

**Is complexity worth paying for? Investigating the perception of wine
complexity for single varietal and blended wines in consumers and experts**

Qian Janice Wang^{1,2} & Charles Spence¹

¹ Crossmodal Research Laboratory, Department of Experimental Psychology, Oxford University, Oxford, UK

²Department of Food Science, Faculty of Science and Technology, Aarhus University, Aarslev, Denmark

DATE: OCTOBER 2018

WORD COUNT: 6800

CORRESPONDENCE TO: Professor Qian Janice Wang, Department of Food Science, Aarhus University, 10 Kirstinebjergvej, Aarslev, 5792, Denmark.

E-mail: qianjanice.wang@food.au.dk

ABSTRACT

Background and Aims: Complexity is widely believed to be a positive trait in wine, but its precise definition remains ambiguous. The present study aims to assess how complexity in wine is perceived by social drinkers and experts, especially which attributes are most linked to their assessment of complexity.

Methods and Results: The study involved three young single-varietal wines (Merlot, Cabernet-Franc, Cabernet-Sauvignon) and three 50-50 blends of those wines. The 50-50 blends were used to assess whether blended wines, which have more chemical complexity, are actually perceived to be complex than single-varietal wines. 87 participants, comprising three expertise levels, tasted all six wines blind, then rated the perceived complexity, liking, intensity, familiarity, quality, and willingness to pay, as well as listing five flavor descriptors that best matched each wine. Finally, participants guessed which, if any, of the wines were blends.

Conclusions: The results demonstrated that wine complexity is a multi-faceted concept that involves the perceived attributes in the wine (quality, intensity, liking) as well as being influenced by participant background (age, gender) and even environmental context (specifically the order in which the wines were tasted). Linguistic analysis on participants' choice of flavour descriptors revealed that greater complexity was associated with winemaking flavours, especially oak. Further analysis demonstrated that complexity was not independently correlated with willingness to pay, but was predicted by a combination of perceived attributes (quality, liking) as well as participant expertise.

Significance of the Study: We demonstrate how complexity and associated willingness to pay are assessed by people of different expertise levels, using both numerical ratings and linguistic analysis. We also test for the first time the role of blending on perceived complexity.

KEYWORDS: COMPLEXITY; BLENDING; QUALITY; WINE; FLAVOUR; NATURAL LANGUAGE PROCESSING; WILLINGNESS TO PAY

1. Introduction

The term complexity is often used to describe quality wines (e.g., Passmore, 2015). In fact, some producers have deliberately aimed to increase the complexity of their wines, in the hope of increasing the wines' consumer appeal (see Parr, 2015). However, complexity remains an ambiguous concept with multiple dimensions (Parr et al., 2011; Schlich et al., 2015). The aim of the present study was to use single-varietal wines and their 50-50 blends to uncover which attributes in wine are most linked to complexity, and whether it is true that the casual wine consumer (or expert) is willing to pay more for more complex wines.

1.1 What kind of complexity are we talking about?

Complexity can potentially come from multiple sources (Holland, 2014). There is chemical complexity, found in the wine itself, where complexity arises from the structural arrangement of the molecules or from the sheer number of different molecules. There is also perceived/inferred complexity, where the focus lies in what the taster takes away from the wine (see Spence & Wang, 2018, for a review). In terms of chemical complexity, there is some evidence to suggest that perceived complexity correlates weakly with the physical complexity of monomolecular odorants (Kermen et al., 2011). In addition, the size of monomolecular odorants has also been found to correlate weakly with the reported pleasantness of the odorant (Khan et al., 2007). Of course, given the hundreds of volatile compounds which exist in wine (Waterhouse et al., 2016), it would be difficult to assess empirically how the physical complexity or size of particular molecules contribute to wine complexity. Alternatively, one can consider the number of different aromatic compounds in wine – logically, the greater number of aromatic compounds should correspond with greater perceived complexity. To the best of our knowledge, such a relationship has yet to be found between the number of compounds in a mixture on the one hand, and the number of perceivable aromas on the other. There is, however, some evidence that such a relationship might resemble an inverse U-curve similar to that found between visual complexity and the number of components elements (e.g., Chipman, 1977). In terms of an upper bound, Weiss and colleagues have found that combinations of 30 or more equi-intense olfactants tend to smell the same, very much like an olfactory equivalent of auditory white (Weiss, Snitz, Yablonka, Khan, Gafsou, Schneidman, & Sobel., 2012). In addition, familiarity can play an important role in the relationship between chemical and perceived complexity. While a single

molecule can trigger multiple aroma descriptors (as used in modern perfumery¹), we have also learned to associate the complex arrays of olfactants with the smell of single objects, such as roses or coffee (Yeshurun & Sobel, 2010).

Perceived complexity can also be considered in multiple ways – for instance, in terms of the number of components perceived, in terms of the temporal evolution of flavours (in the mouth, in the glass, or even in the bottle), or in terms of a holistic integrated percept whose elements may not, in fact, be individuable (see Spence & Wang, 2018, for a review). For instance, some aromatic varieties such as Gewürztraminer or Viognier may have many aromas/flavours perceivable at once (Fielden, 2009). In contrast, an aged red wine might unveil its flavours slowly in the mouth, with primary fruit notes giving way to tertiary flavours (e.g., leather, tobacco) on the finish (Fielden, 2009). Different still, a wine can be considered complex when its components are not individually obvious (Schlich et al., 2015). Finally, it is possible that complexity is simply what people refer to (especially wine novices) when they have trouble analysing a wine, either in terms of identifying elements in a wine (Porcherot & Issanchou, 1998; Ruijschop et al., 2009) or simply describing it in words (Melcher & Schooler, 1996). What has yet to be researched though is how wine tasters internally use these different metrics of judging complexity when they say that a wine is complex.

