Consistency of expert product reviews: an application to wine guides

Gabriel I. Penagos-Londoño[1], Felipe Ruiz-Moreno[2], Ricardo Sellers-Rubio[3], Salvador Del Barrio-García[4], Ana B. Casado-Díaz[5]

[1] Pontifical Xavierian University, Department of Economics, Carrera 7 No. 40 – 62, Bogota – COLOMBIA, Email: penagosi@javeriana.edu.co
[2] University of Alicante, Department of Marketing, Crta. San Vicente s/n. 03690, Alicante – SPAIN, Email: felipe.ruiz@ua.es
[3] University of Alicante, Department of Marketing, Crta. San Vicente s/n. 03690, Alicante – SPAIN, Email: ricardo.sellers@ua.es
[4] University of Granada, Department of Marketing and Marketing Research, Campus Universitario de Cartuja, 18071, Granada – SPAIN, Email: dbarrio@ugr.es
[5] University of Alicante, Department of Marketing, Crta. San Vicente s/n. 03690, Alicante – SPAIN, Email: ana.casado@ua.es

Correspondence concerning this article should be addressed to Felipe Ruiz-Moreno, University of Alicante, Department of Marketing, Crta. San Vicente s/n. 03690, Alicante – SPAIN, Email: felipe.ruiz@ua.es

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record.

Please cite this article as:

Penagos-Londoño G.I., Ruiz-Moreno F., Sellers-Rubio R., Del Barrio-García S., Casado-Díaz A.B. (2022), Consistency of expert product reviews: an application to wine guides, **Wine Economics and Policy**, *Just Accepted.*
DOI: 10.36253/wep-12400

Abstract

Purpose. The purpose of this study is to examine the internal consistency of wine guides by comparing the judgements of expert wine tasters and reviewers. A classification of wines is provided to establish whether expert reviews of similar wines are coherent.

Design/methodology/approach. Sentiment analysis based on natural language processing techniques was used to compare quantitative and qualitative reviews between experts. In addition, a finite mixture model was used to classify wines into categories to analyse internal consistency between ratings.

Findings. The results for a sample of more than 200,000 *Wine Enthusiast* ratings reveal significant differences between expert reviews. This finding indicates that there are no standard criteria for reviewing wines included in the guide.

Originality. Wine guides are amongst the most widely used marketing resources in the wine industry. They provide a signal to consumers about the quality of wines, guiding their purchase decisions. They also influence the reputation of brands and the performance of companies producing these wines. The main contribution of this study is to propose a new way to compare the reviews of wine guide experts.

Keywords: reputation, wine, expert ratings, sentiment analysis, finite mixture model, wine guides

1. Introduction

Information influences users' decision-making processes. However, information asymmetry generally exists in the buyer-seller relationship because each party has a different amount of information about products [1]. Research on experiential and hedonic consumption has shown that consumers' behaviour is affected by "social influence including peer input (word-of-mouth) and judgments of respected experts (professional evaluations)" [2, p.180].

Wine is an experience product whose quality cannot be assessed by consumers before purchase and consumption [3, 4]. This feature of wine increases the complexity of the purchase decision process. Thus, information asymmetries arise between consumers and winemakers in relation to product quality. Accordingly, high- and low-quality products can coexist in the market [5]. Wineries employ different marketing strategies to reduce these asymmetries and inform the market about the quality of their products [6]. Some use advertising in the mainstream media and encourage positive word-of-mouth communication amongst consumers [7, 8]. They also use awards in national and international competitions as part of their branding and communication strategies [6]. Finally, receiving high ratings in well-known wine guides, which are managed by experts and prescribers, can also help reduce information asymmetries between winemakers and consumers.

This study focuses on the social influence of experts in wine guides. Wine guides offer thousands of reviews of wines from around the world, basing their reviews on the opinions of panels of experts who taste these wines. The assumption is that consumers use judgements of wine quality by expert reviewers in wine guides as a source of information to make purchase decisions [9]. These expert reviewers might consequently influence the performance of the wine-producing companies. Previous research has in fact shown that there is a relationship between online reviews and consumer choice and firm sales [10, 11]. However, despite the potential impact on consumers and wineries, the nature and effects of expert opinions in wine guides remains an under-researched topic.

Wine experts usually provide a quantitative (score) and a qualitative (comment) review. The aim of this study is to test the consistency between these two assessments (quantitative and qualitative) of tasted wines. For wine guides to offer a credible source of information, both assessments of the same wine should match. That is, higher scores should be aligned with more positive comments. This analysis can confirm the role of expert evaluations as a credible source of information for consumers. To test the consistency of wine experts' reviews, the qualitative content (i.e. tasting notes) is examined using sentiment analysis based on natural language processing techniques. Then, these reviews and other relevant variables (origin and grape variety) are used to establish whether expert reviews of similar wines are coherent. Coherence is examined by classifying wines according to reviews and wine-related variables. A finite mixture model is employed for this classification. The study context is the *Wine Enthusiast* guide, one of the most prestigious wine guides in the world. The results show significant differences between expert reviews, which raises doubts about the usefulness and credibility of wine guides as a source of information.

