scores for imagery predictors were calculated, as previously described, using the 'self-reported' averages for the VVIQ, VOIQ, VGIQ, VSIQ, VWIQ, VWIQ1, and VWIQ2 repetition. VWIQ scores at times 1 and 2 were averaged into a summarized imagery score, TTVWIQ, since the VWIQs were shown to be redundant. Single ANOVA on 6 Imagery Questionnaire Type, with imagery scores as criteria, showed a significant main effect (22 wine tasters who had completed all sessions, see Fig. 2). Wine tasters generally reported more wine mental images (or more vivid experience representations), on average, compared to those reported for all other imagery tasks, $F(6,126) = 20.12$, $p < .001$, $\eta_p^2 = .49$. Moreover, this effect was distributed evenly on both VWIQ and VWIQ repetitions, thus justifying the use of the combined TTVWIQ average scores in the model analyses.

(Insert Fig. 2 here)

Table 4 shows a list of the top six group consensus words (group probability of occurrence indices or percentages) by sensory condition and wine category (see French words in Appendix A2). Group consensus was clearly very low under all conditions. However, the individuals who selected words that achieved the best consensus in the group also contributed the highest verbal consensus scores.

**Table 4.** Lists of the top six consensus words (%) by sensory condition and wine category (all participants)

| **Visual consensus words (n = 55)** | | | | **Smell consensus words (n = 54)** | | | |
|---|---|---|---|---|---|---|---|
| Premium (459 words) | Group % | Secondary (324 words) | Group % | Premium (670 words) | Group % | Secondary (389 words) | Group % |
| Red | 4.3 | Color | 4.0 | Powerful | 2.8 | Fresh fruit | 3.1 |
| Color | 4.8 | Intense | 4.0 | Fruit | 3.1 | Vanilla | 3.3 |
| Dark | 5.2 | Ruby | 7.7 | Black fruit | 3.1 | Finesse | 3.8 |
| Brilliant | 7.4 | Red | 7.7 | Vanilla | 3.3 | Oaky | 4.6 |
| **Deep** | **7.8** | **Limpid** | **8.3** | **Complexity** | **4.1** | **Red fruit** | **6.4** |
| **Intense** | **7.8** | **Brilliant** | **12.3** | **Oaky** | **5.1** | **Fruity** | **9.0** |
| **Taste consensus words (n = 52)** | | | | **Somesthetic consensus words (n = 47)** | | | |
| Premium (662 words) | Group % | Secondary (423 words) | Group % | Premium (429 words) | Group % | Secondary (229 words) | Group % |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Powerful | 3.2 | Finesse | 4.3 | Rich | 3.6 | Fresh | 3.9 |
| Structure | 3.2 | Tannins | 4.7 | Hot | 3.9 | Soft | 3.9 |
| Oaky | 3.3 | Round | 5.4 | Round | 4.9 | Tannins | 5.2 |
| Long | 3.9 | Fruity | 5.9 | Velvety | 5.1 | Round | 5.7 |
| **Balanced** | **4.7** | **Supple** | **6.1** | **Silky** | **5.6** | **Silky** | **5.7** |
| **Tannins** | **5.4** | **Balanced** | **8.0** | **Tannins** | **5.6** | **Supple** | **6.1** |
| **All senses consensus words (n = 43)** | | | | **All senses rep. consensus words (n = 42)** | | | |
| Premium (775 words) | Group % | Secondary (448 words) | Group % | Premium (668 words) | Group % | Secondary (389 words) | Group % |
| Deep | 2.3 | Fresh | 2.9 | Powerful | 3.0 | Fresh | 3.1 |
| Oaky | 2.8 | Finesse | 3.3 | Complex | 3.3 | Round | 3.1 |
| Balanced | 3.0 | Supple | 3.8 | Balanced | 3.3 | Tannins | 3.1 |
| Long | 3.0 | Red fruit | 4.2 | Long | 3.3 | Supple | 3.6 |
| **Complex** | **3.2** | **Balanced** | **4.9** | **Tannins** | **3.4** | **Balanced** | **5.1** |
| **Tannins** | **3.4** | **Fruity** | **6.3** | **Woody** | **3.7** | **Fruity** | **6.7** |

*Note.* Words in Table 3 are confirmed English translations of the French words originally collected in this study (cf. Paul Cadiau, Lexivin®, 2014).

ANOVA on 6 Sensory Conditions and 2 Wine Categories (Premium vs. Secondary), with verbal consensus scores as the dependent variable, showed significant effects of both variables (for each of the 22 wine tasters who completed all sessions, see Fig. 3). Secondary wines generally captured fewer words, but generated better individual/group consensus scores, $F(1,21) = 51.69$, $p < .001$, $\eta_p^2 = .71$. Visual word consensus scores were also higher, $F(5,105) = 49.72$, $p < .001$, $\eta_p^2 = .70$, compared with all other sensory conditions (consensus indices: 4.5 vs. 2.3).

(Insert Fig. 3 here)

*4.3. Final regression models: Quality and image rating consensus (22 participants)*

In the stepwise regression approach, the inclusion possibility of every predictor for explaining consensus in global quality and image ratings was evaluated. Several different associations of variables were trialed in the models and inclusion decisions were gradually confirmed. This involved selecting the most potent, non-redundant predictors (in the sensory, vocabulary, and imagery groups), proceeding by elimination in stepwise regression analyses. Similar approaches were applied to both quality and image analyses. Only predictors representing a good prediction of the wine rating consensus (quality and image) were retained in the final model.

The final quality rating consensus model (red wine) is shown in Table 5. Regression analyses revealed specific drivers based on age, word consensus scores, total word frequency over all sessions, proportion of premium words over all sessions (negative weight[3]), vividness of wine mental imagery, as well as unisensory smell consensus, and, to a lesser degree, unisensory taste and multisensory visual-taste consensus. Common experiences with wine (consensus words and images, consensus senses), as well as the number of years tasting (age, longevity), were predictors of wine tasters' inclusion in the red wine rating consensus. Both word consensus and proportion suggested that the precision of words over wine descriptions was a very important predictor of consensus (not just total word frequencies). Individual differences in smell consensus substantiated a good overall prediction, as did taste consensus and partitioned visual influence, in combination with taste. Taken together, these results revealed a prediction of consensus on red wine quality ratings based equally on perceptual (31.8%) and psychological prototypes/experience (31.3%).

The final image rating consensus model (red wine) is shown in Table 6. The regression analyses revealed specific drivers based on vividness of taste imagery (negative weight[4]), visual consensus for secondary wine descriptors/words, proportion of double visual descriptors/words for premium and secondary wines (negative weight), proportion of double somesthetic descriptors/words for premium and secondary wines (negative weight), total visual word frequency for premium wines, visual-smell consensus, and, to a lesser degree, unisensory smell consensus (negative weight). First suggestions: negative weights offset the consensus inclusion for image ratings. This was the case for smell consensus prediction, vividness of taste imagery, as well as double words in the visual and somesthetic domains. Strong consensus drivers were obtained for visual vocabulary consensus and word frequencies. Specific visual representations (word/descriptor knowledge) thus contributed to consensus on image ratings in global tastings. Poor word efficiency (double words) in both visual and somesthetic conditions was attributed to wine tasters outside the wine image consensus. Smell image consensus and taste vividness predicted a lack of consensus in global tasting image ratings. While visual-smell was included in the prediction, the role of smell was subordinate to that of vision. Taken together, our results revealed a prediction of mechanisms for consensus about red wine image ratings based equally on perceptual (34%) and psychological prototypes (33%).

[3] Unless otherwise noted (as negative), the reported regression weights are positive.