1.2 The role of blending in complexity

Logically, blending different single-varietal wines should result in a wine with more perceivable complexity as the blend will, at the very least, be more chemically diverse. A study demonstrating the value of such blending was conducted by Singleton and Ough (1962). These researchers selected 34 pairs of similar commercially available single varietal Californian wines from the 1960 vintage that were rated similarly in terms of their quality but which presented somewhat different flavours. From each pair, they also prepared a 50-50 mix of the wines. The participants were presented with the individual wines and also with the blend, though they were blind as to the condition (i.e., individual wine or blend) with the order of presentation randomised. The wines were rated on a 20-point quality scale by the 10 experienced sensory panellists. Notably, the quality scores given to the blends were significantly higher than the mean

¹ For instance, the perfume brand Escentric Molecules, <https://www.escentric.com>.

score of the two wines when evaluated separately. Furthermore, in 7 out of the 34 cases, the blend was rated as having higher quality than the best of the two individual component wines. Singleton and Ough hypothesised that, in at least some cases, an increase in complexity might have been a major factor in enhanced quality ratings, although it remains a hypothesis since the panellists did not rate the complexity of the wines.

1.3 The role of domain-specific expertise on complexity

Parr et al. (2011) developed a model for how people with different levels of expertise assess complexity. Based on interviews with both wine consumers and professionals from New Zealand and Australia, it was found that the casual consumer tends to base their judgments of complexity on the hedonic qualities of the wine as well as on its image/brand (D'Alessandro & Pecotish, 2013; see Plassman et al., 2011, for evidence of perceptual influences of marketing actions). In contrast, wine professionals take into consideration inferred methods of viticulture and wine production. Interestingly, Wang and Spence (2018) demonstrated that, even for social drinkers, flavours associated with the process of winemaking and ageing (such as smoke and vanilla) were more often attributed with those wines perceived to be more complex.

In a more recent study, a series of young Marlborough Sauvignon Blanc wines were tasted by experts, connoisseurs, and consumers, who completed a free-sorting task and a questionnaire designed to assess different attributes (such as harmony, balance, number of components, etc.) which might contribute to the judgement of complexity (Schlich et al., 2015). The authors found more inter-participant agreement amongst experts than the other groups. Furthermore, while experts associated complexity with the number of flavours, harmony, balance, familiarity, and the duration of aftertaste remaining in the mouth; the connoisseurs and consumers associated complexity with flavour intensity.

1.4 Possible correlation between complexity and willingness to pay

As stated previously, it is a widely held belief that complexity is a desirable factor in the quality of wine (Kramer, 2012; Shogren, 2012; Singleton & Ough, 1962). If the New Zealand producers who are experimenting with novel production techniques are correct, more complex wines

should result in greater purchase intent and willingness to pay from consumers. However, as yet, there have been no studies that have looked specifically at the impact of wine complexity on willingness to pay. If quality is viewed as a substitute for complexity, then higher perceived quality has been shown to result in higher willingness to pay by both consumers and experts (Corduas et al., 2013; D'Alessandro & Pecotich, 2013; Mueller & Szolnoki, 2010; Rao & Monroe, 1989; Tellis & Wernerfelt, 1987), although objectively better quality wines are not necessarily preferred by the casual consumer (D'Alessandro & Pecotich, 2013; Goldstein et al., 2008).

1.5 Aims and contributions

The primary aim of the present study is to assess how complexity in wine is perceived by social drinkers and experts, especially which attributes are most linked to their assessment of complexity in a wine. Additionally, in order to ensure that the wines used in the study had different levels of chemical complexity (see Schlich et al, 2015), we used single varietal wines and 50-50 blends of those wines (similar to Singleton and Ough's 1962 study). By gathering participants' ratings of perceived complexity as well as liking and perceived quality, we could therefore test whether complexity (either chemical or perceived) accounts for the increased quality ratings of blended wines as observed by Singleton and Ough. Furthermore, we used natural language processing techniques in order to analyse participants' flavour descriptors, in order to assess their semantic associations with complexity.

2. Methods and materials

2.1 Participants

We recruited 87 participants (38 men, 49 women) were recruited for the study from the University of Oxford Experimental Psychology participant pool and from the Oxford University Blind Tasting Society. Participants were asked to self-identify their wine expertise when signing up for the study. Out of 87 participants, 41 were wine novices (who identified most with the statement "I drink socially but don't know much about wine"), 30 were intermediates (who identified most with the statement "I know which wines I like and have been to some tastings"),

and 16 were experts (who identified most with the statement “I have Wine and Spirits Education Trust (WSET) Level 3 or higher certification or I regularly attend blind tastings”). The experimenter verified the qualifications of those in the expert group to ensure that they in fact met the stated requirements. Participants at different expertise levels were roughly age-matched to control for the effect of age on taste preferences and sensitivity, but the average age of the novices was lower than those of intermediates and experts (see Table 1). Furthermore, the distribution of gender was unequal for the three experience categories, with the ratio of men to women increasing with expertise.

Table 1. Distribution of age and gender for the three different expertise groups in the study.

	Average Age (SD)	Men : Women
Novice (N=41)	25.41 (10.26)	13 : 28
Intermediate (N=30)	31.73 (10.28)	12 : 18
Expert (N=16)	33.38 (15.18)	13 : 3

2.2 Wine samples

A total of six samples were prepared, from three single-varietal red wines and three 50-50 blends of these wines. The single-varietal wines were sourced from the same region (Dr. Frank Winery, Finger Lakes, NY, USA) and all came from the same 2015 vintage. The three single-varietal wines were selected to have similar colour, structure (acid, alcohol, tannin), and price level (see Table 2), but to have different flavours (following the same selection protocol as in Singleton & Ough, 1962). The wines were Cabernet Sauvignon, Merlot, and Cabernet Franc. These are the most typical grape varieties grown in Bordeaux and are commonly blended together as well as made into single varietal wines in the New World (Fielden, 2009).

Table 2. Technical information regarding the three single-varietal wines used in the study

Vintage	Alcohol by Volume	Total Acidity (g/L)	pH	Sugar (g/L)	Oak ageing	Retail price (USD)
---------	----------------------	---------------------------	----	----------------	------------	-----------------------

Dr Frank Merlot	2015	13.0 %	5.9	3.35	< 1	European and American oak, 10 months	21.99
Dr Frank Cabernet Franc	2015	13.0 %	7.5	3.49	< 1	European and American oak, 18 months	21.99
Dr Frank Cabernet Sauvignon	2015	13.0 %	6.9	3.54	2.8	French and European oak, 18 months	24.99

In a pre-test, six trained tasters (members of the Oxford University varsity blind tasting team) evaluated the wines blind. Each taster was asked to choose all of the flavours that they found appropriate for the wines from the list found in the WSET Level 3 Systematic Approach to Wine Tasting guide. The chosen flavours were finalised into a table of 75 descriptors (see Table 3) given to the participants in the study. Each row in the table corresponds to a flavour family, such as red fruit, black fruit, vegetal, oak, etc.