2. Literature review

2.1. Wine guides as a marketing tool

Guides are extremely popular in the wine industry because they offer a point of comparison across brands [12] and provide consumers with a signal of wine quality. Wine guides are based on the opinions of experts and professional tasters, who follow standardised, systematic procedures that aim to provide a rigorous assessment of wines. These experts and tasters are assumed to be independent of wineries, thus helping consumers make informed purchase decisions, as the learning process necessary for consumers to become wine experts themselves takes time [13].

Research has highlighted the effect of wine expert recommendations from a marketing perspective. Parsons and Thompson [14] showed that consumers attribute high credibility to independent wine expert recommendations. Friberg and Grönqvist [15] found a significant effect of positive reviews by experts on the sales of the wines they had tasted. The scores that wines receive in these guides can also influence other marketing variables. A line of research has focused on the effect of expert reviews on wine prices [16]. For instance, studies have shown a positive effect of this type of evaluation on product prices, associated with a greater product reputation [7, 17]. Ashenfelter and Jones [18] showed that the influence of expert ratings on the price of wine is even greater than that of other factors such as terroir conditions or climate, which are commonly used to predict wine prices [19]. Wine research has also used the sensory reviews of experts in wine guides to measure wine quality and brand reputation [20]. Dressler [21] analysed the reputation of German wineries, individually and collectively, using three wine guides (Feinschmecker, Gault Millau and Eichelmann) and found consistent judgements across all three. Focused on Sicilian wines, Roma et al. [9] used experts' scores in wine guides as a proxy of firm (wine) reputation. This approach is common in the wine literature [22]. However, despite this evidence, the impact of a positive expert review on the price of a wine may depend not only on the reputation of the wine itself but also on the reputation of the expert [23, 24] because not all experts or guides have the same reputation and prestige [25].

2.2. The expert-consistency effect

According to dual-process theory [26], individuals' opinions and even behaviours are based on informational and normative influences such as those from expert reviews [27–29]. Information has a greater impact on the receiver if the sender is perceived as credible. Expert information is believed to be more credible and accurate (i.e. consistent) than non-expert information [30, 31].

In the context of wine, it is difficult to identify the factors that each expert considers when making judgements and rating wines because there is no common frame of reference across guides [16, 32]. An expert's rating is not necessarily an objective indicator of the quality of a wine because experts make judgements based on their own personal preferences. Thus, when an expert gives a high rating to a certain wine, it is not intended to convey the idea that the wine is of a higher quality than another wine with a lower rating. This lack of comparability arises because ratings of wines are conditioned by several factors such as origin, vintage, winery, price and even the expectations of the expert. Therefore, a higher score for one wine than for another simply indicates an expert's greater preference for that wine.

Consequently, despite their alleged objectivity (as stated in wine guides), expert reviews cannot be considered absolute objective assessments of wine quality. For instance, they may be biased by experts' personal preferences [33]. Evidence regarding the consistency of expert judgements is somewhat mixed. Some authors have found consistency between different experts' reviews of the same wine (e.g. [34]). However, other authors have expressed concern about inconsistencies between different experts' opinions of wine quality and even inconsistencies in reviews by the same expert over time (e.g. [35–37]). Cao and Stokes [38] reported that personal bias in wine expert reviews translates into different ratings, discriminatory capacity and variability in the ratings of different wines. Likewise, Ashton [35, 39] observed that wine guides focus on a few wines and cannot be considered fair representations of the entire market, noting that even the number of tasters used to issue a rating can influence the rating. These guides continue to be highly important in many markets and are used as a reference by consumers around the world. Therefore, further investigation of the effects of expert consistency/inconsistency is warranted.

2.3. Sentiment analysis: a tool for analysing the consistency of expert reviews

In recent years, natural language processing research techniques have allowed researchers to perform textual and sentiment analysis of reviews by both experts and consumers (e.g. [40–46]). Sentiment analysis is a subfield within natural language processing techniques that focuses on automatically classifying a text through its valence [47]. It enables the extraction of information on opinions about a subject (from users or experts) for a certain product [48, 49]. Previous research has shown that this type of analysis based on the characteristics of the product can provide more precise information than a general analysis of the overall (numerical) assessment [50]. Recent literature reviews have highlighted the importance and uniqueness of sentiment analysis in marketing research [51] and in hospitality and tourism [52].

In the context of wine guides, users typically find two ratings or judgements of a given wine. The first is a numerical score, usually on a scale of 0 to 100 points or 0 to 20 points, depending on the guide. Some guides only publish wines that receive a minimum score of 80 or 85 points. The second rating is a qualitative review based on tasting notes for the wine. These tasting notes consist of a brief literal description of the sensory and organoleptic qualities of the wine [53]. Although numerical scores are easily interpretable, the natural limitations of language hinder and complicate the task of using words to convey what a wine is really like and to describe the sensations that the expert wants to convey [54]. Sometimes, the sensory characteristics of wines are so special or unusual that there may not be the right words to describe it. Furthermore, some authors suggest that the language of professional tasting, which is used to describe the sensory properties of a wine, is based on jargon and vocabulary that is so complex and difficult to decipher that only the experts themselves or the most experienced consumers can understand it. In fact, Peynaud and Blouin [55] found that for professional tasting notes to be effective, consumers must have a high level of understanding about tasting, which is not always the case. Sometimes, these tasting notes may be pretentious, offering little informational validity for consumers [56].