**Table 5.** Final quality rating consensus model.

| **Quality Consensus** | ***F-statistics*** | Spec. Predictors | **β Est.** | ***t*-values** | ***p*-values (*t*)** | **adj. R²** |
|---|---|---|---|---|---|---|
| Quality | 25.81 | Longevity exp./Age | .00 | 5.68 | < .001 | |
| Rating | *df*(8,121) | Consensus words | .06 | 4.55 | < .001 | .61*** |
| Consensus | | Total word freq. | .00 | 5.90 | <.001 | Psych .31 |
| | | Prop. Prem words | -.01 | -6.98 | < .001 | Sens. .30 |
| (all sessions, | | Wine imagery | .26 | 5.66 | < .001 | |
| n = 22) | | Consensus taste | .32 | 2.93 | = .004 | |
| | | Consensus smell | .27 | 4.88 | < .001 | |
| | | Consensus vis-tast. | .21 | 2.63 | = .009 | |

*Note.* Final quality rating consensus model for red wines. Each model was assessed on 22 participants (under test conditions). Model *F*-statistics and adjusted-$R^2$ render the global significance of the models tested, with beta estimate coefficients (β Est.) and their associated *t*-values and *t*-value significance. *** $p < .001$. Total psychological and sensory predictions are each presented under the adjusted-$R^2$ analysis. Regression criteria: quality consensus.

**Table 6.** Final image rating consensus model.

| **Image Consensus** | ***F-statistics*** | Spec. Predictors | **β Est.** | ***t*-values** | ***p*-values (*t*)** | **adj. R²** |
|---|---|---|---|---|---|---|
| Image Rating | 37.54 | Taste imagery | -.18 | -5.62 | = .030 | |
| Consensus | *df*(7,124) | Vis. Cons. Words (Secondary) | .04 | 6.43 | < .001 | .66*** Psych .33 |
| (all sessions, n = 22) | | Visual double words | -.00 | -6.40 | <.001 | Sens. .34 |
| | | Somesthetic double words | -.00 | -3.71 | = .404 | |
| | | Visual word freq. (Premium) | .02 | 4.93 | < .001 | |
| | | Smell consensus | -.18 | -2.28 | < .001 | |
| | | Vis-sm. Consensus | .34 | 4.80 | < .001 | |

*Note.* Final image rating consensus model for red wines. Each model was assessed on 22 participants (under test conditions). Model *F*-statistics and adjusted-$R^2$ render the global significance of the models tested, with beta estimate coefficients (β Est.) and their associated *t*-values and *t*-value significance. *** $p < .001$. Total psychological and sensory predictions are each presented under the adjusted-$R^2$ analysis. Regression criteria: image consensus.

## Discussion

The objective of this study was to gain a better understanding of the sensory and psychological representations that predict decisional consensus in expert wine assessments. The results gave a clearer picture of the decisional processes for red wines. The main results discussed pertain to decisional consensus analyses, by sensory condition, and consensus decision modeling, based on both sensory and psychological variables.

### *5.1. Principal component analyses: wine category, sensory stimuli, and individual differences*

The wine quality responses were depicted with quadrant PCAs. This level of analysis rendered a bi-plot distribution of the wines, in relation to premium and secondary wine groups, defined *a priori*. This was helpful in fitting the wines within each of their projected categories (Abdi & Williams, 2010) and comparing them delimitations under different sensory conditions. In general, wines in the same category (premium vs. secondary) fit the profile predictions for the quality ratings. However, a closer examination of the density plots revealed clearer distinctions in the visual and smell conditions, with secondary wines clearly projecting to the left and premium wines to the right, compared to a confused projection under taste conditions.

In some cases, *a priori* premium wines did not meet the criteria, while some secondary wines did. As previously reported (Ballester *et al*., 2005, p. 351), "Mere membership in a pre-defined category is not enough to ensure that the chosen [wines] possess the typical characteristics […] of that category". Rosch and Mervis (1975) and Mervis and Rosch (1981) stated that the members of a category shared common key attributes, different from those shared by members of another category. Within categories, all wines do not possess every key attribute, but they project on a continuum, which includes "fitter" wines, more typical of the winemaking region under investigation, compared to other wines.

The results of this study matched the consensus for visual, smell, and taste decisions reported in a previous study by Sáenz-Navajas and colleagues (2016b). Individual projections suggested less good agreement for quality decisions in smell (24.1%), taste (16%), and smell-taste (27%), compared to vision (52.9%) and all other sensory conditions (>38.8%). The strong intra- and inter-sense differential points to more acute quality concept disunities when smell and taste are involved. The senses thus offer a specific set of input possibilities, with potential individual differences. The data presented here support more marked individual differences when both smell (Criado et al., 2019; Holley & MacLeod, 1977; Tempere et al., 2011) and taste (Bajec & Pickering, 2008) are involved, compared to those involving vision (Herz & Engen, 1996). This order of differences between the senses was obtained in all scale ratings. However, a gradual trend towards decreasing consensus from higher to lower quality in hedonic, arousal, image, and certainty (of quality) was observed across all the senses.

When comparing the average degree of decisional consensus between sensory conditions, global conditions surpassed bimodal conditions, followed by unimodal conditions. Weaker consensus was observed under sensory controls, suggesting that perceptual variations drove poorer decisional consensus. This order of differences between the senses pointed to sensory distinctions, as well as highlighting the advantage of routine tasting, involving all the senses. Wine tasters were more consistent in usual tasting contexts and technical quality evaluations. A number of interpretations of individual differences in wine assessments have been offered in recent literature. Decisional differences may be driven by differing cognitive strategies, as well as genetic and physiological factors (Tempere *et al.*, 2014a; Tempere *et al.*, 2011). Inter-individual perceptual differences may drive poorer sensory consensus in a similar interpretation (Charters & Pettigrew, 2006; Peynaud, 1987), directly affecting sensory information uptake and, thus, the decisional processes.

The results observed for technical responses emphasized that the consensus between wine tasters was stronger when faced with fewer simultaneous response options. In comparison, emotional decisions probably depended on a wider scope of individual differences and personal interpretations (Engen, 1988; see Charters & Pettigrew, 2006 for a relevant discussion on wine rating consensus, p. 629). In this sense, there is no strong guiding principle in wine-tasting decisions, as the rather low sensory consensus levels predict (< 53%). Furthermore, there is probably no single, established decisional rule, so full consensus at wine tastings is highly unlikely. Traditionally, consensus has not been considered a prerequisite for success in this context. A wine taster who does not fit the consensus may have his/her own valid impression of a wine, especially given his/her level of experience with the wines being tasted (Peynaud, 1987). From that standpoint, a study aiming at a better understanding of consensus should also, logically, consider a slightly different question: What are the sensory and psychological determinants of decisional variances among wine tasters during wine tastings? This question was further addressed using statistical modeling (see next section 5.2.).

In the PCA analyses, with individual responses as variables, the sensory response consensus varied from one bottle to the next. The consensus hierarchy of the senses thus depends on multiple sensory drivers, which are probably object- (i.e, bottle) and not only set-dependent (i.e., premium vs. secondary). For instance, correspondences between scale responses and sensory definitions depend on the level of distinctions perceived intrinsically for each wine (visual, smell, and taste). In this study, the wine selection methods may also have affected the level of distinction available to each sense, as specific controls were given for visual coordinates (normally distributed color coordinates), as well as smell and taste distinctions (sensory faults: yes/no). In this way, the wine set may have afforded more quality distinctions for visual processing, compared to the other senses. This approach thus has its limitations, as the choice of wines probably modulates the wine affordances and sensory distinctions for agreement. Thus, changing the wine potentially changes the sensory analysis possibilities.

The participant outcome of this study deserves further consideration. Participant attrition varied markedly between tasting sessions, with a varying number of wine tasters available. Consequently, the study program may have induced tasting experience inequalities among participants, since each participated in a varying number of tastings. Similarly, the sensory conditions may not be perfectly comparable. However, by analyzing the group response patterns, the results suggested quite similar response trends across sensory conditions, with the exception of taste. Fair reliability was also observed over the six global wine tasting sessions. Consequently, careful consideration was also given to the modeling analyses, as discussed in the next section (5.2.). Notably, the modeling analyses controlled for the number of sessions (and each individual case), also focusing on the tasters who participated in every session (22 participants). This was to avoid possible confusion in interpreting the results due to this issue.