Table 3. List of flavor descriptors given to the participants during the study. Note that column categories were not shown to the participants.

<i>Category</i>	<i>Descriptors</i>
<i>Red fruit</i>	Redcurrant, Cranberry, Raspberry, Strawberry, Red Cherry, Red Plum
<i>Black fruit</i>	Blackcurrant, Blackberry, Bramble, Black Cherry, Black Plum
<i>Dried/Cooked fruit</i>	Fig, Prune, Raisin
<i>Vegetal</i>	Green Bell Pepper, Tomato, Blackcurrant Leaf, Black Olive, Green Olive
<i>Herbal</i>	Eucalyptus, Mint, Medicinal, Lavender, Fennel, Dill
<i>Pungent spice</i>	Black Pepper, White Pepper, Liquorice, Juniper
<i>Dairy</i>	Butter, Cream, Yoghurt
<i>Oak</i>	Vanilla, Toast, Cedar, Charred Wood, Smoke, Cloves, Nutmeg

<i>Kernel</i>	Almond, Coconut, Chocolate, Coffee
<i>Animal</i>	Leather, Meaty, Farmyard
<i>Maturity</i>	Vegetal, Wet Leaves, Forest Floor, Game, Savoury, Tobacco
<i>Mineral</i>	Earth, Petrol, Rubber

2.3 Experiment Design and Procedure

The study uses a mixed design with wine variety and blend as within-participant factors, and wine expertise as the between-participant factor. A total of 10 experimental sessions were held. Tasting sheets were collected and analysed from all sessions.

The study was conducted at the Department of Experimental Psychology at the University of Oxford. Each experimental session lasted for approximately 40 minutes and was held between 5 and 6 pm. Six to twelve participants attended each session. The three 50-50 blends were mixed by the experimenter approximately one hour before the start of the experimental session. Each participant was given a paper questionnaire, a cup of water, and a flight of six wines to taste. The wines were served in 25 mL samples, in six 215 mL ISO glasses (see Figure 1). Each participant received the wines in a different order (determined via a Williams Design Latin Square), and were instructed to taste the wines strictly from left to right, without going back to re-evaluate a wine. After tasting each wine, the participants were asked to choose five descriptors from a table of pre-tested descriptors (see Figure 1) that best fit the wine, as well as assessed the wine in terms of liking, familiarity, complexity, and flavour intensity on a 9-point scale (1=not at all; 9=very much). It was explained to the participants at the start of the experiment that, although the flavour descriptor table was organised in rows, they were free to choose five descriptors from any of the rows, and could choose multiple descriptors from the same row if they so desired. Each participant received the four rating questions (liking, familiarity, complexity, and flavour intensity) in the same order for all wines. The order of the questions was randomised between participants. Finally, the participants evaluated the quality of the wine on a 20-point scale (1=worst quality; 20=best quality, same scale as used by Singleton & Ough, 1962) and estimated their willingness to pay for a full 750 mL bottle of the wine in British pounds (GBP). After they

have assessed all six wines, participants were given a final question, where they were asked to indicate on the questionnaire which if any of the wines were blends.

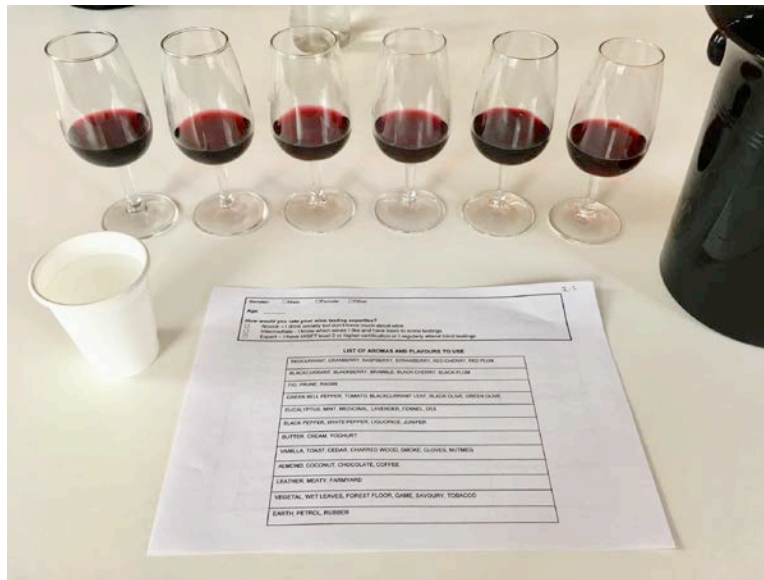


Figure 1. Study setup with paper questionnaire, water cup, flight of six wines ordered from left to right, and spittoon (right).

First, we analysed the participants' guesses for blends in order to assess whether people can tell blended wines from single-varietals. A multivariate analysis of variance (MANOVA) was performed with blend (single varietal or blended) and participant expertise (novice, intermediate, expert) as the between-participant factor to determine whether the blended wines are indeed perceived as more complex and of higher quality. Next, we analysed correlations between perceived complexity and liking, familiarity, quality, willingness to pay, and the numbers of perceived flavours for each group of participants to address the motivating question of whether people are willing to pay more for perceived complexity in wines.

most informative in terms of distinguishing more complex from less complex wines. We also assessed differences in descriptor associations for people at different expertise levels (see Gawel, 1997; Langlois et al., 2011).

3. Results

3.1 Can people tell blends from single varietal wines?

Table 4 illustrates the distributions of yes/no guesses for single varietal wines and blended wines, for all participants and divided by expertise groups. A “yes” guess means a participant thinks the wine is a blend, and “no” if otherwise. Therefore, we would expect that the percentage of yes to no guesses should be higher for the 50-50 blends than for the single varietal wines. Chi squared tests of independence revealed that, overall expertise groups, the distribution of yes-no guesses between single varietal wines and blended wines is not significantly different from random. However, when each group was analysed separately, we observed that novices and intermediates had significantly different patterns of yes-no guesses for single varietal wines versus blended wines. Interestingly, it was the experts who did not have different guessing patterns for single varietal versus blended wines.