Therefore, sentiment analysis based on each of the characteristics considered in the tasting notes could offer a broader and more accurate illustration of how experts review a wine. From an analytical perspective, the opinions of experts require analysis at the sentence level [57]. This sentence-level focus is necessary because experts who review wines consider different characteristics or attributes and generally have a different opinion on each of these aspects. Although many sentiment analysis tools can easily divide comments into negative, positive or neutral, a textual review of a given wine may contain phrases with different polarities because experts may have different feelings about each characteristic of the wine. For instance, the standard tasting phases (i.e. sight, smell and taste) may have different polarities, with some aspects being rated positively, others negatively and others neutrally. In addition, there may be different degrees of positive or negative opinions. Accordingly, reviews cannot be qualified simply as positive, negative or neutral. Instead, they include a series of additive perceptions that create a nuanced rating and provide specific information on each of the aspects evaluated by the expert. For instance, some characteristics of the wine (e.g. in the olfactory phase of tasting) may be rated positively, whereas others (e.g. related to the palate) may be negatively rated.

In sum, sentiment analysis techniques could lead to precise inference of the overall numerical score for the wine. Therefore, these techniques are particularly useful for examining the opinions of experts about the wines in a guide. Nguyen et al. [58] recently employed a similar approach, focusing on so-called online expert users.

3. Method

This study focuses on reviews by 19 professional wine tasters from the *Wine Enthusiast* guide between 1999 and 2019. *Wine Enthusiast Magazine* is one of the most prestigious international magazines in the sector, together with *The Wine Advocate* (Robert Parker). Each review included qualitative tasting notes, in which the expert gave a judgement on the tasted wine, a quantitative score of the wine (from 80 to 100 points), and some additional characteristics such as price, origin and grape variety (see Figure 1). The wines were from 43 countries and their price ranged from 4 dollars to 3,400 dollars. After the elimination of outliers and missing cases, the final sample contained 201,004 reviews.

FIGURE 1. **Sample *Wine Enthusiast* guide review**

Source: *Wine Enthusiast* website (2021).

The method had two stages. The first stage involved that quantitative ratings as well as qualitative reviews were compared among the different experts in the guide. Reviews published in the guide were made by 19 experts, as well as some other anonymous reviewers. Although the comparison of quantitative ratings was straightforward, the comparison of qualitative reviews required prior analysis of tasting notes using sentiment analysis. This analysis was carried out using the AFINN lexicon. AFINN consists of 2,477 words in English that express a certain degree of positive or negative sentiment. This corpus of words, produced by Finn Arup Nielsen between 2009 and 2011, contains a rating for words ranging from −5 (*most negative sentiment*) to +5 (*most positive sentiment*). This lexicon displays the information in two columns: the word next to its corresponding value (e.g. “awesome” - 4 or “awful” -3). In this study, the sentiment value of the expert review was calculated as the sum of the polarity of each of the words used in the review. In essence, each review was divided into sentences, and each sentence into words. To evaluate one sentence of the review, each word was assigned a value according to the AFINN lexicon. Adding up the values of all words in the sentence gave an evaluation of that specific comment. Once this process had been performed for all sentences in the review, the evaluations of each sentence or comment were summed to give an overall score for the review. Because an expert review covers different aspects, different opinions can be found in the same review. That is, the same review might contain both positive comments (e.g. regarding palate) and negative comments (e.g. regarding nose). However, the additive procedure employed in this study gave an overall evaluation of the intensity (value) and polarity (positive/negative) of the review based on the evaluation of each comment in the review. Compared to the alternative of using the average of the individual evaluations of each word, this additive procedure accounted for the length of the review because there is evidence that longer reviews provide greater added value to the tasting note of the wine [53]. In addition, it provided a broader ranking of the review than a simple classification as positive, negative or neutral.

In the second stage, the wines were classified according to their characteristics using techniques based on cluster analysis. The starting assumption was that wines in a given group were homogeneous but different from the wines in other groups. Each wine was defined by a set of variables related to its review (qualitative and quantitative), origin and grape. The objective of this stage was to group similar wines by comparing specific vectors for the set of variables used in this study. An $N \times d$ matrix was created for this analysis, where the columns were the variables, and the rows were the observations. Each observation (i.e. row) was a vector of dimension d, denoted as x_i. The data set was denoted as $x = (x_i)_{i \in \{1,\cdots,N\}}$. Each observation had d_{cont} continuous variables in $\mathbb{R}^{d_{cont}}$ and d_{cat} categorical variables, with $\{1, \cdots, m_j\}$ levels for each nominal variable j. Hence, $d_{cont} + d_{cat} = d$.