### *5.2. Predicting the decisional consensus: sensory, vocabulary, and mental imagery*

This section discusses the statistical models of the red wine decisional consensus in greater detail. The results showed that both representational and sensory factors explained a significant proportion of the decisional consensus. The wine tasters probably identified wine attributes in a variable manner, temporally (sequence), quantitatively (sensory thresholds), and qualitatively (cognitive mechanisms). As previously discussed, consensus varied depending on the sense modality and the type of scale, which is also a good indication of the involvement of both representational and sensory influences in the decisional outcomes. These variations informed the consensus prediction analyses. Throughout the tastings, each wine taster gradually constructed his/her own latent set of wine quality and image hierarchies. These hierarchies were reduced to individual factor scores, used in the predictions. They defined each individual’s clarity of wine representations, according to their distance from the group consensus.

Two response models were analysed for the decisional outcomes on the quality and image scales. The wine tasters' consensus levels varied substantially under certain sensory and representational conditions. For instance, very low consensus levels were observed under taste conditions, as well as for the words describing taste inputs in the word exercises. In contrast, consensus levels were significantly higher under visual tasting conditions and during visual word exercises. Notably, the wine experts reported more vivid imagery for wines, compared to the average vividness for visual, smell, taste, and somesthetic images. These results, in close correspondence to a recent study by Croijmans and collaborators (2020; see also Croijmans, Speed, Arshamian, & Majid, 2019), suggest that expertise shapes multimodal imagery for wine with repeated experience, since wine images were reported as more vivid compared to images in other, unrelated sensory contexts. Greater consensus was also achieved for words in secondary wine descriptions, suggesting a better availability of quality distinctions and clearer word representations for lower-quality wines. Premium wines, which are presumably better quality, may be more difficult to describe, since the wine tasters are more experienced at judging wines by their faults or negative qualities, rather than their positive qualities (Jackson, 2009; Peynaud, 1987).

During wine tastings, the wine tasters conceivably uploaded sensory inputs, under *a priori* expectations, knowledge, and psychological representations of wines (Brochet & Dubourdieu, 2001), as well as mental standards (Costell, 2002). The senses may work collaboratively with task-ready or useful mental representations. In this sense, prior representations or images are useful, since they activate quite similar brain locations that contribute to the task being performed (Kosslyn, Ganis, & Thompson, 2001; Plailly, Delon-Martin, & Royet, 2012). Wine tasting also requires working memory and executive functions in tasting and decision making (Alba & Hutchinson, 1987; Castriota-Scanderberg *et al.*, 2005). This is further helped by organized knowledge and experience (Hughson & Boakes, 2002), a clearly-defined wine vocabulary (Gawel, 1997; Lehrer, 1975; Melcher & Schooler, 1996), and common wine images or representational prototypes (Torri *et al.*, 2013). The search starts with sensory data, as tasters activate their wine-tasting sequence, which is further influenced by cognitive mechanisms during progression towards a decision on wine quality.

In the quality model, extensive experience with wine, as well as the number of years tasting, promoted consensus belongingness in global wine decisions. This suggested that common quality representations of the wines were shared by the tasters. Knowledge of descriptor words was also decisive, as predicted by the proportion of premium and secondary words, total word frequency, and total word consensus scores. On the one hand, representational variates were determinant. On the other hand, skilled nose, mouth, and eyes also contributed to consensus belongingness. Decisional processes called equally upon both the psychological (31.3%) and sensory predictors (senses: 31.8%). In this process, the senses represent skill and knowledge depicts the representations as drivers of consensus (or individual differences). These may be particularly influential in substantiating the guiding principle: what tasting factors to consider in rating wine quality. In the case of expert wine tasters, sensory abilities and semantic representations are often thought to develop collaboratively in an experiential nexus (Melcher & Schooler, 1996; Parr, Hearterbell & White, 2002). Training and expertise in wine tasting enhance the taster's analytical abilities (see Arvisenet, Guichard & Ballester, 2016) and multimodal, mental wine imagery (Croijmans *et al*., 2020; Croijmans *et al*., 2019). Training should thus involve word exercises, with wine tasters continuously seeking precision in regional descriptors and representations. Mental imagery certainly offers an interesting avenue for expertise training (Croijmans *et al*., 2020; Tempere *et al*., 2014b).

In terms of global decisions on wine quality, smell was the key sensory predictor of inclusion in a decisional consensus. The lack of inclusion of visual consensus points to a distributional effect. These results suggest a more common visual assessment of color in judging quality (e.g., dark vs. light red). Vision may act as the gatekeeper, letting in the first impression of the wine (Delwiche, 2012; Spence, 2010). There is a well-established visual color bias, as visual analysis often precedes further steps involving the other senses in wine assessments (Morrot, Brochet & Dubourdieu, 2001). However, the involvement of all the senses opens the way for additional signals. A broader level of sensory skills then influences wine decisions (skilled nose, tasting). An experienced taster obtains an initial, visual impression, potentially signals unappealing smells, and decides whether the eyes were right, indeterminate, or wrong. He/she then decides whether the wine is a good exemplar of the premium quality category or not. Taste representations may only provide confirmation, in combination with the representations available from the other senses (e.g., taste balance). In contrast, a less-experienced taster may favor a ‘lesser’ wine, rendering imbalance and incompatibility in rating the quality categories. However, compatible signals should be easier to integrate with experience, compared to incompatible signals (Prescott, 2012). Similarly informed (in-)compatibilities between the senses should thus predict consensus quality decisions. A comprehensive approach for wine tasters would, thus, include sensory and representational balance, for a clearer quality assessment in the final decisional ratings.

The quality and image models produced a distinctive set of significant predictors that may be considered domain-specific. For the image model, the results again revealed an equal relationship between the sensory (34%) and psychological predictors (33%). Visual word variables linked positively, whereas word redundancy (poor lexical content) linked negatively to image consensus. Taste imagery prediction also had negative weight. In this way, participants reporting vivid taste images with less diversified word descriptors exhibited lower consensus belongingness for the image ratings. This negative relationship also applied to smell consensus, whereas visual influences (sensory, word knowledge, quality descriptors) dominated the image consensus. These results suggest that active common representations are both visually- and verbally-driven, at least for this wine set and taster population. This further suggests a model based on both depictive (image) and descriptive (word) representations (Anderson, 1979; Kosslyn et al., 2001; Pylyshyn, 2003). This level of results also indicates that a stable wine-tasting approach requires the comprehensive use of sensory and representational content.

*5.3. Conclusion*

This study revealed better decisional consensus under visual sensory conditions, compared to all other conditions (all senses, smell, taste). Smell was the most important sensory predictor for consensus in quality decisions (together with wine odors). In contrast, the most important sensory predictor for consensus in imagery decisions was vision. Consensus in quality and image decisions also depended on equal predictions from both sensory and psychological representations. Knowledge shapes decisiveness in wine assessment and this relates directly to the number and nature of the sensory inputs available at tasting (visual, smell, taste). Further knowledge and vividness of imagery representations are derived from the regional wine experience and training of the wine tasters. Consequently, perceptual mechanisms and cognitive processes govern wine distinctions and decisions (i.e., what to look for in relation to quality and image representational prototypes). The senses thus integrate as basic information pickup (in bottom-up processing) and skill anchors (in top-down processing), closely related to the wine representations used in decision-making by wine tasters in a particular wine production area. Further research in this direction will help to delineate expert wine tasting decisions and promote informed professional training in this field.