Table 4. Distribution of guesses as to whether participants thought a wine was a blend (“yes”) or not (“no”), for all participants and for each expertise group. Chi squared tests of independence were conducted to assess whether the distribution of yes-no guesses were different for single varietal wines versus 50-50 blends.

	Single Varietal Wines		50-50 Blends		χ^2	<i>p</i>
	Yes guesses	No guesses	Yes guesses	No guesses		
Overall (N=87)	105 (40%)	156 (60%)	125 (48%)	136 (52%)	3.11	.08
Novice (N=41)	49 (40%)	74 (60%)	71 (58%)	52 (42%)	7.88	.005
Intermediate (N=30)	37 (41%)	53 (59%)	53 (59%)	37 (41%)	5.69	.02
Expert (N=16)	19 (40%)	29 (60%)	19 (40%)	29 (60%)	0	1

To understand how people would decide whether a wine was a blend, we calculated Spearman correlation coefficients between people's blending guess (yes or no) and complexity, liking, intensity, familiarity, quality, willingness to pay, and the order in which the wines were tasted (see Table 4). The results revealed that the probability of a participant guessing that a wine was a blend was significantly negatively correlated with the rating of familiarity ($r_{521} = -.12$, $p = .005$). In other words, wines that seemed less familiar were more likely to be guessed as blends.

Table 4. Spearman correlation coefficients (N=521) between participants' guesses as to whether the wine is a blend, and complexity, liking, intensity, familiarity, quality, willingness to pay, and the order in which the wine was tasted. * indicates significance at 0.01 level.

	Complexity	Liking	Intensity	Familiarity	Quality	Willingness to pay	Wine order
Blend guess (yes=1, no=0)	.05	-.03	.02	-.12 *	-.05	-.01	-.04

To understand whether single-varietal or blended wines influenced participant ratings, a rm-MANOVA was conducted with blend status (single variety or 50-50 blend) and participant expertise (novice, intermediate, or expert) as between-participant factor, and complexity, liking, intensity, familiarity, quality, and willingness to pay as measures. The results revealed that there was no significant effect of blend status on any ratings ($F(6, 508) = 0.20$, $p = .98$, *Wilks' Lambda* = .99), but there was a significant effect of wine expertise ($F(12, 1018) = 10.86$, $p < .0005$, *Wilks' Lambda* = .79). Further univariate ANOVAs demonstrated that expertise had a significant effect on complexity ($F(2, 513) = 3.08$, $p = .047$, $\eta_p^2 = 0.01$), intensity ($F(2, 513) = 6.39$, $p = .002$, $\eta_p^2 = 0.02$), familiarity ($F(2, 513) = 27.43$, $p < .0005$, $\eta_p^2 = 0.10$), quality ($F(2, 513) = 11.42$, $p < .0005$, $\eta_p^2 = 0.04$), and willingness to pay ($F(2, 513) = 30.37$, $p < .0005$, $\eta_p^2 = 0.11$). Expertise did not have a significant effect on liking ($F(2, 513) = 2.16$, $p = .12$, $\eta_p^2 = 0.01$). Post-hoc comparisons with Bonferroni corrections (see Figure 2) revealed that experts found the wines to be significantly more intense than did the novices. Furthermore, compared to novices and

intermediates, experts found the wines to be more familiar, of higher quality, and they were willing to pay more.

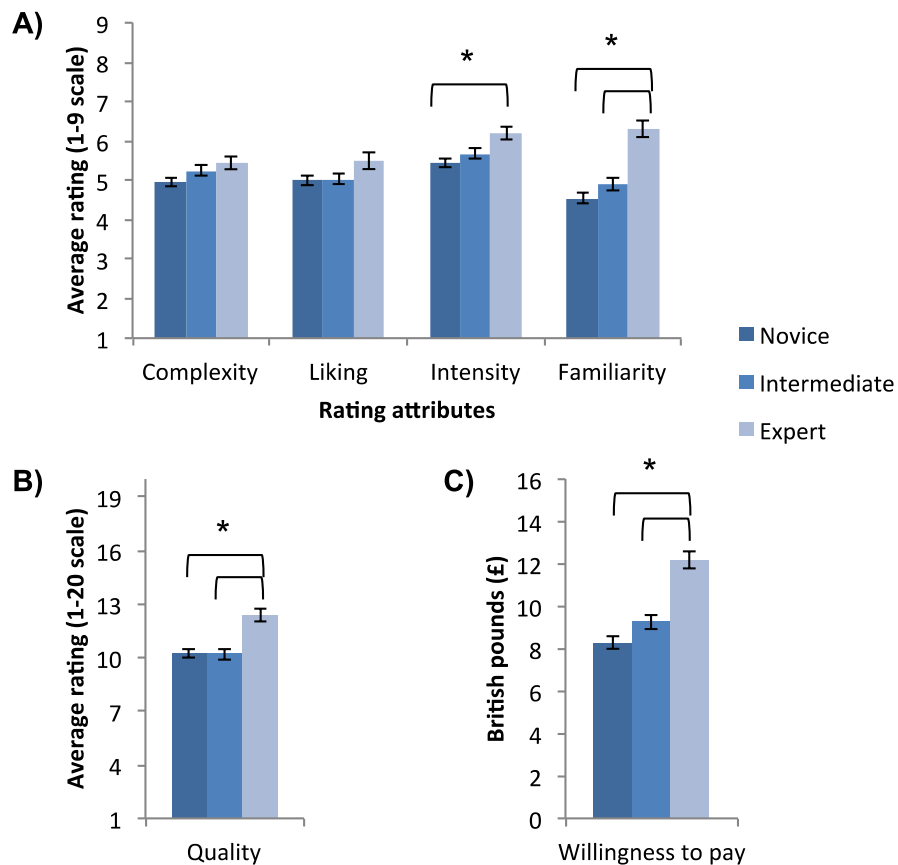


Figure 2. Average ratings for complexity, liking, intensity, familiarity (Figure 2A), quality (Figure 2B), and willingness to pay (Figure 2C) for each expertise group. Error bars indicate standard error of the means. Asterisks ‘*’ indicate significant differences according to a Bonferroni test ($p < .05$).

