

RUNNING HEAD: WINE COMPLEXITY

Wine Complexity:
An Empirical Investigation

Qian Janice Wang & Charles Spence

Crossmodal Research Laboratory, Oxford University

WORD COUNT: 6500 WORDS

DATE: FEBRUARY 2018

CORRESPONDENCE TO:

Dr. Qian Janice Wang, Department of Experimental Psychology, New Radcliffe
House, University of Oxford, Oxford, OX2 6BW, UK.

E-mail: qian.wang@psy.ox.ac.uk

ABSTRACT

Complexity is a term that is often invoked by people when writing appreciatively about the taste, aroma/bouquet, and/or flavour of wine. However, it is not clear what exactly wine complexity refers to. The present study was designed to uncover which attributes are most strongly linked to the social drinker's perception of complexity in wine. Notably, unlike previous studies of wine complexity, we assessed the temporal component of complexity by acquiring information from participants at the various stages of smelling, tasting, and aftertaste. Furthermore, natural language processing techniques were used to analyse participants' flavour descriptors in order to assess their semantic associations with complexity. Eight wines, chosen for their ability to showcase various aspects of complexity, were tasted in three flights, grouped by dry white, red, and sweet wines. Participants rated the perceived liking, quality, and complexity of each wine, as well as listing flavours of the wines perceived at different stages (aroma, in-mouth, post-swallowing). The results demonstrated that complexity was positively correlated with liking and with quality, but not with the price of the wines or the number of flavours detected. Furthermore, semantic analysis revealed that participants used more consistent vocabulary describe wines that they perceived to be more complex. We also observed similar consistency trends for wines that were liked more, as well as wines rated to be lower quality. In general, secondary and tertiary flavours (derived from fermentation or from ageing) were more often used to describe more complex wines. These results reveal intriguing patterns in how social drinkers assess perceive/infer wine complexity, as well as elucidating the relationship between complexity, quality, and liking.

KEYWORDS: COMPLEXITY; QUALITY; WINE; FLAVOUR; NATURAL LANGUAGE PROCESSING

1. Introduction

The term complexity is often used to describe some desirable attribute of a wine. For instance, just take the following quote from long-time wine critic Matt Kramer:

“The single greatest standard used in assessing the quality of a wine is complexity. The more times you can return to a glass of wine and find something different in it—in the bouquet, in the taste—the more complex the wine. The very greatest wines are not so much overpowering as they are seemingly limitless.” (Kramer, *Wine Spectator*, 2012).

The belief that complexity is a positive quality has driven some wine producers to explicitly try to develop a more complex product offering so as to appeal more to consumers. For instance, Parr (2015) describes one such commercial winemaking project that deliberately aimed at increasing complexity in New Zealand Sauvignon Blanc wines via innovative grape-growing and winemaking practices. However, what exactly does ‘complexity’ mean when it comes to wine (or, for that matter, any other drink)?

Intuitively, complexity in perception must come from multiple elements. However, we could be talking about physical/chemical complexity, where complexity arises from the structural arrangement of the individual molecules (in other words, the nature, connectivity, and orientation of its component elements¹) or from the number of different molecules. Complexity could also refer to perceived/inferred complexity, where the focus lies in what the taster takes away from the wine based on sensory elements that are perceived within it.

In terms of physical complexity, there is some evidence at least that perceived complexity seems to correlate weakly with the physical complexity of monomolecular odorants (Kermen et al., 2011). However, given that wine is made up of many hundreds of different volatile compounds, this certainly is not the appropriate level at which to address wine complexity. Alternatively, one can consider the number of different volatile aromatic compounds to be found in wine – intuitively, the greater number of aromatic compounds, the greater the perceived complexity. However, there

¹ In fact, chemists have developed a molecular complexity index taking into account both the elements that make up the molecule, and structural features of the molecule such as its symmetry and the number and types of bonds of its component elements (Hendrickson et al., 1987).

is evidence to suggest that we can only detect a limited number of aromas (around three) in mixtures (Jinks & Laing, 1999; Laing et al., 2002; Marshall et al., 2006). In addition, there is no simple linear mapping from physical/chemical complexity to perceived complexity, since what smells “like a rose” consists of a complex array of compounds, and what smells complex may consist of a single molecule (Sell, 2006; Yeshurun & Sobel, 2010). Therefore, in practical terms and according to the literature, what people are concerned with when speaking of complexity in the world of wine (or other flavour experiences) would seem instead to be a notion of complexity that is *inferred*, by the taster, from elements that are perceived in the wine.