**Acknowledgments**

The authors would like to acknowledge the support received from *La Région Nouvelle-Aquitaine*, *Le Conseil Interprofessionnel du Vin de Bordeaux (Bordeaux Wine Council)* and *Le Conseil des Vins de Saint-Emilion.*

## References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(4), 433-459.

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of Consumer Expertise. *Journal of Consumer Research, 13*(4), 411.

Anderson, J. R. (1979). Further arguments concerning representations for mental imagery: A response to Hayes-Roth and Pylyshyn. *Psychological Review, 86*(4), 395-406.

Andrade, J., May, J., Deeprose, C., Baugh, S., & Ganis, G. (2013). Assessing vividness of mental imagery: The Plymouth Sensory Imagery Questionnaire. *British Journal of Psychology, 105*(4), 547-563.

Arshamian, A., & Larsson, M. (2014). Same same but different: The case of olfactory imagery. *Frontiers in Psychology, 5*, 1-8.

Arvisenet, G., Guichard, E., & Ballester, J. (2016). Taste-aroma interaction in model wines: Effect of training and expertise. *Food Quality and Preference, 52*, 211-221.

Ballester, J., Dacremont, C., Fur, Y. L., & Etiévant, P. (2005). The role of olfaction in the elaboration and use of the Chardonnay wine concept. *Food Quality and Preference, 16*(4), 351-359.

Ballester, J., Patris, B., Symoneaux, R., & Valentin, D. (2008). Conceptual vs. perceptual wine spaces: Does expertise matter? *Food Quality and Preference, 19*(3), 267-276.

Brochet, F., & Dubourdieu, D. (2001). Wine Descriptive Language Supports Cognitive Specificity of Chemical Senses. *Brain and Language, 77*(2), 187-196.

Cadiau, P. (2014). *Lexivin*. Pernand-Vergelesses: Paul Cadiau.

Castriota-Scanderbeg, A., Hagberg, G. E., Cerasa, A., Committeri, G., Galati, G., Patria, F., Frackowiak, R. (2005). The appreciation of wine by sommeliers: A functional magnetic resonance study of sensory integration. *NeuroImage*, *25*(2), 570-578.

Charters, S., & Pettigrew, S. (2006). Conceptualizing product quality: The case of wine. *Marketing Theory*, *6*(4), 467-483.

Costell, E. (2002). A comparison of sensory methods in quality control. *Food quality and Preference*, 13(6), 341-353.

Croijmans, I., Speed, L., Arshamian, A., & Majid, A. (2020). Expertise Shapes Multimodal Imagery for Wine. *Cognitive Science*, *44*(5).

Croijmans, I., Speed, L., Arshamian, A., & Majid, A. (2019). Measuring Multisensory Imagery of Wine: the Vividness of Wine Imagery Questionnaire. *Multisensory Research*, *32*(3), 179-195.

Deltour, J. J. (1993). Echelle de vocabulaire de Mill Hill de J. C. Raven. Braine-le-Chateau, Belgium: éditions l'Application des Techniques Modernes SPRL.

Delwiche, J. F. (2012). You eat with your eyes first. *Physiology & behavior*, *107*(4), 502-504.

Denis, M. (1979). Les images mentales. Paris, Presses Universitaires de France.

Engen, T. (1988). The Acquisition of Odour Hedonics. In: *Perfumery: The Psychology and Biology of Fragrance*, S. van Toller and G. H. Dodd, eds., pp. 79-90, London: Chapman Hall.

Etaio, I., Albisu, M., Ojeda, M., Gil, P., Salmerón, J., & Elortondo, F. P. (2010). Sensory quality control for food certification: A case study on wine. Method development. *Food Control*, *21*(4), 533-541.

Franc C., David F., De Revel, G. (2009). Multi-residue off-flavour profiling in wine using stir bar sorptive extraction-thermal desorption-gas chromatography-mass spectrometry. *Journal of Chromatography A*, *1216*(15), 3318-3327.

Gawel, R. (1997). The Use Of Language By Trained And Untrained Experienced Wine Tasters. *Journal of Sensory Studies*, *12*(4), 267-284.

Gawel, R., & Godden, P. W. (2008). Evaluation of the consistency of wine quality assessments from expert wine tasters. *Australian Journal of Grape and Wine Research*, *14*(1), 1-8.

Gilbert, A. N., Crouch, M., & Kemp, S. E. (1998). Olfactory and visual mental imagery. *Journal of Mental Imagery, 22*(3-4), 137-146.

Gravetter, F. J., Wallnau, L. B., & Forzano, L. B. (2018). *Essentials of Statistics for the Behavioral Sciences*. Boston: Cengage Learning.

Herz, R. S., & Engen, T. (1996). Odor memory: Review and analysis. *Psychonomic Bulletin & Review*, *3*(3), 300-313.

Herz, R. S., & Schooler, J. W. (2002). A naturalistic study of autobiographical memories evoked by olfactory and visual cues: testing the Proustian hypothesis. *American Journal of Psychology*, *115*(1), 21-32.

Holley, A., & Mcleod, P. (1977). Transduction et codage des informations olfactives chez les vertébrés.

Hopfer, H., & Heymann, H. (2014). Judging wine quality: Do we need experts, consumers or trained panelists? *Food Quality and Preference*, *32*, 221-233.

Hughson, A. L., & Boakes, R. A. (2001). Perceptual and cognitive aspects of wine expertise. *Australian Journal of Psychology*, *53*(2), 103-108.

Hughson, A. L., & Boakes, R. A. (2002). The knowing nose: The role of knowledge in wine expertise. *Food Quality and Preference*, *13*(7-8), 463-472.

Hughson, A. L., & Boakes, R. A. (2009). Short Article: Passive perceptual learning in relation to wine: Short-term recognition and verbal description. *Quarterly Journal of Experimental Psychology*, *62*(1), 1-8.

ISO 3591 (1977). Sensory analysis. Sensory analysis apparatus. Specification for winetasting glass.

Jackson, R. S. (2009). *Wine tasting: A professional handbook*. Amsterdam: Elsevier Academic Press.

Kobayashi, M., Takeda, M., Hattori, N., Fukunaga, M., Sasabe, T., Inoue, N., Watanabe, Y. (2004). Functional imaging of gustatory perception and imagery: "top-down" processing of gustatory signals. *NeuroImage*, *23*(4), 1271-1282.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*(9), 635-642.

Lawless, H. T. (1984). Flavor description of white wine by "expert" and nonexpert wine consumers. *Journal of Food Science*, *49*(1), 120-123.

Lehrer, A. (1975). Talking about Wine. *Language*, *51*(4), 901-923.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17-24.

Melcher, J. M., & Schooler, J. W. (1996). The Misremembrance of Wines Past: Verbal and Perceptual Expertise Differentially Mediate Verbal Overshadowing of Taste Memory. *Journal of Memory and Language*, *35*(2), 231-245.

Mervis, C. B., & Rosch, E. (1981). Categorization of Natural Objects. *Annual Review of Psychology*, *32*(1), 89-115.

Newell, A. and Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In: *Cognitive skills and their acquisition*, Anderson, J. R., ed., pp. 1–55, Lawrence Erlbaum Associates, Hillsdale, NJ.

Oberfeld, D., Hecht, H., Allendorf, U., & Wickelmaier, F. (2009). Ambient Lighting Modifies The Flavor Of Wine. *Journal of Sensory Studies*, *24*(6), 797-832.

Parr, W. V., Heatherbell, D., & White, K. G. (2002). Demystifying wine expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical senses*, 27(8), 747-755.

Parr, W., Mouret, M., Blackmore, S., Pelquest-Hunt, T., & Urdapilleta, I. (2011). Representation of complexity in wine: Influence of expertise. *Food Quality and Preference*, *22*(7), 647-660.

Peynaud, E. (1987). *The taste of wine: The art and science of wine appreciation*. London: Macdonald Orbis.