According to Singleton and Ough’s (1962) original study, the average quality scores of the 50-50 blends were significantly better than the average of the scores of the two constituent wines (and as it follows, also better than the lower scoring of the two constituent wines). Furthermore, “*in no case did a blend receive a lower mean score than the lowest-scoring wine of the pair*” (Singleton & Ough, 1962, p. 195). We therefore re-ran the same calculations on our dataset. However, we did not replicate Singleton and Ough’s results. In fact, the average quality rating of the blends ($Q_{\text{blend}} = 10.69$, $SE = 0.24$) was not significantly different from the quality rating of the lower scoring wine ($Q_{\text{lower}} = 10.31$, $SE = 0.27$, $t(260) = 1.32$, $p = .19$), the higher scoring wine ($Q_{\text{higher}} = 10.84$, $SE = 0.26$, $t(260) = -0.50$, $p = .62$), or the average score of the two

constituent wines ($Q_{\text{average}} = 10.57$, $SE = 0.21$, $t(260)=0.45$, $p = .65$). Table 5 reveals that, in the case of the Cabernet Franc – Merlot pair, the average blend score was actually lower than the lowest-scoring wine of the pair.

Table 5. Average quality ratings for each pair of single-varietal wines and its 50-50 blend. Pairs of wines are listed with the lower scoring wine first, then higher scoring wine. M = Merlot, CF = Cabernet Franc, CS = Cabernet Sauvignon. Standard error of the mean shown in parentheses.

	Wine 1 (lower score)	Wine 2 (higher score)	Blend
CF – M	10.22 (.48)	10.48 (.44)	10.03 (.45)
M – CS	10.48 (.44)	11.02 (.47)	10.90 (.38)
CF – CS	10.22 (.48)	11.02 (.47)	11.13 (.40)

3.2 Understanding Complexity

Initial correlation analysis between complexity, liking, intensity, familiarity, quality, wine order, age, gender, wine expertise, and willingness to pay revealed a significant positive correlation between complexity and liking, intensity, familiarity, quality, willingness to pay, wine order, age, and expertise (see Table 6). However, given that many of these attributes are co-correlated, we then conducted multiple regression analysis with complexity as the dependent measure, in order to understand the effect of each predictor independent of the effects of all other predictors. The results revealed that perceived quality, intensity, liking, participant age, wine order, and participant gender were able to predict the variance in complexity (58%, see Table 7). This reveals that complexity is a nuanced construct that not only depends on the perceived attributes in the wine, but also, to some extent, on participant's background (e.g., age and gender) as well as the environmental context (e.g., the order in which the wine was tasted). That said, stepwise multiple regressions revealed that the majority of the variance (55%) could be explained by perceived quality, intensity and liking.

Table 6. Pearson's correlation coefficients ($N=521$) between complexity, liking, intensity, familiarity, quality, wine order, age, gender, wine expertise, and willingness to pay * indicates significance at .05 level, and ** indicates significance at the .01 level.

	Liking	Intensity	Familiarity	Quality	Wine order	Gender	Age	Expertise	Willingness to pay
Complexity	.57 **	.52 **	.29 **	.66 **	.23 **	.08	.19 **	.10 *	.50 **
Liking	1.0	.26 **	.44 **	.68 **	.11 *	.05	.14 **	.08	.57 **
Intensity		1.0	.21 **	.38 **	.08	-.04	.09 *	.15 **	.31 **
Familiarity			1.0	.34 **	.02	-.16 **	.07	.29 **	.35 **
Quality				1.0	.18 **	.05	.07	.16 **	.65 **
Wine order					1.0	.00	.00	.00	.11 *
Gender						1.0	-.11 *	-.33 **	-.05
Age							1.0	.29 **	.19 **
Expertise								1.0	.31 **

Table 7. Multiple regression analysis with complexity as dependent measure. * indicates significance at 0.05 level, and ** indicates significance at 0.01 level.

	β	t-Stat.
Quality	.35	7.62 **
Intensity	.31	9.93 **
Liking	.19	4.38 **
Age	.12	3.91 **
Wine Order	.11	3.86 **
Gender (1=male, 2=female)	.07	2.28 *
Willingness to pay	.05	1.19
Expertise	-.04	-1.22

Familiarity	.02	.52
Adjusted R²	.58	
Num. Obs.	518	

3.3 Linguistic analysis

A frequency count of the participants' flavour descriptors revealed that black and red fruits were the predominant flavours in the wines (see Table 8). To understand which flavour descriptors are more associated with greater or lesser complexity in wine, we conducted linguistic analysis using the Python Natural Language Toolkit (nltk, Bird et al., 2014). A Naïve Bayes classifier was written to categorise more from less complex wines based on the flavor descriptors chosen (complexity ratings were split according to the median rating of 5/9). The classifier calculates the prior probability of each word occurring in a tasting note categorised as either low or high complexity (Bird et al., 2014, Chapter 6). The prior probability distribution reveals which descriptors might be most informative (i.e., used more for tasting notes belonging to the high complexity group rather than the low complexity group, or vice versa). Table 9 below shows the top 10 most informative words. Note that for the wines used in the present study (i.e., young Bordeaux varieties grown in a cool climate region), greater complexity was associated predominately with flavours that come from winemaking (e.g., coconut, toast, vanilla, butter).

Table 8. Top 5 most frequently used descriptors for the 6 wines tasted in the present study. M = Merlot, CF = Cabernet Franc, CS = Cabernet Sauvignon. Actual frequency count shown in parentheses. There were a total of 435 descriptors provided by 87 participants for each wine.

M		CF	CS	M+CF	CF+CS		M+CS
Cranberry (20)		Cranberry (20)	Vanilla (22)	Cranberry (20)	Blackcurrant (23)		Cranberry (19)
Black (18)	cherry	Vanilla (19)	Blackberry (19)	Red plum (16)	Charred (19)	wood	Vanilla (17)

Red cherry (17)	Redcurrant (19)	Black (18)	cherry	Blackcurrant (16)	Cranberry (17)	Raspberry (16)
Blackberry (15)	Blackcurrant (17)	Black (18)	pepper	Strawberry (14)	Smoke (17)	Blackcurrant (16)
Blackcurrant (15)	Black (17)	cherry	Redcurrant (18)	Red cherry (13)	Redcurrant (15)	Black pepper (16)

Table 9. Top 10 most informative words descriptors indicating whether a wine belongs in the lower or higher complexity group (higher complexity defined as higher than 5 on a 9 point scale), with its related likelihood ratios.