To classify the observations into groups that could be interpreted in a meaningful way, an unsupervised learning method was used. It was hypothesised that there existed hidden or latent variables (unobserved random variables) for all data points in the data set that associated a specific cluster to each observation. Thus, the latent variable model was a mixture model.

In a mixture model, K distributions are mixed, and it is assumed that each observation belongs to one of them. The latent variable z_i for observation i corresponds to one of the distributions in the mixture. In other words, the latent variable z_i is the cluster to which observation x_i belongs. If the number of clusters is K, then $z_i \in \{1, \cdots, K\}$, and the set of latent variables is denoted as $z = (z_i)_{i \in \{1,\cdots,N\}}$. In a mixture model, the data generation process is assumed to be $p(z, x) = p(z_i)p(x_i|z_i = k)$. Here,

$p(z_i)$ is a multinomial distribution, where $\eta_k = Pr(z_i = k)$ is the probability that observation i belongs to cluster k. The set of probabilities $\eta = (\eta_k)_{k\in\{1,\cdots,K\}}$ are referred to as the mixing weights. Furthermore, $\phi_k(x_i|\theta_k) = p(x_i|z_i = k)$ is the probability distribution of the data in cluster k, and θ_k are the parameters of this distribution. The probability density function is given as follows:

$$f(x_i|\theta) = \sum_{k=1}^{K} \eta_k \phi_k(x_i|\theta_k)$$

where $\theta = (\theta_k)_{k\in\{1,\cdots,K\}}$ is the set of all parameters for the distributions in the mixture, including the mixing weights.

For continuous variables, the cluster distributions were multivariate Gaussian distributions $\phi_k(x_i|\theta_k) = N(x_i|\mu_k, \Sigma_k)$, where the parameters of the distribution k, $\theta_k = \{\mu_k, \Sigma_k\}$ were the mean vector μ_k and covariance matrix Σ_k. Categorical variables were assumed to be independent multivariate multinomial variables distributed conditional on the latent variable. Therefore, $\phi_k(x_i|\theta_k) = \mathcal{M}(x_i|\alpha_k)$ for $\alpha_k = \left(\alpha_{jk}\right)_{j\in\{1,\cdots,d_{cat}\}}$, where α_{jk} is the vector of parameters (event probabilities) for the multinomial distribution associated with variable j in cluster k, and its dimension is m_j.

For the estimation of the parameters, the R package Rmixmod version 2.1.5 was used. This package maximises the log-likelihood with an expectation maximisation (EM) algorithm as follows:

$$\mathcal{L}(\Theta) = \sum_{i=1}^{N} \ln f(x_i|\theta)$$

for $\Theta = \{\eta, \theta\}$, the set of all parameters of the mixture.

Once the wines had been classified into similar groups, the differences between the expert reviews of the wines belonging to each cluster were analysed. The data processing and estimation was carried out in MATLAB.

4. Results

In the first stage, the quantitative and qualitative expert reviews in the guide were compared. The average score of the tasted wines was 88.81 points (SD = 3.03), with a minimum of 80 points and a maximum of 100. The experts used an average of 40.56 words in their descriptions of wines (SD = 11.28), with a minimum of three words and a maximum of 135. The average sentiment score was 3.2 points (SD = 7.02), with a minimum of -33 points and a maximum of 41. The average price was 36.62 dollars (SD = 43.17), with a minimum of 4 dollars and a maximum of 3,400 dollars.

Table I presents the average quantitative and sentiment ratings for each expert. It also shows the average number of words used by each expert in the tasting notes. There are statistically significant differences between the experts' quantitative ratings. There are also differences in the nuances provided in the tasting notes, as reflected by the differences in the number of words used and the sentiment ratings for the experts.

TABLE I. **Ratings of wines according to experts**

Expert	No. of wines tasted	Average quantitative score	Average of sentiment rating	Average number of words
Alexander Peartree	1,637	87.14	-1.28	41.26
Anna Lee C. Iijima	8,061	89.37	0.83	41.38
Anne Krebiehl MW	7,661	91.02	5.27	47.17
Carrie Dykes	268	86.45	1.10	42.75
Christina Pickard	2,349	88.97	1.72	57.00
Fiona Adams	408	86.72	-3.91	49.77
Jeff Jenssen	783	88.08	-1.39	35.75
Jim Gordon	9,083	88.71	4.71	38.12
Joe Czerwinski	5,842	88.66	0.24	40.96
Kerin O'Keefe	20,055	89.12	-1.88	38.03
Lauren Buzzeo	2,886	88.00	3.18	50.53
Matt Kettmann	13,910	90.21	-0.43	44.40
Michael Schachner	20,004	86.99	0.28	42.42
Mike DeSimone	956	89.07	-0.44	43.21
Paul Gregutt	13,824	89.34	4.61	43.48
Roger Voss	40,124	88.90	8.58	37.47
Sean P. Sullivan	9,197	88.67	1.74	38.39
Susan Kostrzewa	1,170	86.89	6.03	39.71
Virginie Boone	17,578	89.67	2.75	38.71
Nameless	25,208	87.81	4.10	38.96
Total	**201,004**	**88.81**	**3.20**	**40.55**
F		**1158.84 ($p < 0.000$)**	**3534.31 ($p < 0.000$)**	**1351.94 ($p < 0.000$)**