Plailly, J., Delon-Martin, C., & Royet, J. P. (2012). Experience induces functional reorganization in brain regions involved in odor imagery in perfumers. Human brain mapping, 33(1), 224-234.

Prescott J. (2012). Multimodal chemosensory interactions and perception of flavor. In: *The Neural Bases of Multisensory Processes*, Murray MM, Wallace MT, eds. Boca Raton (FL): CRC Press/Taylor & Francis.

Pylyshyn, Z. (2003). Return of the mental image: Are there really pictures in the brain? *Trends in Cognitive Sciences*, *7*(3), 113-118.

Rapp R. (2014) Using word association norms to measure corpus representativeness. In: *Computational Linguistics and Intelligent Text Processing*, Gelbukh A., eds. Lecture Notes in Computer Science, *8403*. Springer, Berlin, Heidelberg.

Raven, J. C., Styles, I., & Raven, M. A. (1998). Raven’s Progressive Matrices: SPM plus test booklet. Oxford, England: Oxford Psychologists Press/San Antonio, TX: The Psychological Corporation.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive psychology*, *7*(4), 573-605.

Sáenz-Navajas, M., Avizcuri, J., Ballester, J., Fernández-Zurbano, P., Ferreira, V., Peyron, D., & Valentin, D. (2015). Sensory-active compounds influencing wine experts and consumers perception of red wine intrinsic quality. *LWT - Food Science and Technology*, *60*(1), 400-411.

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

Sáenz-Navajas, M., Ballester, J., Fernández-Zurbano, P., Ferreira, V., Peyron, D., & Valentin, D. (2016a). Wine quality perception: A sensory point of view. *Wine Safety, Consumer Preference, and Human Health*, 119-138.

Sáenz-Navajas, M., Campo, E., Avizcuri, J. M., Valentin, D., Fernández-Zurbano, P., & Ferreira, V. (2012). Contribution of non-volatile and aroma fractions to in-mouth sensory properties of red wines: Wine reconstitution strategies and sensory sorting task. *Analytica Chimica Acta*, *732*, 64-72.

Sauvageot, F., Urdapilleta, I., & Peyron, D. (2006). Within and between variations of texts elicited from nine wine experts. *Food Quality and Preference*, *17*(6), 429-444.

Schifferstein, H. N. J. (2006). The relative importance of sensory modalities in product usage: a study of self-reports. *Acta Psychologica*, *118*, 293-318.

Schifferstein, H. N., & Desmet, P. M. (2008). Tools Facilitating Multi-sensory Product Design. *The Design Journal*, *11*(2), 137-158.

Schifferstein, H. N. J., & Spence, C. (2008). Multisensory product experience, 133–161. In: *Product Experience*, Hendrik N. J. Schifferstein, Paul Hekkert, eds., pp. 133-161, Elsevier Science.

Solomon, G. E. (1990). Psychology of Novice and Expert Wine Talk. *The American Journal of Psychology*, *103*(4), 495-417.

Spence, C. (2010). The color of wine, part 2. *World Fine Wine*, *29*, 112-119.

Tempere, S., Cuzange, E., Malak, J., Bougeant, J. C., de Revel, G., & Sicard, G. (2011). The training level of experts influences their detection thresholds for key wine compounds. *Chemosensory perception*, *4*(3), 99.

Tempere, S., Hamtat, M., Bougeant, J., Revel, G. D., & Sicard, G. (2014a). Learning Odors: The Impact of Visual and Olfactory Mental Imagery Training on Odor Perception. *Journal of Sensory Studies*, *29*(6), 435-449.

Tempere, S., Cuzange, E., Schaaper, M. H., De Lescar, R., De Revel, G., & Sicard, G. (2014b). "Brett character" in wine: Is there a consensus among professional assessors? A perceptual and conceptual approach. Food quality and preference, 34, 29-36.

Torri, L., Dinnella, C., Recchia, A., Naes, T., Tuorila, H., & Monteleone, E. (2013). Projective Mapping for interpreting wine aroma differences as perceived by naïve and experienced assessors. *Food Quality and Preference*, *29*(1), 6-15.

Trochim, W. M., & Donnelly, J. P. (2008). *Research methods knowledge base*. Mason, OH: Atomic Dog/Cengage Learning.

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

Varela, P., & Gambaro, A. (2006). Sensory Descriptive Analysis Of Uruguayan Tannat Wine: Correlation To Quality Assessment. *Journal of Sensory Studies*, *21*(2), 203-217.

World Medical Association Declaration of Helsinki. (2013). *Jama, 310*(20), 2191.

# Appendix

A1. Criteria and predictors in Quality consensus and Image consensus models

| Criteria and Predictors | Model | Mean | SD |
|---|---|---|---|
| **Quality Criteria Consensus** | Quality | 0.67 | 0.18 |
| **Image Criteria Consensus** | Image | 0.25 | 0.21 |
| **Gender (Biographical)** | both | 32% Female | |
| **Age (Biographical/experience)** | both | 49.15 | 9.25 |
| **Visual Consensus** | Quality | 0.63 | 0.25 |
| | Image | 0.21 | 0.19 |
| **Smell Consensus** | Quality | 0.33 | 0.22 |
| | Image | 0.17 | 0.16 |
| **Taste Consensus** | Quality | 0.24 | 0.18 |
| | Image | 0.09 | 0.09 |
| **Smell-Visual Consensus** | Quality | 0.62 | 0.21 |
| | Image | 0.22 | 0.21 |
| **Visual-Taste Consensus** | Quality | 0.48 | 0.23 |
| | Image | 0.28 | 0.24 |
| **Smell-Taste Consensus** | Quality | 0.45 | 0.23 |
| | Image | 0.15 | 0.12 |
| **VVIQ, self-reported visual imagery** | both | 3.67 | 0.42 |
| **VOIQ, self-reported smell imagery** | both | 3.77 | 0.42 |
| **VGIQ (new), self-reported taste imagery** | both | 3.68 | 0.40 |
| **VSIQ (new), self-reported somesthetic imagery** | both | 3.76 | 0.37 |
| **VWIQ (new), self-reported wine imagery** | both | 4.20 | 0.28 |
| **TTVWIQ, average of 2 VWIQ repetitions** | both | 4.23 | 0.27 |
| **Mill Hill vocabulary, total verbal intelligence** | both | 27.67 | 3.05 |
| **Word frequency total, premium wines** | both | 12.37 | 6.08 |
| **Word frequency total, secondary wines** | both | 8.43 | 4.76 |

*A2 continued…*

| | | | |
|---|---|---|---|
| **Word frequency total, all wines** | both | 21.44 | 11.55 |
| **Visual word frequency, premium** | both | 9.00 | 3.98 |
| **Visual word frequency, secondary** | both | 6.63 | 3.14 |
| **Visual word frequency, all wines** | both | 15.63 | 6.74 |
| **Smell word frequency, premium** | both | 12.39 | 4.27 |
| **Smell word frequency, secondary** | both | 7.64 | 3.51 |
| **Smell word frequency, all wines** | both | 20.03 | 7.27 |
| **Taste word frequency, premium** | both | 10.60 | 2.83 |
| **Taste word frequency, secondary** | both | 8.28 | 3.81 |
| **Taste word frequency, all wines** | both | 18.25 | 5.82 |
| **Somesthetic word frequency, premium** | both | 8.62 | 3.51 |
| **Somesthetic word frequency, secondary** | both | 5.86 | 2.95 |
| **Somesthetic word frequency, all wines** | both | 15.48 | 7.68 |
| **Global word frequency, premium 1** | both | 17.26 | 7.77 |
| **Global word frequency, secondary 1** | both | 11.84 | 6.30 |
| **Global word frequency, all wines 1** | both | 30.09 | 14.23 |
| **Global word frequency, premium 2** | both | 16.09 | 6.26 |
| **Global word frequency, secondary 2** | both | 10.07 | 5.10 |
| **Global word frequency, all wines 2** | both | 28.26 | 13.95 |
| **% Premium total words** | both | 60.43 | 7.38 |