Descriptor	Likelihood	Ratio
Coconut	High : low	4.3 : 1
Farmyard	Low: high	2.9 : 1
Game	Low: high	2.7 : 1
Toast	High : low	2.5 : 1
Vanilla	High : low	2.1 : 1
Butter	High : low	1.9 : 1
Vegetal	Low: high	1.8 : 1
Charred wood	Low: high	1.7 : 1
Red cherry	Low: high	1.7 : 1

To better understand how wine expertise might influence the perception of complexity, we divided the participants into two groups, novices (N=41) and more experienced tasters (N=46, intermediates and experts). A similar linguistic analysis (see Table 10) revealed that beginners used a more consistent vocabulary when describing wines of lower complexity, attributed to animal or vegetal flavours (e.g., game, vegetal, farmyard). On the other hand, more experienced tasters associated riper fruit and new oak flavours with greater complexity (e.g., chocolate, coconut, toast) and unripe flavours (e.g., cranberry, green bell pepper) with less complex wines.

Table 10. Left: Top 10 most informative words descriptors for novice tasters, indicating whether a wine belongs in the lower or higher complexity group, with its related likelihood ratios. Right: top 10 most informative descriptors for experienced tasters.

Novice tasters	Experienced tasters (intermediate+expert)
-----------------------	--

Word	Likelihood	Ratio	Word	Likelihood	Ratio
Game	Low: high	4.7 : 1	Chocolate	High: low	3.3 : 1
Vegetal	Low: high	4.0 : 1	Coconut	High: low	3.0 : 1
White pepper	High: low	4.0 : 1	Cranberry	Low: high	2.9 : 1
Farmyard	Low: high	2.8 : 1	Yoghurt	Low: high	2.8 : 1
Rubber	Low: high	2.7 : 1	Toast	High: low	2.7 : 1
Cedar	Low: high	2.6 : 1	Farmyard	Low: high	2.4 : 1
Yoghurt	High: low	2.5 : 1	Green bell pepper	Low: high	2.4 : 1
Almond	High: low	2.2 : 1	Vanilla	High: low	2.3 : 1
Red cherry	Low: high	2.1 : 1	Fig	Low: high	2.3 : 1

3.4 Are people willing to pay more for complexity?

Initial correlation analysis revealed a significant positive correlation between complexity and willingness to pay ($r_{521} = .50, p < .0005$, see Table 7). However, given that most of the attributes are correlated, we calculated the partial correlation between complexity and willingness to pay, while controlling for the other variables. The results demonstrated that there was not a significant correlation between complexity and willingness to pay when controlling for the other factors ($r_{509} = .05, p = .23$).

A multiple regression analysis (see Table 11) revealed that perceived quality, wine expertise, and wine liking were able to significantly explain the variance (50%) in the participants' willingness to pay. Stepwise multiple regression showed that quality was the most significant predictor of willingness to pay (explaining 43% of the variance).

Table 11. Multiple regression analysis with willingness to pay as dependent measure. * indicates significance at 0.05 level, and ** indicates significance at 0.01 level.

	β	t-Stat.
Quality	.43	8.83 **
Wine expertise	.19	5.23 **
Liking	.20	4.21 **
Age	.06	1.77

Complexity	.06	1.19
Familiarity	.04	.97
Intensity	.02	.55
Wine order	-.003	-.10
Gender (0=male, 1=female)	.003	.09
Adjusted R²	.50	
Num. Obs.	518	

A similar linguistic analysis on the willingness to pay data revealed that top 10 most informative terms associated with higher or lower willingness to pay (higher willingness to pay defined as greater than £9 according to the median) compared to lower willingness to pay (see Table 12). For the wines used in the present study (i.e., young Bordeaux varietals grown in a cool climate region), greater willingness to pay was once again associated with flavours that come from winemaking (e.g., vanilla, butter, chocolate, toast). That said, it is interesting that “yoghurt”, which is also a term commonly associated with winemaking (malolactic fermentation) is associated with lower willingness to pay. Perhaps this is due to the tartness association with yoghurt, which might be undesirable, whereas those flavours associated with ripeness are associated with higher willingness to pay. Once again, Table 13 reveals that the more experienced tasters are the ones with consistent winemaking associations, favouring vanilla, chocolate, and nutmeg. Novices associated fresh fruit and herbaceous characteristics with higher willingness to pay.

Table 12. Top 10 most informative words descriptors indicating whether a wine belongs in the lower or higher willingness to pay group (higher willingness to pay defined as £10 or higher), with its related likelihood ratios.

Descriptor	Likelihood	Ratio
Yoghurt	Low: high	4.3 : 1
Vanilla	High : low	2.9 : 1
Petrol	Low: high	2.7 : 1

Wet leaves	Low: high	2.5 : 1
Butter	High : low	2.1 : 1
Chocolate	High : low	1.9 : 1
Toast	High : low	1.8 : 1
Coconut	High : low	1.8 : 1
Medicinal	Low: high	1.8 : 1
Strawberry	High : low	1.8 : 1

Table 13. Left: Top 10 most informative words descriptors for novice tasters, indicating whether a wine belongs in the lower or higher willingness to pay group, with its related likelihood ratios. Right: top 10 most informative descriptors for experienced tasters.