Source: Authors

In the second stage, the wines were classified according to their characteristics using techniques based on cluster analysis. The proposed model was estimated for K = 2, . . . ,7 clusters in relation to the wines appearing in this guide. To identify the clusters, four variables were used: the quantitative rating, sentiment score of the tasting note, country of origin of the wine and grape variety. The model selection criterion was the Bayesian information criterion (BIC; [59] Schwarz 1978). This criterion suggested that K = 4 was the number of groups that best fit the data (see Table II). External validation is also desirable to confirm the usefulness of the cluster solution. External validation consisted of examining whether there were also intercluster differences in variables other than those used to classify the wines. This external validation served as an exploratory investigation of the influence of the cluster structure and main characteristics [60]. To this end, the price variable was also examined (see Table II).

TABLE II. **Descriptive analysis of clusters with mean and standard deviation (in parentheses)**

	Variables used in the cluster analysis				External validation
	Quantitative rating	**Qualitative (sentiment) rating**	**Main country origin**	**Main grape variety**	**Price**
Best quality N = 56,043	90.09 (2.77)	10.26 (5.74)	France	Red & White	41.50 (65.01)
Affordable N = 48,321	85.29 (1.74)	1.33 (4.47)	America, France and Spain	Red & White	21.10 (16.40)
Over-priced N = 67,789	90.00 (2.23)	0.08 (5.75)	United States and Italy	Red	47.24 (37.42)
Smart choice N = 28,851	89.41 (2.15)	-0.02 (5.65)	United States	White	28.80 (25.02)
TOTAL N = 201,004	88.81 (3.03)	3.21 (7.02)	N.A.	N.A.	36.62 (43.16)

Source: Authors

The empirical findings reveal some interesting differences between the clusters. The first group, "top-of-the-range wines (*best quality*)", consists of wines with a well-above-average rating based on both sentiment and quantitative ratings. These wines are also on average more expensive. It consists of red and white wines, mainly from France. The second group, "low-price wines (*affordable/low cost*)", consists of wines with a below-average quantitative score but with a slightly positive sentiment rating. The average price of wines in this group is well below the average for the entire sample. This group includes white and red wines from North and South America, France and Spain. The third group, "*overpriced* wines", consists of wines with a neutral sentiment rating but a roughly average quantitative score. These wines' average price is well above the average for the entire sample. They are mostly red wines from the United States and Italy. Finally, the fourth group, "best-value wines (*smart choice*)", consists of wines with a roughly average quantitative score and a below-average qualitative rating. They also have a lower-than-average price. This group mainly consists of white wines from the United States.

The differences between the four groups were significant for the four variables considered in the analysis. In addition, for the external validation of the four clusters, ANOVA was used to test whether the prices differed between clusters. The price variable (4064.87; < 0.0001) was significantly different between clusters, thereby externally validating the classification presented in this research.

Once the wines had been classified into homogeneous groups, the average sentiment evaluations of the tasters were calculated for each group. The results indicate that the differences between the experts' reviews differ significantly, which shows that there are no standard criteria for reviewing the wines in the guide (see Table III). This result reinforces the earlier idea (see Table I) that tasting notes might differ amongst wine experts, even when the tasted wines are similar and receive a comparable quantitative rating.

TABLE III. **Test of differences of experts' sentiment ratings**

	F	*p* value
Group 1. Best quality	382.65	$p < 0.001$
Group 2. Affordable	110.97	$p < 0.001$
Group 3. Overpriced	295.44	$p < 0.001$
Group 4. Smart choice	151.12	$p < 0.001$

Source: Authors

5. Conclusions

Wine guides written by professional and expert tasters are widely used in the wine industry to market wine, providing important information signals for consumers around the world. However, despite the importance of these guides, some authors have expressed doubts about the consistency of the scores and reviews they provide. The objective of this study was to analyse the internal consistency of the scores and reviews of the experts and professional tasters writing for a specific guide. The method included sentiment analysis of the tasting notes and a novel clustering technique that identified groups of wines with similar characteristics.

The results show considerable divergence between the qualitative and quantitative assessments by professional tasters in the *Wine Enthusiast* wine guide. Although most consumers trust the guide to reduce their information asymmetries with respect to winemakers, disparity in the criteria used by the guide's experts raises doubts over its effectiveness as a source of reliable, verified, standardised information for consumers. In fact, even when wines are grouped according to their characteristics, there are still discrepancies amongst experts. Therefore, it cannot be said that the guide follows a single, uniform set of criteria for its wine reviews.

These results have managerial implications for the wine sector. First, the results have implications for wineries whose wines are tasted by experts writing for this guide. These wineries should be aware that experts' personal preferences may affect their judgements. Hence, knowing the personal tastes and background of each expert could help wineries improve the ratings of their wines. Second, these results are important for the management of the guide itself. The reputation and prestige of a particular guide is the basis of consumers' trust in that guide, which is considered a reliable and independent source of information. If the reviews in the guide are inconsistent and the experts do not reach a consensus when rating wines, doubts may arise about the reliability of these reviews, depending on which expert tasted the wine. These doubts could ultimately affect the publication's reputation.