That said, inferred complexity can be thought of in multiple ways – for instance, in terms of the number of components perceived, in terms of the temporal evolution of flavours in the mouth (or in the bottle), or in terms of a holistic integrated percept whose elements may not be individuable. For instance, a heady Gewürztraminer or Viognier gives rise to many flavours that are perceivable at once, whereas a Chablis might give rise to a sensation of seamless minerality rather than any specific range of flavours (Robinson & Harding, 2015). Different still, an aged claret might unveil its flavours slowly in the mouth, with bright acidity and blackcurrants slowly giving way to leather and cigar box flavours later on. And, over a longer timescale, the same claret, if left in the cellar, will likely develop and take on different characteristics as it ages in the bottle. Might all of these different ways of perceiving a wine give the same resulting inference concerning complexity in the mind of the taster?

It remains unclear in exactly which way wine writers refer to complexity when talking about wine, although from Matt Kramer’s quote (see above), he at least would seem to be referring to a combination of complexity in terms of temporal evolution in addition to having many different flavours. Or take Master of Wine Alex Hunt, who defines complexity as having multiple flavours as well as flavour diversity, but also refers to a kind of “intricate subtlety, requiring sustained mental effort to grasp, as with a complex puzzle²” (Oxford Companion to Wine). Adopting a more scientific (rather than anecdotal) approach, Parr et al. (2011) developed a model for how people with different levels of expertise think about complexity in wine (see Figure 1).

² Along the lines of mental difficulty, Snitz and colleagues (2016) have proposed a novel, robust, and quantitative method for measuring intricacy (a related but not exchangeable term with complexity) that depends on more intricate stimuli evoking a larger variance in the response of observers.

Compiling interviews with both wine consumers and professionals from New Zealand and Australia, these researchers concluded that, in general, the casual consumer is driven more by the image/brand (D'Alessandro & Pecotish, 2013; see also Plassman et al., 2011, for evidence of perceptual influences of marketing actions) and hedonic qualities of the wine, whereas the wine professional tends to give more weight to inferred methods of viticulture and wine production.

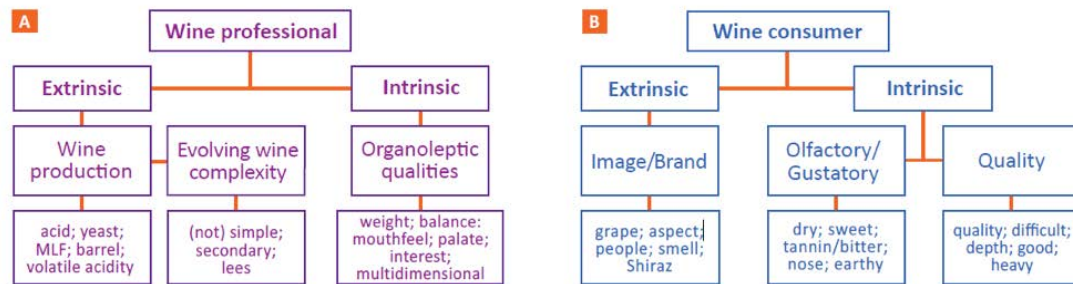


Figure 1. Dimensions of inferred/perceived complexity in wine as a function of the level of wine expertise. MLF = malolactic fermentation. [Figure reprinted from Parr et al. (2011, p. 657).]

So what, exactly, are wine professionals taught about complexity? For a professional trade perspective, one need look only at the Wine and Spirits Education Trust, a globally accredited provider of education and qualifications in both wine and spirits. According to the WSET, complexity can either result from fruit character alone – when the flavours span multiple categories such as floral, herbaceous, citrus fruit, stone fruit, etc. – or from a combination of primary (fruit-based), secondary (from wine-making), or tertiary (from bottle ageing) aromas (WSET Level 4 Diploma candidate assessment guide). This follows Parr et al.’s (2011) model of complexity (see Figure 1), where the wine professional is trained to think about how the flavours are linked to wine production (i.e., yeast, lees, MLF, barrel, volatile acidity) as well as intrinsic qualities in the wine.

In terms of the actual tasting experience, Schlich and colleagues (2015) addressed the role of domain-specific expertise in ratings of the perceived complexity of Sauvignon Blanc wines. 13 New Zealand Sauvignon Blancs (including 10 from the aforementioned complexity innovation winemaking program) were evaluated by experts (oenologists in this case), connoisseurs (non-professionals with a great deal of

experience in wine tasting), and consumers via a free sorting task and a complexity questionnaire. The authors found that, while experts associated complexity with the number of flavours, with harmony, with balance, with the length of finish (duration of aftertaste remaining in the mouth), and with familiarity; connoisseurs and consumers correlated complexity with intensity. Additionally, there was evidence that experts had closer agreement amongst themselves in terms of their evaluation of complexity as compared to either connoisseurs or consumers. Interestingly, there was no evidence that the ease of identifying individual flavours in a wine enhanced the perceived complexity, seemingly adding support to the holistic notion of complexity, one that goes beyond the perception of multiple individual elements. One drawback of this study, however, is that the wines were not perceived as particularly complex to begin with by the French participants who took part.