Novice tasters			Experienced tasters (intermediate + expert)		
Descriptor	Likelihood	Ratio	Descriptor	Likelihood	Ratio
Green olive	High: low	2.8: 1	Yoghurt	Low: high	2.8 : 1
Lavender	High: low	2.8 : 1	Wet leaves	Low: high	2.5 : 1
Smoke	Low: high	2.6 : 1	Vanilla	High: low	2.4 : 1
Blackcurrant leaf	High: low	2.4 : 1	Petrol	Low: high	2.3 : 1
Prune	Low: high	2.2 : 1	Chocolate	High: low	2.2 : 1
Cedar	Low: high	2.2 : 1	Nutmeg	High: low	2.1 : 1
Blackberry	High: low	1.7 : 1	Medicinal	Low: high	2.0 : 1
Chocolate	High: low	1.7 : 1	Green olive	Low: high	1.9 : 1
Redcurrant	High: low	1.7 : 1	Farmyard	Low: high	1.9 : 1
Black pepper	High: low	1.7 : 1	Coconut	High: low	1.9 : 1

4. Discussion

The results of the present study demonstrate that wine complexity is a multi-faceted concept that involves the perceived attributes in the wine (quality, intensity, liking) as well as participant background (age, gender) and even environmental context (order of wines tasted). Linguistic analysis on participants' choice of flavour descriptors revealed that, for the wines used in the study (i.e., young Bordeaux varieties), greater complexity was associated with winemaking flavours, especially the usage of oak. This is similar to the results observed in Wang and Spence (2018), where oak ageing flavours were also correlated with perceived complexity. Moreover, chemical analyses have revealed that wines with higher concentrations of oak ageing related

compounds, such as eugenol, were also judged to be of higher quality by Spanish consumers (Sáenz-Navajas et al., 2015). Of course, it remains to be seen whether this association stems from learned knowledge (for instance, the fact that wines aged in new oak barrels are generally more expensive, due to the cost of the oak), or from the intrinsic appreciation for such flavours (or both).

Further analysis in the present study demonstrated that complexity was not independently correlated with willingness to pay, once other factors such as wine liking and quality were controlled for. Instead, willingness to pay was predicted by an orthogonal combination of perceived attributes (quality, liking) as well as participant expertise. The fact that quality and liking contributed independently to willingness to pay as well as complexity implies that participants were able to separate, at least in part, their subjective enjoyment of the wine from a more objective notion of a wine's quality. Furthermore, the fact that willingness to pay was driven partially by expertise might be explained by the fact that familiarity has been shown to increase liking for foods (Raudenbush & Frank, 1999; Stallberg-White & Pliner, 1999). The experts' greater preference for wines, then, could explain their increased willingness to pay. Alternatively, it is possible that the wine novices in the present study, who are younger than more experienced wine tasters, might simply be used to drinking and paying for cheaper wines.

Contrary to the predictions emerging out of Singleton and Ough's (1962) work, blending the wines did not necessarily make them appear to be of higher quality or more complex. This might have been due to the fact that the wines used were all predominantly red- and black fruit-driven (see Table 9), which might have made it difficult for participants to tell blends apart from single varietal wines. Given this was the case, any non-fruit flavours such as vanilla and charred wood might have been used to distinguish the wines. Certainly, the presence of winemaking-related flavours such as vanilla was associated with higher complexity. However, as these flavours were found in both single variety wines as well as blends, this wasn't sufficient for participants to distinguish the blends from non-blends. In addition, in the original Single and Ough (1962) study, the quality analysis was performed over averages of 34 pairs of wines and their blends, including both red and white wines. In the present study, only three pairs of wines and their blends were used. Interestingly, when given the opportunity to guess whether a wine is blended or not only based on having tasted the wine blind, those with little or no experience guessed "yes

it's a blend" significantly higher percentage of the time for blended wines rather than single-varietal wines. This was not the case for the experts, who guessed "yes it's a blend" equally often for blended wines and single-varietal wines. The failure of tasters to discern blends has also been shown for the case of whisky (Smith et al., 2017), where participants were asked to distinguish blended whiskies from single malts, and Champagne (Harrar et al., 2013), where participants were asked to estimate the proportion of Chardonnay in the wines.

One limitation of the study was that only young Bordeaux varieties were tested, making the results difficult to generalise for a study of complexity in general. Unlike in the original Singleton and Ough (1962) study, where 34 pairs of wines were used with sometimes unusual combinations, both blends and single varieties wines of Bordeaux grapes (Merlot, Cabernet Franc, Cabernet Sauvignon) are commonly available. Therefore, blends in the Singleton and Ough study might have been easier to distinguish due to their lack of familiarity, whereas all the wines in the present study might have tasted familiar to participants. It was certainly the case that the experts, who rated all the wines to be significantly more familiar than did the novices and intermediates, also performed the worst in terms of guessing whether the wines were blends or single varietals. Furthermore, it can be argued that blended wines are rarely 50-50 blends – instead, famous blended wines such as Champagne, and Port commonly involve a complex mixture of wines from different plots, vintages, winemaking styles, etc. (Fielden, 2009). Therefore, using 50-50 blends of young red wines as a way to induce complexity might be overly simplistic and lack ecological validity when compared to the real world of fine wine.

Another drawback of the present study is that we assessed participants' willingness to pay without taking into consideration their purchasing power. Granted, age, gender, and expertise were factored into our pricing model, which somewhat takes into account people's spending habits on wine. Our participant base was comprised of current Oxford residents, which while not ideal in terms of participant diversity, might make our results more consistent in terms of purchasing power considerations. However, we do not take into account the inequality in age and gender between the three groups of participants (for instance, the ratio of men to women were very different in the three expertise groups, see Table 1). This could have skewed our results as it is known that there exist sex/gender and age differences in terms of olfactory sensitivity (Brand & Millot, 2001; Hummel et al., 2003) and identification (Larsson et al., 2000).

Finally, the present study highlights the importance of personal factors – especially wine expertise – in wine evaluation. In general, experts found the wines more intense, more familiar, of higher quality, and were willing to pay more for the wines as compared to novices. Moreover, unlike novices, experts associated oak ageing flavours such as vanilla and chocolate with both higher complexity and higher willingness to pay. This has important implications for growing wine markets, where wine preferences and purchasing patterns might shift with an increasingly knowledgeable and experienced consumer population.

ACKNOWLEDGMENTS

The authors would like to thank the Leverhulme International Network Grant entitled *Evaluating Methods of Aesthetic Enquiry across Disciplines* (IN-2015-016) for stimulating and provocative discussion on the topic of complexity. We would also like to thank the Dr. Frank winery for generously providing the wines for the study.