Finally, regarding the limitations of this study, only one guide (*Wine Enthusiast*) was analysed. It is not possible to extrapolate these results to other specialist publications within the sector. Furthermore, the sentiment analysis was carried out using a specific lexicon. Although this lexicon has been widely used in academic studies, it is not the only available alternative, nor is it specific to the wine sector. These limitations open new research opportunities that should be addressed in the future. Future research could also explore the effect of reviewer expertise in the context of wine guides. Reviewer expertise has already been shown to influence reviewer ratings in the context of hotel and restaurant review platforms [58]. Finally, future research could extend this analysis to other markets where guides based on expert reviews are also common. Examples include the film and television industry, where sentiment analysis techniques have already been used to study expert and consumer opinions [2] but not to study specialised guides (e.g. Rotten Tomatoes).

References

[1] L. Fan, X. Zhang, and L. Rai, When should star power and eWOM be responsible for the box office performance? - An empirical study based on signaling theory, J. Retail. Consum. Serv., 62, 102591, 2021, https://doi.org/10.1016/j.jretconser.2021.102591.

[2] R. Niraj and J. Singh, Impact of User-Generated and Professional Critics Reviews on Bollywood Movie Success, Australas. Mark. J., 23(3), 179–187, 2015, https://doi.org/10.1016/j.ausmj.2015.02.001.

[3] E. Oczkowski, Hedonic Wine Price Functions and Measurement Error, Econ. Rec., 77(239), 374–382, 2001, https://doi.org/10.1111/1475-4932.00030.

[4] S. Petropoulos, C. S. Karavas, A. T. Balafoutis, I. Paraskevopoulos, S. Kallithraka, and Y. Kotseridis, Fuzzy logic tool for wine quality classification, Comput. Electron. Agric., 142, 552–562, 2017, https://doi.org/10.1016/j.compag.2017.11.015.

[5] G. A. Akerlof, The Market for 'Lemons': Quality Uncertainty and the Market Mechanism, Q. J. Econ., 84(3), 488–500, 1970, https://doi.org/10.2307/1879431.

[6] U. R. Orth and P. Krška, Quality signals in wine marketing: the role of exhibition awards, Int. Food Agribus. Manag. Rev., 4(4), 385–397, 2001.

[7] R. Sellers-Rubio, F. Mas-Ruiz, and F. Sancho-Esper, Firm reputation, advertising investment, and price premium: The role of collective brand membership in high-quality wines, Agribusiness, 34(2), 351–362, 2018, https://doi.org/10.1002/agr.21526.

[8] T. Spawton, Marketing planning for wine, Int. J. Wine Mark., 2(2), 2–49, 1990.

[9] P. Roma, G. Di Martino, and G. Perrone, What to show on the wine labels: a hedonic analysis of price drivers of Sicilian wines, Appl. Econ., 45(19), 2765–2778, 2013, https://doi.org/10.1080/00036846.2012.678983.

[10] K. Floyd, R. Freling, S. Alhoqail, H. Y. Cho, and T. Freling, How Online Product Reviews Affect Retail Sales: A Meta-analysis, J. Retail., 90(2), 217–232, 2014, https://doi.org/10.1016/j.jretai.2014.04.004.

[11] N. Hu, L. Liu, and J. J. Zhang, Do online reviews affect product sales? The role of reviewer characteristics and temporal effects, Inf. Technol. Manag., 9(3), 201–214, 2008, https://doi.org/10.1007/s10799-008-0041-2.

[12] A. J. Blair, C. Atanasova, L. Pitt, A. Chan, and Å. Wallstrom, Assessing brand equity in the luxury wine market by exploiting tastemaker scores, J. Prod. Brand Manag., 26(5), 447–452, 2017, https://doi.org/10.1108/JPBM-06-2016-1214.

[13] K. A. Latour and J. A. Deighton, Learning to Become a Taste Expert, J. Consum. Res., 46(1), 1–19, 2019, https://doi.org/10.1093/jcr/ucy054.

[14] A. G. Parsons and A. Thompson, Wine recommendations: who do I believe?, Br. Food J., 111, 9, 1003–1015, 2009, https://doi.org/10.1108/00070700910992899.

[15] R. Friberg and E. Grönqvist, Do Expert Reviews Affect the Demand for Wine?, Am. Econ. J. Appl. Econ., 4(1), 193–211, 2012, https://doi.org/10.1257/app.4.1.193.

[16] A. Albright, P. Pedroni, and S. Sheppard, Uncorking Expert Reviews with Social Media: A Case Study Served with Wine, p. 19, 2018.