In the present study, we set out to assess how complexity in wine is perceived by social drinkers, especially which attributes are most linked to their assessment of complexity in wine. We used three flights of wines with white wines, red wines, and dessert wines, at various ages and prices, to showcase the spectrum of wine complexity. By means of this selection, we also hoped to ensure that, unlike with the Schlich et al.'s (2015) study, the participants would find at least some of the wines reasonably complex. Notably, different from previous studies on wine complexity, we aimed to get a sense of the temporal component of complexity by acquiring information at the various stages of smelling, tasting, and aftertaste. Furthermore, natural language processing techniques were also used to analyse participants' flavour descriptors, in order to provide a preliminary assessment of any semantic associations they had with complexity.

2. Methods and materials

2.1 Participants

18 participants (7 female, 11 male) of age 28-62 years ($M=46.2$, $SD=11.3$) took part at the Complexity and Wine tasting event as part of the Leverhulme International Network 'Evaluating Methods of Aesthetic Enquiry across Disciplines' workshop held at Somerville College, Oxford, UK. The participants were researchers in art

history, music, psychology, philosophy, and neuroscience. All of the participants gave their informed consent to take part in the study.

2.2 Wines

A flight of 8 different wines, grouped into 3 flights, were selected to showcase different aspects of wine complexity (see Table 1, see Table 5 for price information). The first flight consisted of three white wines designed to showcase different interpretations of complexity, from many aromas present at once, to the idea of complexity as a single percept (Parr, 2015), to the disjoint nose and palate. The second flight of three red wines demonstrate a single variety, the same variety in a blend, and the same blend with aging (see Singleton & Ough, 1962). The final flight (pair) of sweet wines were chosen to showcase the effect of modern versus traditional winemaking on the same type of grapes (from the same winery, in fact). All wines, as 30 mL samples, were served in standard 270 mL wine glasses, at room temperature (20° C).

Flight #	Wine #	Function	Wine
1	1	Many aromas/flavours at a given moment in time	Domaines Princes Schlumberger Les Abbés Gewürztraminer 2013
	2	Many aromas/flavours at a given time, but somehow harmonised together into a single note	Jean-Marc Brocard Chablis Premier Cru Vaucoupin 2014
	3	Example of complexity from disjoint nose and palate. A wine that smells sweet but tastes dry.	Filipa Pato Branco 2016
2	4	Young single grape variety, simple flavours and not much development on the palate	Château Vartely 2016
	5	Young blended grape varieties, complex primary/secondary flavours and some development on the palate over time	Château Cantemerle 2012
	6	Aged blended grape varieties, complex primary/secondary/tertiary flavours and much	Château Cantemerle 2001

development on the palate over time		
3	7	Modern "clean" winemaking of partly botrytised grapes Disznókő Tokaji Late Harvest 2013
	8	Traditional style of winemaking: barrel aging of partly botrytised grapes Disznókő Tokaji 1413 Szamorodni 2012

Table 1. Wines used for the study and their function in demonstrating different aspects of wine complexity in a tutored wine tasting.

2.3 Procedure

Each participant was presented with the wines, one flight at a time. All of the wines from each flight were served at once, although participants tasted and rated the wines sequentially. All of the participants tasted the wines in the same order. An aroma wheel (Noble et al., 1984) was supplied to each participant in order to provide aroma/flavour notes, if needed.

A paper-and-pencil questionnaire was used. All of the wines were first tasted blind (i.e., without the participants seeing the bottle and knowing what the wine was, ahead of time), and only disclosed after the questionnaire for the flight had been collected. The participants provided their self-rated level of wine expertise on a scale of 3 (1=novice, 2=intermediate, 3=expert) before the tasting began. For flight 1 (wines 1-3), the participants were asked to rate how much they liked the wine, the aromas and flavours perceived, the quality of the wine, and the complexity of the wine. Liking and complexity were presented as a series of 9 boxes (see Figure 2), and quality was presented as four options from the lowest “adequate” to the highest “an excellent wine of great character”. The quality ratings were taken from Decanter World Wine Awards’ (DWWA’s) descriptions for seal of approval, bronze, silver, and gold metal qualifications.

How much do you like the wine?

dislike ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ like
 extremely neither like nor dislike extremely

What aromas do you perceive on the nose?

What flavours do you perceive in the mouth?

Please rate the quality of the wine:

☐ Adequate
☐ A well-made, straightforward and enjoyable wine
☐ A very accomplished wine
☐ An excellent wine of great character

How complex is the wine?

extremely ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ extremely
 simple complex

Figure 2. Sample questionnaire used for flight 1.

For flights 2 and 3, in addition to the liking, quality, and complexity questions from flight 1, the participants were also asked to note on a timeline at which point in time they perceived different flavours in the wine (see Figure 3).