REFERENCES

- Brand, G., & Millot, J.-L. (2001). Sex differences in human olfaction: Between evidence and enigma. *The Quarterly Journal of Experimental Psychology*, **54B**, 259-270.
- Chipman, S. F. (1977). Complexity and structure in visual patterns. *Journal of Experimental Psychology: General*, **106**, 269-301.
- Corduas, M., Cinquanta, L., & Ievoli, C. (2013). The importance of wine attributes for purchase decisions: A study of Italian consumers' perception. *Food Quality and Preference*, **28**, 407-418.
- D'Alessandro, S., & Pecotish, A. (2013). Evaluation of wine by expert and novice consumers in the presence of variations in quality, brand, and country of origin cues. *Food Quality and Preference*, **28**, 287-303.
- Fielden, C. (2009). *Exploring the world of wines and spirits*. London, UK: Wine & Spirit Education Trust.
- Gawel, R. (1997). The use of language by trained and untrained experienced wine tasters. *Journal of Sensory Studies*, **12**, 267-284.
- Goldstein, R., Almenberg, J., Dreber, A., Emerson, J. W., Herschkowitsch, A., & Katz, J. (2008). Do more expensive wines taste better? Evidence from a large sample of blind tastings. *Journal of Wine Economics*, **3**, 1-9.
- Harrar, V., Smith, B., Deroy, O., & Spence, C. (2013). Grape expectations: How the proportion of white grape in Champagne affects the ratings of experts and social drinkers in a blind tasting. *Flavour*, **2**:25.
- Holland, J. H. (2014). *Complexity: A very short introduction*. Oxford, UK: Oxford University Press.
- Hummel, T., Futschik, T., Frasnelli, J., & Hüttenbrink, K.-B. (2003). Effects of olfactory function, age, and gender on trigeminally mediated sensations: A study based on the lateralization of chemosensory stimuli. *Toxicology Letters*, **140-141**, 273-280.

Kermen, F., Chakirian, A., Sezille, C., Joussain P., Le Goff, G., Ziessel, A., Chastrette, M., Mandaïron, N., Didier, A., Rouby, C., & Bensafi, M. (2011). Molecular complexity determines the number odour notes and the pleasantness of smells. *Scientific Reports*, **1**:206.

Khan, R. M., Luk, C.-H., Flinker, A., Aggarwal, A., Lapid, H., Haddad, R., & Sobel, N. (2007). Predicting odor pleasantness from odorant structure: Pleasantness as a reflection of the physical world. *Journal of Neuroscience*, **27**, 10015-10023.

Kramer, M. (2012). How to really taste wine. *Wine Spectator*, **December 18th**.
<http://www.winespectator.com/webfeature/show/id/47792>.

Langlois, J., Dacremont, C., Peyron, D., Valentin, D., & Dubois, D. (2011). Lexicon and types of discourse in wine expertise: The case of *vin de garde*. *Food Quality and Preference*, **22**, 491-498.

Larsson, M., Finkel, D., & Pedersen, N. L. (2000). Odor identification: Influences of age, gender, cognition, and personality. *Journal of Gerontology: Psychology Sciences*, **55B**, 304-310.

Melcher, J., & Schooler, J. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language*, **35**, 231-245.

Mueller, S., & Szolnoki, G. (2010). The relative influence of packaging, labelling, branding and sensory attributes on liking and purchase intent: Consumers differ in their responsiveness. *Food Quality and Preference*, **21**, 774-783.

Parr, W. V. (2015). Unravelling the nature of perceived complexity in wine. *Practical Winery & Vineyard*, **January**, 5-8.

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

Passmore, N. (2015). Six great wines to pair with sushi. *Forbes*, **August 6th**.
<https://www.forbes.com/sites/nickpassmore/2015/08/06/six-great-wines-to-pair-with-sushi/#51d67abc655b>.

Porcherot, C., & Issanchou, S. (1998). Dynamics of liking for flavoured crackers: Test of predictive value of a boredom test. *Food Quality and Preference*, **9**, 21-29.

Rao, A. R., & Monroe, K. B. (1989). The effect of price, brand name, and store name on buyers' perceptions of product quality: An integrative review. *Journal of Marketing Research*, **36**, 351-357.

Raudenbush, B., & Frank, R. A. (1999). Assessing food neophobia: The role of stimulus familiarity. *Appetite*, **32**, 261-271.

Ruijschop, R. M. A. J., Boelrijk, A. E. M., Burgering, M. J. M., de Graaf, C., & Westerterp-Plantenga, M. S. (2009). Acute effects of complexity in aroma composition on satiation and food intake. *Chemical Senses*, **35**, 91-100.

Sáenz-Navajs, M.-P., Avizcuri, J.-M., Ballester, J., Fernandez-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**, 400-411.

Schlich, P., Maraboli, M., Urbano, C., & Parr, W. V. (2015). Perceived complexity in Sauvignon blanc wines: Influence of domain-specific expertise. *Australian Journal of Grape and Wine Research*, **21**, 168-178.

Shogren, E. (2012). Thank the simple wasp for that complex glass of wine. *The Salt (NPR)*, August 31st. <http://www.npr.org/sections/thesalt/2012/08/02/157606554/thank-the-simple-wasp-for-that-complex-glass-of-wine>.

Singleton, V. L., & Ough, C. S. (1962). Complexity of flavour and blending of wines. *Journal of Food Science*, **27**, 189-196.

Smith, B. C., Sester, C., Ballester, J., & Deroy, O. (2017). The perceptual categorisation of blended and single malt Scotch whiskies. *Flavour*, **6**:5.

Spence, C., & Wang, Q. (J.) (2018). On the meaning(s) of complexity in the chemical senses. *Chemical Senses*, **43**, 451-461.

Stallberg-White, C., & Pliner, P. (1999). Effect of flavor principles on willingness to taste unfamiliar foods. *Appetite*, **33**, 209-221.

Tellis, G. J., & Wernerfelt, B. (1987). Competitive price and quality under asymmetric information. *Marketing Science*, **6**, 240-253.

Wang, Q. J., & Spence, C. (2018). Wine complexity: An empirical investigation. *Food Quality & Preference*, **68**, 238-244.

Waterhouse, A. L., Sacks, G. L., & Jeffery, D. W. (2016). *Understanding wine chemistry*. Hoboken, NJ: Wiley-Blackwell.

Yeshurun, Y., & Sobel, N. (2010). An odor is not worth a thousand words: From multidimensional odors to unidimensional odor objects. *Annual Review of Psychology*, **61**, 219-241.