[17] S. González-Hernando, V. Iglesias-Argüelles, and C. González-Mieres, ¿Cómo afectan las prescripciones de terceras partes a las evaluaciones del consumidor y a la prima de precio del producto?, Catedra Fundación Ramón Areces de Distribución Comercial, 1304, 2013. Accessed: Apr. 08, 2021. [Online]. Available: https://ideas.repec.org/p/ovr/docfra/1304.html

[18] O. Ashenfelter and G. V. Jones, The Demand for Expert Opinion: Bordeaux Wine*, J. Wine Econ., 8(3), 285–293, 2013, https://doi.org/10.1017/jwe.2013.22.

[19] O. Ashenfelter, D. Ashmore, and R. Lalonde, Bordeaux Wine Vintage Quality and the Weather, CHANCE, 8, 4, 7–14, 1995, https://doi.org/10.1080/09332480.1995.10542468.

[20] B. A. Benjamin and J. M. Podolny, Status, quality, and social order in the California wine industry, Adm. Sci. Q., 44(3), 563–589, 1999.

[21] M. Dressler, Strategic winery reputation management – exploring German wine guides, Int. J. Wine Bus. Res., 28(1), 4–21, 2016, https://doi.org/10.1108/IJWBR-10-2014-0046.

[22] L. Benfratello, M. Piacenza, and S. Sacchetto, Taste or reputation: what drives market prices in the wine industry? Estimation of a hedonic model for Italian premium wines, Appl. Econ., 41(17), 2197–2209, 2009, https://doi.org/10.1080/00036840701222439.
[23] H. H. Ali and C. Nauges, The Pricing of Experience Goods: The Example of en primeur Wine, Am. J. Agric. Econ., 89(1), 91–103, 2007.
[24] D. A. Reinstein and C. M. Snyder, The Influence of Expert Reviews on Consumer Demand for Experience Goods: A Case Study of Movie Critics*, J. Ind. Econ., 53(1), 27–51, 2005, https://doi.org/https://DOI.org/10.1111/j.0022-1821.2005.00244.x.
[25] V. Odorici and R. Corrado, Between Supply and Demand: Intermediaries, Social Networks and the Construction of Quality in the Italian Wine Industry, J. Manag. Gov., 8(2), 149–171, 2004, https://doi.org/10.1023/B:MAGO.0000026542.18647.48.
[26] M. Deutsch and H. B. Gerard, A study of normative and informational social influences upon individual judgment, J. Abnorm. Soc. Psychol., 51, 629–636, 1955, https://doi.org/10.1037/h0046408.
[27] M. Y. Cheung, C. Luo, C. L. Sia, and H. Chen, Credibility of Electronic Word-of-Mouth: Informational and Normative Determinants of On-line Consumer Recommendations, Int. J. Electron. Commer., 13(4), 9–38, 2009, https://doi.org/10.2753/JEC1086-4415130402.
[28] A. Naujoks and M. Benkenstein, Who is behind the message? The power of expert reviews on eWOM platforms, Electron. Commer. Res. Appl., 44, 101015, 2020, https://doi.org/10.1016/j.elerap.2020.101015.
[29] P. Racherla and W. Friske, Perceived 'usefulness' of online consumer reviews: An exploratory investigation across three services categories, Electron. Commer. Res. Appl., 11(6), 548–559, 2012, https://doi.org/10.1016/j.elerap.2012.06.003.
[30] A. S. Lo and S. S. Yao, What makes hotel online reviews credible? An investigation of the roles of reviewer expertise, review rating consistency and review valence, Int. J. Contemp. Hosp. Manag., 31(1), 41–60, 2019, https://doi.org/10.1108/IJCHM-10-2017-0671.
[31] S. Quaschning, M. Pandelaere, and I. Vermeir, When Consistency Matters: The Effect of Valence Consistency on Review Helpfulness, J. Comput.-Mediat. Commun., 20(2), 136–152, Mar. 2015, https://doi.org/10.1111/jcc4.12106.
[32] I. Olkin, Y. Lou, L. Stokes, and J. Cao, Analyses of Wine-Tasting Data: A Tutorial, J. Wine Econ., 10(1), 4–30, 2015, https://doi.org/10.1017/jwe.2014.26.
[33] S. Castriota, D. Curzi, and M. Delmastro, Tasters' bias in wine guides' quality evaluations, Appl. Econ. Lett., 20(12), 1174–1177, 2013.

[34] E. T. Stuen, J. R. Miller, and R. W. Stone, An Analysis of Wine Critic Consensus: A Study of Washington and California Wines, J. Wine Econ., 10(1), 47–61, 2015, https://doi.org/10.1017/jwe.2015.3.

[35] R. H. Ashton, Reliability and consensus of experienced wine judges: Expertise within and between, J. Wine Econ., 7(1), 70–87, 2012.

[36] R. H. Ashton, Is there consensus among wine quality ratings of prominent critics? An empirical analysis of red Bordeaux, 2004-2010, J. Wine Econ., 8(2), 225, 2013.

[37] R. T. Hodgson, An examination of judge reliability at a major US wine competition, J. Wine Econ., 3(2), 105–113, 2008.

[38] J. Cao and L. Stokes, Evaluation of Wine Judge Performance through Three Characteristics: Bias, Discrimination, and Variation*, J. Wine Econ., 5(1), 132–142, 2010, https://doi.org/10.1017/S1931436100001413.