How much do you like the wine?

Dislike extremely ☐ Dislike moderately ☐ Neither like nor dislike ☐ Like moderately ☐ Like extremely ☐

On the timeline below, please note when you taste which flavour:

Drink Swallow

Please rate the quality of the wine:

☐ Adequate ☐ A well-made, straightforward and enjoyable wine ☐ A very accomplished wine ☐ An excellent wine of great character

How complex is the wine?

Extremely simple ☐ Moderately complex ☐ Extremely complex ☐

Figure 3. Sample questionnaire used for flights 2 and 3.

The study lasted for around 45 minutes and the participants were debriefed afterwards.

3. Results

All except one participant (N=17) reported novice level of wine expertise (the remaining participant reported an intermediate level).

3.1 Overall complexity, liking, quality, number of flavours

Pearson's correlation analysis revealed significant correlations amongst complexity, liking, quality, and number of flavours detected (see Table 2). Notably, complexity was correlated with liking, quality, and with the number of flavours that were detected.

	Complexity	Liking	Quality	Number of flavours
Complexity	1	.62*	.53*	.26*
Liking		1	.36*	.23*
Quality			1	.12
Number of flavours				1

Table 2. Pearson's correlation values between pairs of wine attributes (complexity, liking, quality, and number of flavours). Asterisks "*" indicate significance at $p < .01$ level.

Furthermore, partial correlation coefficients were calculated for those pairs of attributes found to be significant, while controlling for all other attributes (see Table 3). The results suggest that the influences of liking and quality on complexity are independent, and that liking and quality are orthogonal attributes.

Pairs of attributes	Partial correlation coefficient	Controlling for:
Complexity & liking	.51*	Quality, number of flavours
Complexity & quality	.42*	Liking, number of flavours
Complexity & number of flavours	.15	Liking, quality
Liking & quality	.06	Complexity, number of flavours
Liking & number of flavours	.10	Quality, complexity

Table 3. Partial correlation values between pairs of wine attributes (complexity, liking, quality, and number of flavours). Asterisks “*” indicate significance at $p < .01$ level.

3.2 Role of wine type – dry white, red, or sweet?

A multivariate analysis of variance (MANOVA) with wine type (dry white, red, or sweet) as the dependent factor and liking, quality, complexity, and number of flavours as measures revealed an overall significant effect of wine type ($F(8,230) = 7.09, p < .0005$, Wilks' $\Lambda = .62$). More significantly, wine type had a significant effect on liking ($F(2,117) = 6.88, p < .005, \eta_p^2 = 0.15$), complexity ($F(2,117) = 14.34, p < .005, \eta_p^2 = 0.20$), and on the number of flavours that were reported ($F(2,117) = 16.39, p < .005, \eta_p^2 = 0.22$). In general, fewer flavours were detected in the dry white wines and they were liked less than red and sweet wines (see Table 4). In addition, the red wines were perceived as being more complex than the dry white wines. There was no significant difference in quality for the different wines ($F(2,117) = 1.56, p = .21$).

	Wine	Mean	Std error	Comparison
Complexity (1-9 scale)	Dry white	4.35	.27	<i>a</i>
	Red	6.40	.27	<i>b</i>
	Sweet	5.30	.39	<i>ab</i>
Liking (1-9 scale)	Dry white	4.90	.29	<i>a</i>
	Red	6.56	.30	<i>b</i>
	Sweet	6.70	.43	<i>b</i>
Quality (1-4 scale)	Dry white	2.39	.14	<i>a</i>
	Red	2.73	.14	<i>a</i>
	Sweet	2.65	.20	<i>a</i>
Number of flavours	Dry white	2.45	.19	<i>a</i>
	Red	3.94	.19	<i>b</i>
	Sweet	3.65	.28	<i>b</i>

Table 4. Average ratings and standard errors of wine attributes, divided by wine type (dry white, red, sweet). Within each rating group (complexity, liking, quality, number of flavours), samples that do not share a letter are significantly different from each other ($p < .05$).

3.3 Role of wine price - Do people rate more expensive wines as more complex?

The prices of the wines were also incorporated into the analysis in order to assess whether wine price was correlated with any reported wine attributes. For each wine, the average values of complexity, liking, quality, and number of flavours were calculated (see Table 5). Pearson's correlation coefficients were then calculated between these average ratings and the price of the wines. Price correlated significantly with perceived quality ($r_s = .78, p = .02$) but not with complexity ($r_s = .53, p = .17$), liking ($r_s = .28, p = .50$), or the numbers of flavours ($r_s = .26, p = .53$).