| | | | |
|---|---|---|---|
| **% Premium visual words** | both | 57.25 | 7.15 |
| **% Premium smell words** | both | 62.58 | 6.90 |
| **% Premium taste words** | both | 59.66 | 6.47 |
| **% Premium somesthetic words** | both | 59.64 | 8.26 |
| **% Premium global words 1** | both | 60.66 | 7.37 |
| **% Premium global words 2** | both | 62.29 | 6.66 |
| **Total double-word count** | both | 40.81 | 30.82 |
| **Visual double-word count** | both | 56.76 | 28.18 |
| **Smell double-word count** | both | 37.96 | 31.57 |
| **Taste double-word count** | both | 38.98 | 27.12 |
| **Somesthetic double-word count** | both | 35.78 | 35.28 |
| **Global double-word count 1** | both | 46.22 | 28.07 |
| **Global double-word count 2** | both | 27.10 | 22.65 |
| **Total consensus words. Premium** | both | 2.18 | 1.09 |
| **Total consensus words. Secondary** | both | 3.14 | 1.73 |
| **Total consensus words. All** | both | 5.29 | 2.67 |
| **Visual consensus words. Premium** | both | 4.04 | 0.82 |
| **Visual consensus words. Secondary** | both | 5.59 | 2.05 |
| **Visual consensus words. All** | both | 9.63 | 2.58 |
| **Smell consensus words. Premium** | both | 1.75 | 0.73 |
| **Smell consensus words. Secondary** | both | 2.61 | 0.79 |
| **Smell consensus words. All** | both | 4.36 | 1.31 |
| **Taste consensus words. Premium** | both | 2.11 | 0.72 |
| **Taste consensus words. Secondary** | both | 2.97 | 1.22 |
| **Taste consensus words. All** | both | 5.04 | 1.83 |
| **Somesthetic consensus words. Premium** | both | 2.15 | 0.73 |
| **Somesthetic consensus words. Secondary** | both | 2.89 | 1.13 |
| **Somesthetic consensus words. All** | both | 4.77 | 1.78 |
| **Global consensus words. premium 1** | both | 1.47 | 0.36 |
| **Global consensus words. secondary 1** | both | 2.21 | 0.67 |
| **Global consensus words. all 1** | both | 3.56 | 1.06 |
| **Global consensus words. premium 2** | both | 1.63 | 0.36 |
| **Global consensus words. secondary 2** | both | 2.67 | 1.12 |
| **Global consensus words. all 2** | both | 4.15 | 1.10 |

*Note: n = 22 population of tasters who completed all sessions 1-6; 77 predictors, 2 criteria (Quality Consensus Model, Image Consensus Model). Statistics: Averages, Standard Deviations.*

*A2. List of the top six consensus words in French (%) by sensory condition and wine category*

| **Visual consensus words (n = 55)** | | | | **Smell consensus words (n = 54)** | | | |
|---|---|---|---|---|---|---|---|
| Premium (459 words) | Group % | Secondary (324 words) | Gro up % | Premium (670 words) | Group % | Secondary (389 words) | Group % |
| Rouge | 4.3 | Couleur | 4.0 | Puissant | 2.8 | Fruits frais | 3.1 |
| Couleur | 4.8 | Intense | 4.0 | Fruits | 3.1 | Vanille | 3.3 |
| Sombre | 5.2 | Rubis | 7.7 | Fruits noirs | 3.1 | Finesse | 3.8 |
| Brillant | 7.4 | Rouge | 7.7 | Vanille | 3.3 | Boisé | 4.6 |
| **Profond** | **7.8** | **Limpide** | **8.3** | **Complexité** | **4.1** | **Fruits rouges** | **6.4** |
| **Intense** | **7.8** | **Brillant** | **12.3** | **Boisé** | **5.1** | **Fruits** | **9.0** |
| **Taste consensus words (n = 52)** | | | | **Somesthetic consensus words (n = 47)** | | | |
| Premium (662 words) | Group % | Secondary (423 words) | Gro up % | Premium (429 words) | Group % | Secondary (229 words) | Group % |
| Puissant | 3.2 | Finesse | 4.3 | Gras | 3.6 | Fluide | 3.9 |
| Structuré | 3.2 | Tanins | 4.7 | Chaud | 3.9 | Frais | 3.9 |
| Boisé | 3.3 | Rond | 5.4 | Rond | 4.9 | Tannins | 5.2 |
| Long | 3.9 | Fruité | 5.9 | Velouté | 5.1 | Rond | 5.7 |
| **Equilibre** | **4.7** | **Souple** | **6.1** | **Soyeux** | **5.6** | **Soyeux** | **5.7** |
| **Tanins** | **5.4** | **Equilibre** | **8.0** | **Tannins** | **5.6** | **Souple** | **6.1** |
| **All senses consensus words (n = 43)** | | | | **All senses rep. consensus words (n = 42)** | | | |
| Premium (775 words) | Group % | Secondary (448 words) | Gro up % | Premium (668 words) | Group % | Secondary (389 words) | Group % |
| Profond | 2.3 | Frais | 2.9 | Puissant | 3.0 | Frais | 3.1 |
| Bois | 2.8 | Fin | 3.3 | Complexe | 3.3 | Rond | 3.1 |
| Equilibre | 3.0 | Souple | 3.8 | Equilibre | 3.3 | Tanins | 3.1 |
| Long | 3.0 | Fruits Rouges | 4.2 | Longueur | 3.3 | Souple | 3.6 |
| **Complexe** | **3.2** | **Equilibre** | **4.9** | **Tanins** | **3.4** | **Equilibre** | **5.1** |
| **Tanins** | **3.4** | **Fruité** | **6.3** | **Boisé** | **3.7** | **Fruité** | **6.7** |

## S1. Supplementary materials: Vividness of Mental Imagery Questionnaires

### Vividness of Gustatory Imagery Questionnaire

These items describe taste sensations that may or may not evoke clear taste representations. Reflect on the taste representations that come to mind for each food item and evaluate the degree of realism and vividness of the resulting 'taste image' on a scale from 1 (No taste image) to 5 (Perfectly realistic taste image, as clear as the actual perception of taste).

**Five point scale – For each item, circle the corresponding intensity.**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| No taste image. You only know that you are thinking about the taste. | Vague, imprecise taste image. | Taste image neither clear nor unclear. | Relatively realistic, clear taste image. | Perfectly realistic, as clear as the actual taste perception. |

☞ **Imagine the following situation: You are participating in a cooking class and asked to reflect on different in-mouth tastes and their associated food products.**

The head chef asks you to reflect on **tastes associated with acidity**:
Imagine:

| | | | | | |
|---|---|---|---|---|---|
| The taste of lemon juice | 1 | 2 | 3 | 4 | 5 |
| The taste of vinegar pickles | 1 | 2 | 3 | 4 | 5 |
| The taste of natural yogurt | 1 | 2 | 3 | 4 | 5 |
| The taste of gooseberries | 1 | 2 | 3 | 4 | 5 |

The head chef asks you to reflect on **tastes associated with bitterness**:
Imagine:

| | | | | | |
|---|---|---|---|---|---|
| The taste of endives | 1 | 2 | 3 | 4 | 5 |
| The taste of arugula salad | 1 | 2 | 3 | 4 | 5 |
| The taste of black chocolate 85% | 1 | 2 | 3 | 4 | 5 |
| The taste of coffee (no sugar) | 1 | 2 | 3 | 4 | 5 |

The head chef asks you to reflect on **salty tastes**:
Imagine:

| | | | | | |
|---|---|---|---|---|---|
| The taste of oysters | 1 | 2 | 3 | 4 | 5 |
| The taste of plain potato chips/crisps | 1 | 2 | 3 | 4 | 5 |
| The taste of fine parmesan | 1 | 2 | 3 | 4 | 5 |
| The taste of table salt | 1 | 2 | 3 | 4 | 5 |