[39] R. H. Ashton, Improving Experts' Wine Quality Judgments: Two Heads Are Better than One*, J. Wine Econ., 6(2), 160–178, 2011, https://doi.org/10.1017/S1931436100001577.

[40] F. Caviggioli, L. Lamberti, P. Landoni, and P. Meola, Technology adoption news and corporate reputation: sentiment analysis about the introduction of Bitcoin, J. Prod. Brand Manag., 29(7), 877–897, 2020, https://doi.org/10.1108/JPBM-03-2018-1774.

[41] B. Chen, V. Velchev, J. Palmer, and T. Atkison, Wineinformatics: A Quantitative Analysis of Wine Reviewers, Fermentation, 4(4), 82, 2018, https://doi.org/10.3390/fermentation4040082.

[42] N. Kotonya, P. De Cristofaro, and E. De Cristofaro, Of Wines and Reviews: Measuring and Modeling the Vivino Wine Social Network, in 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, 2018, 387–392. https://doi.org/10.1109/ASONAM.2018.8508776.

[43] E. Lefever, I. Hendrickx, and A. van den Bosch, Very quaffable and great fun: applying NLP to wine reviews, in Computational Linguistics in the Netherlands, 2015. Accessed: Apr. 08, 2021. [Online]. Available: http://hdl.handle.net/1854/LU-7198586

[44] E. Lefever, I. Hendrickx, and I. Croijmans, Discovering the Language of Wine Reviews: A Text Mining Account, Proc. Elev. Int. Conf. Lang. Resour. Eval. LREC 2018, 3297-3302, 2018, Accessed: Apr. 08, 2021. [Online]. Available: https://eprints.whiterose.ac.uk/137244/p. 6, 2018.

[45] S. Moon and W. A. Kamakura, A picture is worth a thousand words: Translating product reviews into a product positioning map, Int. J. Res. Mark., 34(1), 265–285, 2017, https://doi.org/10.1016/j.ijresmar.2016.05.007.

[46] G. J. Thebaud, P. Kraus, and M. Mondaresnezhad, Applying Text Data Analytics Techniques to Wine Reviews, presented at the AMCIS 2020 TREOs, 2020, 53, 2.

[47] B. Pang and L. Lee, Opinion mining and sentiment analysis, Found. Trends Inf. Retrieva, 2(1–2), 1–135, 2008.

[48] M. Joshi, P. Prajapati, A. Shaikh, and V. Vala, A survey on Sentiment Analysis, Int. J. Comput. Appl., 163(6), 34–38, 2017, https://doi.org/10.5120/ijca2017913552.

[49] X. Wang, T. Zhou, X. Wang, and Y. Fang, Harshness-aware sentiment mining framework for product review, Expert Syst. Appl., 187, 115887, 2022, https://doi.org/10.1016/j.eswa.2021.115887.

[50] A. Gandomi and M. Haider, Beyond the hype: Big data concepts, methods, and analytics, Int. J. Inf. Manag., 35(2), 137–144, 2015, https://doi.org/10.1016/j.ijinfomgt.2014.10.007.

[51] M. Rambocas and B. G. Pacheco, Online sentiment analysis in marketing research: a review, J. Res. Interact. Mark., 12(2), 146–163, 2018, https://doi.org/10.1108/JRIM-05-2017-0030.

[52] F. Mehraliyev, I. C. C. Chan, and A. P. Kirilenko, Sentiment analysis in hospitality and tourism: a thematic and methodological review, Int. J. Contemp. Hosp. Manag., 34(1), 46–77, 2021, https://doi.org/10.1108/IJCHM-02-2021-0132.

[53] C. D. Ramirez, Do Tasting Notes Add Value? Evidence from Napa Wines, J. Wine Econ., 5(1), 143–163, 2010.

[54] K. Lehrer and A. Lehrer, Winespeak or critical communication? Why people talk about wine, 2008.

[55] E. Peynaud and J. Blouin, The Taste of Wine: The Art Science of Wine Appreciation. John Wiley & Sons, 1996.

[56] R. E. Quandt, On Wine Bullshit: Some New Software?, J. Wine Econ., 2(2), 129–135, 2007, https://doi.org/10.1017/S1931436100000389.

[57] R. Feldman, Techniques and applications for sentiment analysis, Commun. ACM, 56, 4, 82–89, Apr. 2013, https://doi.org/10.1145/2436256.2436274.

[58] P. Nguyen, X. (Shane) Wang, X. Li, and J. Cotte, Reviewing Experts' Restraint from Extremes and Its Impact on Service Providers, J. Consum. Res., 47(5), 654–674, 2021, https://doi.org/10.1093/jcr/ucaa037.

[59] G. Schwarz, Estimating the Dimension of a Model, Ann. Stat., 6(2), 1978, https://doi.org/10.1214/aos/1176344136.

[60] R. P. Dant and G. T. Gundlach, The challenge of autonomy and dependence in franchised channels of distribution, J. Bus. Ventur., 14(1), 35–67, 1999, https://doi.org/10.1016/S0883-9026(97)00096-7.