Wine #	Complexity (SE)	Liking (SE)	Quality (SE)	Number of flavours (SE)	Price (£)	Source
1	5.24 (.40)	6.94 (.35)	2.29 (.19)	2.29 (.22)	14.50	The Wine Society
2	3.47 (.51)	3.35 (.47)	2.76 (.41)	2.59 (.30)	19.50	The Wine Society
3	4.33 (.53)	4.33 (.45)	2.07 (.15)	2.47 (.29)	9.95	The Wine Society
4	6.19 (.37)	6.50 (.47)	2.44 (.18)	4.25 (.34)	5.95	The Wine Society
5	5.94 (.43)	5.88 (.44)	2.56 (.22)	3.69 (.35)	29.50	Berry Bros. & Rudd
6	7.06 (.32)	7.31 (.40)	3.19 (.16)	3.88 (.40)	50.00	Berry Bros. & Rudd
7	5.36 (.59)	6.71 (.57)	2.57 (.27)	3.71 (.41)	15.99	Oxford Wine Company

8	5.22 (.74)	6.67 (.67)	2.78 (.22)	3.56 (.56)	14.99	Ocado
---	------------	---------------	---------------	---------------	-------	-------

Table 5. Average values (with standard error in parentheses) of complexity, liking, quality, and number of perceived aromas for each wine in the study, as well as the price and place of purchase.

3.4 Does the number of aromas vs. flavours predict complexity?

In the first flight of white wines, the participants had to report both what they smelled in the glass (aromas) and what they tasted in the mouth (flavours). There was a significant correlation between the number of aromas perceived and the number of flavours perceived ($r_{51} = .54, p < .005$). However, for the white wines, neither the number of aromas ($r_{48} = .13, p = .37$, controlling for number of flavours), nor the number of flavours ($r_{48} = -.26, p = .07$, controlling for number of aromas), correlated with complexity.

3.5 Do mid-palate vs. post-swallow flavours predict complexity?

In the second flight (of red wines) and the third flight (of sweet wines), the participants had to report at which point in time they experienced the flavours (see Figure 2). The reported flavours were then divided into those experienced in the mouth before swallowing, and those experienced after swallowing. The number of flavours perceived before swallowing was not correlated with the number of flavours perceived after swallowing ($r_{85} = -.11, p = .30$). Interestingly, complexity was significantly correlated with the number of flavours perceived before swallowing ($r_{85} = .29, p = .008$), but not after swallowing ($r_{85} = .03, p = .79$).

3.6 Are certain flavour descriptors associated with greater complexity?

Flavour descriptors were transcribed for each wine. For each participant, each tasting note (i.e., the list of descriptors for each wine) was categorised as either low complexity (rating of 6 or less) or high complexity (rating of 7 or more), with the split point determined by the median complexity rating. A Naïve Bayes classifier was written using the Python Natural Language Toolkit (nltk, Bird et al., 2014) to

categorise more from less complex wines based on the tasting note provided. Here it is worth noting that despite its simplicity, the Naïve Bayes classifier is widely used for text classification applications such as email spam detection and document categorisation (Kibriya et al., 2005). The classifier calculates the prior probability of each descriptor occurring in a tasting note categorised in either the low or high complexity group (Bird et al., 2014, Chapter 6). The classifier would then use this prior probability information in order to estimate the likelihood of a new tasting note belonging to the low or high complexity group. For the sake of the present study, however, we were only interested in calculating the prior probability of the descriptor to uncover which descriptors might be most informative (i.e. used for tasting notes belonging in one complexity group more than another). Each participant provided 8 tasting notes, which gave a total of 144 tasting notes for training the Naïve Bayes model. While this is admittedly a rather small sample size, Naïve Bayes is relatively robust when dealing with small sample sizes since all it requires is the calculation of prior probabilities (Bo et al., 2016; Domingos & Pazzani, 1997; Koller & Sahami, 1997).

The Naïve Bayes classifier revealed which flavour descriptors were most informative, in terms of distinguishing more complex from less complex, or simpler, wines. In other words, some flavour descriptors appeared much more frequently in complex wines as compared to simpler wines, or vice versa, which makes their appearance more valuable in terms of predicting which category (complex vs. simple) a given tasting note is likely to fall in to.

Table 6 below shows the top 10 most informative flavour descriptors. Note that “smoke” was the most indicative flavour term for a complex wine, with odds of 7.6:1 for describing a wine as complex rather than simple. In other words, the term “smoke” was 7.6 times more likely to appear in the tasting note for a wine that was judged to be relatively complex, than a wine that was judged to be relatively simple.