The head chef asks you to reflect on **sweet tastes**:
Imagine:

| | | | | | |
|---|---|---|---|---|---|
| The taste of one teaspoonful of honey | 1 | 2 | 3 | 4 | 5 |
| The taste of chocolate-hazelnut spread | 1 | 2 | 3 | 4 | 5 |
| The taste of strawberry jam | 1 | 2 | 3 | 4 | 5 |
| The taste of nougat | 1 | 2 | 3 | 4 | 5 |

Finally, the head chef asks you to reflect on the **following tastes**:
Imagine:

| | | | | | |
|---|---|---|---|---|---|
| The taste of a licorice stick | 1 | 2 | 3 | 4 | 5 |
| The taste of a piece of fatty bacon | 1 | 2 | 3 | 4 | 5 |
| The taste of soy sauce | 1 | 2 | 3 | 4 | 5 |
| The taste of smoked fish | 1 | 2 | 3 | 4 | 5 |

## Vividness of Tactile/Body Imagery Questionnaire

These items describe tactile and body sensations that may or may not evoke clear representations. Reflect on the tactile and body representations that come to mind for each item and evaluate the degree of realism and vividness of the resulting 'body and tactile image', on a scale from 1 (No body or tactile image) to 5 (Perfectly realistic tactile and body image, as clear as the actual perception).

**Five point scale – For each item, circle the corresponding intensity.**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| No tactile/body image. You only know that you are thinking about the sensation. | Vague, imprecise tactile/body image. | Tactile/body image neither clear nor unclear. | Relatively realistic, clear tactile/body image. | Perfectly realistic, as clear as the actual tactile/body perception. |

**Imagine touching:**

| | | | | | |
|---|---|---|---|---|---|
| Hot sand at the beach | 1 | 2 | 3 | 4 | 5 |
| A soft bath towel after your shower | 1 | 2 | 3 | 4 | 5 |
| The pointy end of a needle | 1 | 2 | 3 | 4 | 5 |
| Cold icy water | 1 | 2 | 3 | 4 | 5 |
| A fur coat | 1 | 2 | 3 | 4 | 5 |

**Imagine the following body sensations:**

| | | | | | |
|---|---|---|---|---|---|
| Relaxing in a warm water bath | 1 | 2 | 3 | 4 | 5 |
| Having a sore throat | 1 | 2 | 3 | 4 | 5 |
| Jumping into a pool | 1 | 2 | 3 | 4 | 5 |
| Walking fast in the cold | 1 | 2 | 3 | 4 | 5 |
| Lifting a heavy table | 1 | 2 | 3 | 4 | 5 |

**Imagine the tactile sensations when:**

| | | | | | |
|---|---|---|---|---|---|
| You taste an unripe grape | 1 | 2 | 3 | 4 | 5 |
| You take a bite from a piece of brioche | 1 | 2 | 3 | 4 | 5 |
| You take a sip of milk | 1 | 2 | 3 | 4 | 5 |
| You take a sip of steamy hot chocolate | 1 | 2 | 3 | 4 | 5 |
| You have melting ice on your tongue | 1 | 2 | 3 | 4 | 5 |

**While chewing, you imagine the shape and texture in your mouth:**

| | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |

Of a grape
Of a piece of corn
Of a chocolate chip
Of a cashew nut
Of a raspberry

| | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |

**You imagine the effect these foods might have in your mouth:**
The pungency of a hot pepper ("bite with teeth")
The pungency of a piece of ginger
The pungency of Dijon mustard
The pungency of wasabi sauce
The pungency of a piece of radish
The astringency of a piece of raw artichoke

| | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |

## Vividness of Wine Imagery Questionnaire

The items presented below refer to some situations that may relate closely to your professional experience in the wine sector. They may or may not evoke specific representations for you. Reflect on the representations that come to mind for each item and evaluate the degree of realism and vividness of the resulting image, on a scale from 1 (No image) to 5 (Perfectly realistic image, as clear as the actual perception).

**Five point scale – For each item, circle the corresponding intensity.**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| No wine image. You only know that you are thinking about the sensation. | Vague, imprecise wine image. | Wine image neither clear nor unclear. | Relatively realistic, clear wine image. | Perfectly realistic, as clear as the actual wine perception. |

☞ **As close as possible to your professional experience in the vineyards**

Imagine:
The smell of the tipcart on arrival of the harvest
The smell of musty grapes
The smell of a vat during fermentation
The smell of SO2
The smell of barrel wood

| | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |

Imagine:
The acidity of grape berries approaching maturity
The "sweetness" of berries in the grape basket
The bitterness of a wine after excessive aging in new wood
The taste of a balanced wine
The acidity of a wine before malolactic fermentation
The acidity of a poor vintage

| | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |
| 1 | 2 | 3 | 4 | 5 |

Imagine:

| | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |

| | | | | | |
|---|---|---|---|---|---|
| The mouthfeel of silky tannins | 1 | 2 | 3 | 4 | 5 |
| The mouthfeel of a wine with a dry finish | 1 | 2 | 3 | 4 | 5 |
| The mouthfeel of a wine lacking tannic structure | 1 | 2 | 3 | 4 | 5 |
| The mouthfeel of a wine described as "high-alcohol" | 1 | 2 | 3 | 4 | 5 |
| The mouthfeel resulting from unripe grape seeds | 1 | 2 | 3 | 4 | 5 |
| The mouthfeel of a wine described as hard, tough, and astringent | | | | | |

Imagine these actions:

| | | | | | |
|---|---|---|---|---|---|
| Tipping a glass (to look at the wine) | 1 | 2 | 3 | 4 | 5 |
| Swirling wine in a glass | 1 | 2 | 3 | 4 | 5 |
| Sucking in air when you have wine in your mouth | 1 | 2 | 3 | 4 | 5 |
| Leaning over to spit wine into a spittoon | 1 | 2 | 3 | 4 | 5 |

Imagine a wine bottle:

| | | | | | |
|---|---|---|---|---|---|
| When you touch the bottle, you notice that it is smooth | 1 | 2 | 3 | 4 | 5 |
| The temperature of the bottle in contact with your hand | 1 | 2 | 3 | 4 | 5 |
| The weight of a bottle of wine when you lift it | 1 | 2 | 3 | 4 | 5 |

Imagine:

| | | | | | |
|---|---|---|---|---|---|
| The color of a red Bordeaux wine from a recent vintage | 1 | 2 | 3 | 4 | 5 |
| A light red wine from a recent vintage | 1 | 2 | 3 | 4 | 5 |
| A wine that looks cloudy when you pour it into your glass | 1 | 2 | 3 | 4 | 5 |
| The hue of a mature wine | 1 | 2 | 3 | 4 | 5 |