Flavour descriptor	Likelihood	Ratio
Smoke	Complex : simple	7.6 : 1
Pepper	Complex : simple	3.2 : 1
Honey	Complex : simple	3.1 : 1
Grape	Complex : simple	2.9 : 1
Pungent	Complex : simple	2.9 : 1
Liquorice	Complex : simple	2.9 : 1

Jam	Complex : simple	2.9 : 1
Meat	Complex : simple	2.9 : 1
Orange	Complex : simple	2.9 : 1
Citrus	Simple : complex	2.8 : 1

Table 5. Top 10 most informative flavour descriptors from participants' tasting notes of all 8 wines, that indicates whether a wine is 'more complex' or 'more simple', with its related likelihood ratios. For instance, "smoke" is 7.6 times more likely to be used to describe a wine that was rated as 'more complex' than a wine rated 'more simple'.

Interestingly, by observing the relative likelihood ratios of a flavour being used to describe complex or simple wines, it can be seen that the most informative words are used to classify more complex rather than simpler wines. This implies that people use more consistent language when describing a more complex wine as compared to a simpler wine. Of course, as seen in Section 3.2, red wines and sweet wines (at least the ones chosen in the tasting) were rated in general to be more complex than white wines, and the most indicative words are indeed the same ones used to describe those particular wines. To ensure a fair comparison, we then only analysed the descriptors used for the three red wines (see Table 6), which are related in that they all contain at least a major proportion of the Cabernet Sauvignon grape. Interestingly, the top three descriptors associated with complexity are associated with oak ageing, which makes sense as winemaking introduces an additional dimension of flavours not found in the grapes themselves. It also agrees with the model put forward by Parr et al. (2011) (see Figure 1) as well as WSET guidelines where the effect of barrel fermentation, as a method of wine production, adds to perceived complexity. Additionally, 7 out of the top 10 most informative descriptors (i.e., those descriptors that were best at categorising a wine, because they were used predominately for either one group or another) were used to describe complex, rather than simple, wines. This implies that participants were somehow more consistent in using words to describe complex wines, or, on the other hand, simple wines can be described in different ways, but the more complex wines in this case all had traits in common.

Flavour descriptor	Likelihood	Ratio
Smoke	Complex : simple	3.8 : 1

Resin	Complex : simple	2.4 : 1
Vanilla	Complex : simple	2.4 : 1
Mineral	Simple : complex	2.2 : 1
Earth	Complex : simple	1.7 : 1
Cherry	Complex : simple	1.7 : 1
Meat	Complex : simple	1.7 : 1
Liquorice	Complex : simple	1.7 : 1
Sweet	Simple : complex	1.6 : 1
Cranberry	Simple : complex	1.6 : 1

Table 6. Top 10 most informative flavour descriptors from participants' tasting notes of the three red wines, that indicates whether a wine is 'more complex' or 'more simple', with its related likelihood ratios.

A similar analysis was conducted for wine liking (not liked = rating 7 or lower, liked = rating 8 or higher) and for quality (lower quality = rating 2 or lower, higher quality = rating 3 or higher), again restricted only to the flight of red wines since they are roughly comparable. For liking, the top 10 most informative words were all used to describe liked rather than disliked wines, which again implies that the participants (casual consumers in this case) might use more consistent language when it comes to describing wines they like. The reasoning goes that participants evidently shared vocabulary in common for wines they liked, which explains the high likelihood ratio of words such as "vanilla" and "earth". On the other hand, descriptors associated with disliked wines might have varied widely, so there were not any terms used consistently by participants. In contrast, when it comes to quality (of Bordeaux blend red wines, at least), there are many more informative words to describe lower, rather than higher, quality wines. Once again, participants might have had a more consistent idea of what constituted a low quality wine, as opposed to a high quality one.

Flavour	Likelihood	Ratio	Flavour	Likelihood	Ratio
Vanilla	Like : dislike	4.2 : 1	Leather	Higher quality : lower quality	3.9 : 1
Earth	Like : dislike	3.0 : 1	Cranberry	Lower quality : higher quality	2.7 : 1
Red	Like : dislike	2.4 : 1	Spice	Lower quality : higher quality	2.7 : 1
Fruit	Like : dislike	2.0 : 1	Tar	Lower quality : higher quality	2.7 : 1
Chocolate	Like : dislike	1.8 : 1	Bitter	Lower quality : higher quality	2.7 : 1
Pungent	Like : dislike	1.8 : 1	Pepper	Lower quality : higher quality	2.3 : 1
Raspberry	Like : dislike	1.8 : 1	Smoke	Higher quality : lower quality	2.2 : 1
Fig	Like : dislike	1.8 : 1	Blackberry	Higher quality : lower quality	1.8 : 1

Tobacco	Like : dislike	1.8 : 1	Raspberry	Lower quality : higher quality	1.6 : 1
Smoke	Like : dislike	1.8 : 1	Tannins	Lower quality : higher quality	1.6 : 1

Table 7. Left: top 10 most informative flavour descriptors indicating whether a wine is more liked or disliked, with its related likelihood ratios. Right: top 10 most informative flavour descriptors that indicates whether a wine is rated as higher or lower quality, with its related likelihood ratios.

Discussion

What does the present study reveal about complexity in wine as perceived by social drinkers? First, our results demonstrate that, at least for social drinkers, complexity is correlated with liking and quality, but not with the number of flavours detected (once controlling for liking and quality). The relationship between complexity and such positive qualities is in line with prior studies for both wine professionals and consumers (Parr et al., 2011). Interestingly, liking and quality had independent effects on complexity, even for our participant pool of mostly social drinkers³. This result therefore goes against previous research suggesting that novices tend to conflate liking and quality (Lawless et al., 2007). On the other hand, the fact that we found no relationship between complexity and the number of flavours perceived is similar to Schlich et al.'s (2015) findings, whereby perceptual separability of individual flavours of a wine was not a factor positively associated with perceived complexity. This supports the theory that there might be a more integrated, holistic assessment of wine complexity that goes beyond noticing elemental components of a wine.

In the selection of wines used for the present study, red wines were perceived as more complex than whites. Of course, without controlling for other variables (such as the presence of oak), it is difficult to say whether social drinkers perceive red wines to be more complex in general, or whether instead this was true only for the particular choice of white and red wines in the present study.

The wines used in the present study covered a wide spectrum of prices (see Table 5), but perceived complexity was not correlated with price (contrary to what some New

³ It should be noted that this particular participant pool all held doctorate degrees, therefore it might be attributed to having a more mindful/intellectual approach to wine tasting.

Zealand producers might be hoping for, see Parr, 2015). In fact, the only measured attribute that was correlated with price was quality, where the higher priced wines were, in fact, perceived to be of higher quality. This is different from Hopfer et al.'s (2015) findings suggesting that there were no significant correlations between retail bottle price and points awarded during wine competitions, but this might be explained by the fact that Hopfer et al. used 27 California Cabernet Sauvignon wines from similar vintages. In the present study, on the other hand, the quality-price correlation might be driven by the fact that the obviously-aged (2001 vintage) Chateau Cantemerle was the most complex and expensive wine of the group. This result therefore supports Singleton and Ough's (1962) early suggestion that flavour complexity is one of the primary effects produced by proper ageing.

To answer the question of the importance of the timeframe of flavour perception in judgments of complexity, we found that, for the red and sweet wines, complexity was correlated with the number of flavours perceived in the mouth before swallowing, but not after swallowing. In other words, the flavours perceived on the finish did not seem to impact participants' ratings. This could have to do with the fact that people are not used to paying attention to those flavours that may remain in the mouth post-swallowing (Neely & Borg, 1999; Noble, 1994). This can be contrasted with experienced wine tasters, who have been trained to evaluate the length and complexity of finish as a part of the quality of a wine (Charters & Pettigrew, 2007). In the future, time-based methodologies such as Temporal-Choose-All-That-Apply (TCATA) might be used to assess the exact time frame of participants' flavour perceptions. TCATA is especially appropriate as it measures simultaneous perceived flavours over time, therefore allowing researchers to answer both the question of *how many* flavours are perceived at once as well as *when* they are perceived. It would be especially interesting in future research to compare the temporal assessment of flavours for both novices and trained wine tasters, in order to test the hypothesis that trained wine tasters use the full temporal extent of the tasting experience in their judgment (e.g., quality/complexity/preference) of the wine.

Finally, the preliminary semantic analysis of participants' word choice reported here revealed that flavour descriptors associated with secondary and tertiary flavours (e.g., smoke, resin, vanilla, earth, leather) were perceived as more complex, were rated as

of higher quality, and were liked more. On the other hand, attributes that imply a lack of balance arising from a bitter or pungent element (e.g., cranberry, spice, tar, bitter, pepper) are more often used for those wines that are rated to be of lower quality (Saenz-Navajas et al., 2015, also found that wines containing astringent-related chemical compounds were rated to be of lower quality by consumers, for Rioja red wines).

Another interesting result to emerge from the lexical analysis was that social drinkers in this study seemed to use more consistent language when it came to describing wines with more complexity, compared to those of lower complexity. In other words, some descriptors, such as smoke, resin, and vanilla, were used more consistently for wines that seemed to be complex rather than those wines that appeared simple. Similarly, all of the top 10 most informative descriptors for like/dislike category were used more often to distinguish a wine that was liked than one that was disliked, which implies that the participants agreed more on what were seen as positive characteristics, and perhaps differed in their opinions on what negative traits for wines were. Finally, more descriptors were informative of lower rather than higher quality, implying that while it is easier to discern a low quality wine, it is harder to describe a good quality wine – or that high quality wines can vary in style and description (although leather seems to be prominent descriptor of high quality). This information is highly relevant for wine producers to understand which wine attributes appeal (or not) to the novice drinkers. For instance, the present results revealed an agreement on what traits appear in a more likable Bordeaux red wine, as well as which traits are associated with lower quality wine.