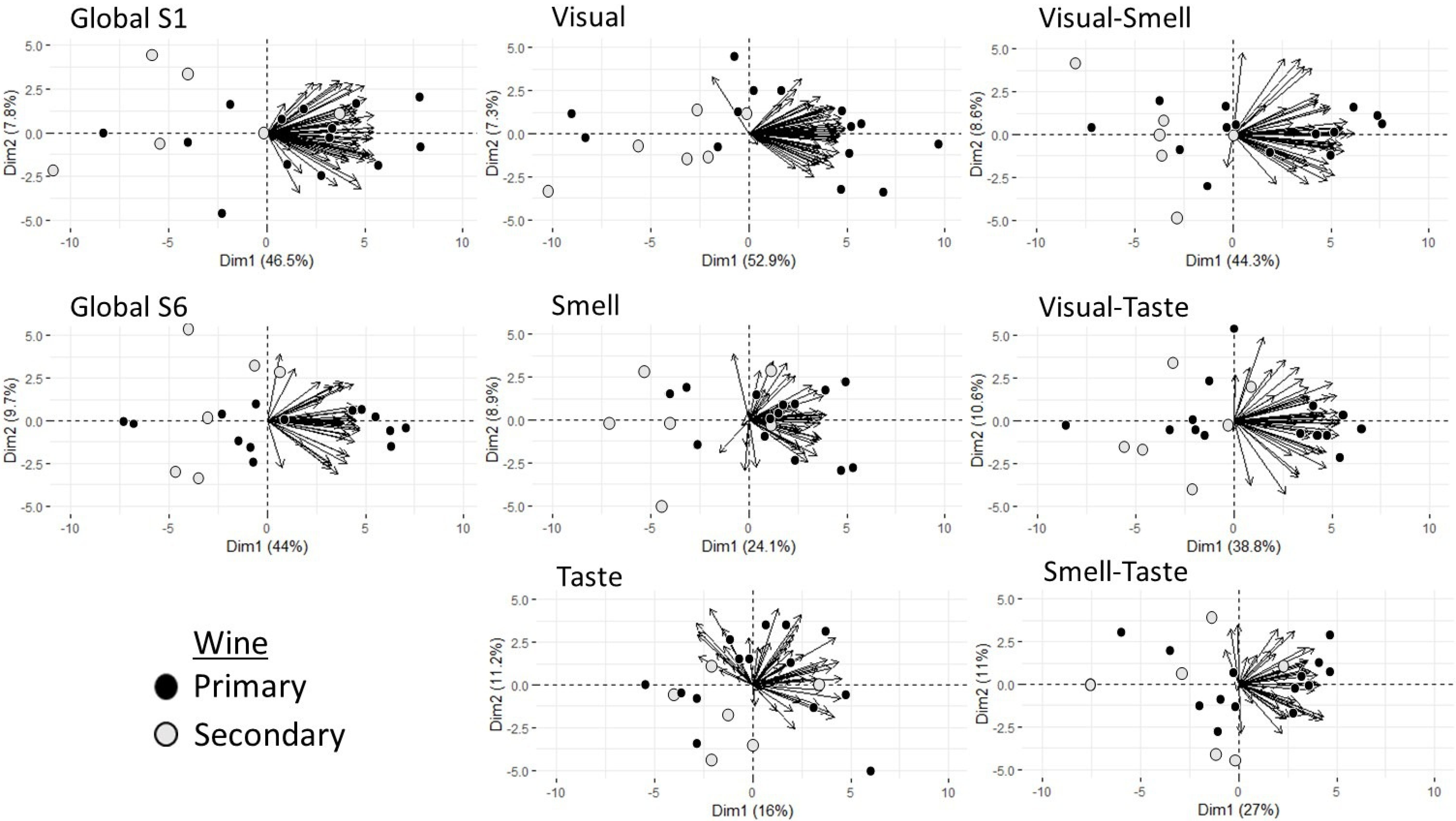
| | | | | | |
|---|---|---|---|---|---|
| The purplish hue of fermenting must | 1 | 2 | 3 | 4 | 5 |
| The color of a wine described as "spoilt" | 1 | 2 | 3 | 4 | 5 |
| An aged wine with light brownish-red tinges | 1 | 2 | 3 | 4 | 5 |
| The clarity of a bright red wine | 1 | 2 | 3 | 4 | 5 |
| Wine "tears" running slowly down the inside of a glass | 1 | 2 | 3 | 4 | 5 |

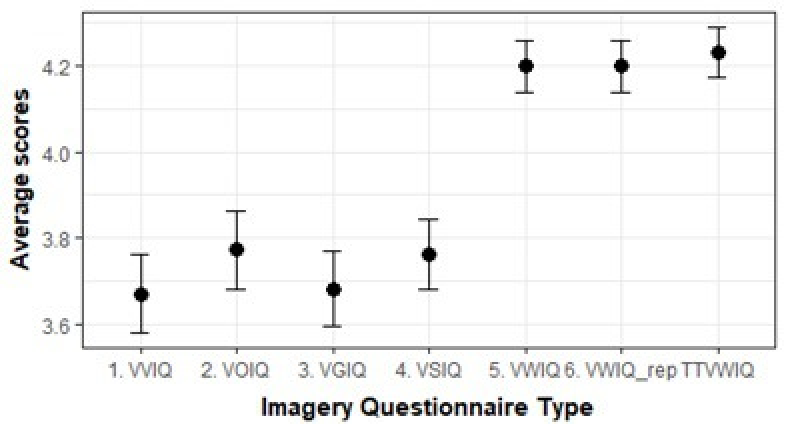
## Figure Captions

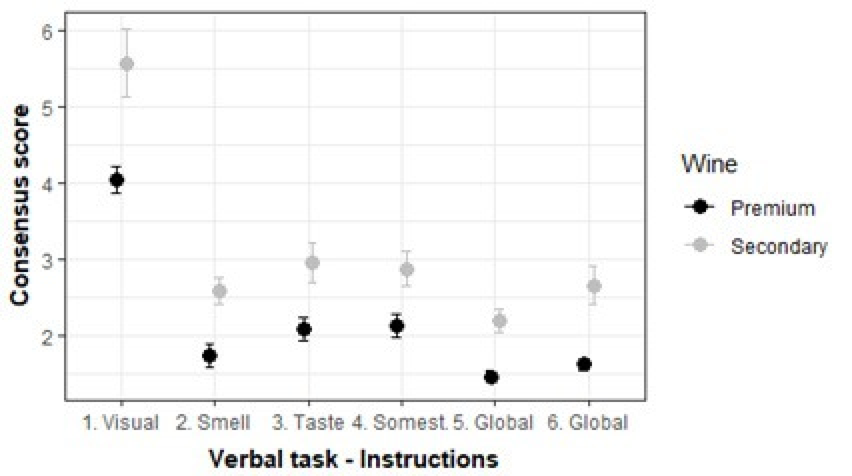
**Fig. 1.** Principal component analysis (PCAs) biplots for wine quality ratings, with level of agreement among experts (individuals, arrows) and wine ratings (Premium, black; Secondary, gray). Total agreement is based on the percentage explanation of the original variability on Dim1, with the inter-individual spread (arrows) visually substantiating the level of coherence or consensus among wine experts. Global tasting conditions: Global S1 and Global S6; Unimodal tasting conditions: visual, smell, and taste; Bimodal tasting conditions: visual-smell, visual-taste, and smell-taste.

**Fig. 2.** Self-assessment imagery scores by Imagery Questionnaire Type (VVIQ, VOIQ, VGIQ, VSIQ, VWIQ, VWIQ repetition, TTVWIQ) for 22 wine tasters who completed all sessions. Two repetitions of VWIQ, averaged into TTVWIQ scores.

**Fig. 3.** Average consensus score by verbal task instruction (Sensory condition) and wine category (Premium vs. Secondary) for 22 wine tasters who completed all sessions.


Global S1
Dim1 (46.5%)
Dim2 (7.8%)
Visual
Dim1 (52.9%)
Dim2 (7.3%)
Visual-Smell
Dim1 (44.3%)
Dim2 (8.6%)
Global S6
Dim1 (44%)
Dim2 (9.7%)
Smell
Dim1 (24.1%)
Dim2 (8.9%)
Visual-Taste
Dim1 (38.8%)
Dim2 (10.6%)
Taste
Dim1 (16%)
Dim2 (11.2%)
Smell-Taste
Dim1 (27%)
Dim2 (11%)
Wine
Primary
Secondary


Average scores
4.2
4.0
3.8
3.6
1. VVIQ
2. VOIQ
3. VGIQ
4. VSIQ
5. VWIQ
6. VWIQ_rep
TTVWIQ
Imagery Questionnaire Type



Consensus score
6
5
4
3
2
1. Visual
2. Smell
3. Taste
4. Somest.
5. Global
6. Global
Verbal task - Instructions
Wine
Premium
Secondary