Granted, the semantic analysis is limited to the flight of three Bordeaux blend red wines used in the tasting. Given the relatively small sample of wines and tasting notes, more comprehensive studies should be conducted in order to draw any definitive conclusions. It should especially be noted that, for now, the results here are not generalizable beyond the specific three Bordeaux-style wines used in the tasting, without further research involving more styles and vintages of wines. In terms of acquiring bigger datasets, similar machine learning analyses can be performed on large corpuses of wine tasting notes (such as CellarTracker, with over 2 million wine reviews), especially in combination with known complexity ratings (such as reflected

from wine competition awards) to extract more robust findings. Nevertheless, these preliminary findings – via a combination of sensory science and machine learning techniques – do provide a glimpse into intriguing behavioural differences in how the casual consumer navigate the triad of wine complexity, liking, and quality.

FUNDING

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.

REFERENCES

- Bird, S., Klein, E., & Loper, E. (2014). *Natural language processing with Python*. Sebastopol, CA: O'Reilly Media.
- Bo, T., Kay, S., He, H. (2016). Toward optimal feature selection in Naïve Bayes for text categorization. *IEEE Transactions on Knowledge and Data Engineering*, **28**, 2508-2521.
- Charters, S., & Pettigrew, S. (2007). The dimensions of wine quality. *Food Quality and Preference*, **18**, 997-1007.
- 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.
- Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning*, **29**, 103-130.
- Hendrickson, J. B., Huang, P., & Toczko, A. G. (1987). Molecular complexity: a simplified formula adapted to individual atoms. *Journal of Chemical Information and Modeling*, **27**, 63-67.
- Hopfer, H., Nelson, J., Ebeler, S. E., & Heymann, H. (2015). Correlating wine quality indicators to chemical and sensory measurements. *Molecules*, **20**, 8453-8483.
- Hunt, A. (2015). "Complexity". *Oxford Companion to Wine Online*. Accessed December 7th, 2017. <https://www.jancisrobinson.com/ocw/detail/complexity>.
- Jinks, A., & Laing, D. G. (1999). Temporal processing reveals a mechanism for limiting the capacity of humans to analyze mixtures. *Cognitive Brain Research*, **8**, 311-325.
- Kibriya, A. M., Frank, Pfahringer, B., and Holmes, G. (2005). Multinomial naive Bayes for text categorization revisited. In *AI 2004: Advances in Artificial Intelligence*, (pp. 488-499). New York, NY: Springer.
- Koller, D., & Sahami, M (1997). Hierarchically classifying documents using very few

words. *Proceedings of the Fourteenth International Conference on Machine Learning (ML-97)*, pp. 170-178.

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

Laing, D. G., Link, C., Jinks, A., & Hutchinson, I. (2002). The limited capacity of humans to identify the components of taste mixtures and taste–odor mixtures. *Perception*, **31**, 617-635.

Lawless, H. T., Liu, Y.-F., & Goldwyn, C. (1997). Evaluation of wine quality using a small-panel hedonic scaling method. *Journal of Sensory Studies*, **12**, 317-332.

Marshall, K., Laing, D. G., Jinks, A. L., & Hutchinson, I. (2006). The capacity of humans to identify components in complex odor-taste mixtures. *Chemical Senses*, **31**, 539-545.

Neely, G., & Borg, G. (1999). The perceived intensity of caffeine aftertaste: Tasters versus nontasters. *Chemical Senses*, **24**, 19-21.

Noble, A. C. (1994). Bitterness in wine. *Physiology & Behavior*, **56**, 1251-1255.

Noble, A. C., Arnold, R. A., Masuda, B. M., Pecore, S. D., Schmidt, J. O., & Stern, P. M. (1984). Progress towards a standardized system of wine aroma terminology. *American Journal of Enology and Viticulture*, **35**, 107-109.

Parr, W. V. (2015). Unraveling 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.

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing actions can modulate neural representations of experienced pleasantness. *Proceedings of the National Academy of Sciences of the USA*, **105**, 1050-1054.

Robinson, J., & Harding, J. (Eds.). (2015). *The Oxford companion to wine* (4th Ed.). Oxford, UK: Oxford University Press.

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.

Sell, C. (2006). Structure-odor relationships: On the unpredictability of odor. *Angewandte Chemie International Edition*, **45**, 6254-6261.

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

Snitz, K., Arzi, A., Jacobson, M., Secundo, L., Weissler, K., & Yablonka, A. (2016). A cross modal performance-based measure of sensory stimuli intricacy. *PLoS ONE*, **11**, e0147449.

Wine and Spirits Education Trust (2017). *WSET Level 4 Diploma in Wines and Spirits Candidate Assessment Guide, Part 2: Tasting*. Issue 12, August 2017.