ARTICLE

The impact of wine tasters' expectations on wine quality ratings and willingness-to-pay

Philippe Masset and Steffen Raub

EHL Hospitality Business School, HES-SO, University of Applied Sciences and Arts Western Switzerland, Route de Cojonnex 18, 1000 Lausanne, Switzerland
Corresponding author: Philippe Masset; email: philippe.masset@ehl.ch

Abstract
Is it possible to exploit cognitive biases so that a non-professional taster prefers one wine to several other absolutely identical wines? To address this question, three complementary experiments were carried out. Each time, five wines were tasted blind in a tasting laboratory by 24 to 34 tasters. Converging evidence from the experiments shows that participants were not capable of identifying that some of the wines they were tasting were absolutely identical. Moreover, the results show that by providing information about the wines' ratings, prices, or reputation, tasters' expectations can be modified, and, as a result, their evaluations of the wines can be altered. Specifically, we show that it is possible to modify the ranking between different wines and to get tasters to prefer a wine over other absolutely identical wines. Finally, a surprising finding was that experienced tasters express stronger opinions and adapt their evaluations more strongly after being given manipulative information on the wines they taste.

> *A real connoisseur does not drink wine but tastes of its secrets.*
>
> —Salvador Dali

I. Introduction

Research on wine tasting, evaluation, and expertise has grown significantly over the past 20 years. This is due to the fact that wine offers a fascinating laboratory for studying research questions that extend beyond this noble beverage. Wine is indeed an experience good, with a strong cultural and aesthetic dimension. It is complex from both a chemical and sensory viewpoint and comes in a variety of forms

(grape variety/blends, style, aromatic profile, etc.). Research on the topic reflects the multifaceted nature of wine as it spans fields as diverse as food and beverage sciences, economics, marketing, enology, and psychology (see Spence (2020) for a recent review). Storchmann (2012) even identifies research on wine evaluation and expertise as among the three most relevant in wine economics.

Spence (2020, p. 2) notes that "a wide variety of cognitive and perceptual factors […] influence the wine-drinking experience." This combination between a psychological (cognitive) and a mostly experience-based (sensory) dimension implies that wine evaluation is a difficult exercise that requires expertise that substantially differs from other realms (Ashton, 2017). Extant research suggests that wine tasters may be notoriously unreliable at evaluating the quality of wines served blind and even at identifying perfectly identical wines as such when served blind. This holds both for professional wine judges (Hodgson, 2008) as well as for non-professional wine enthusiasts (Weil, 2007). For instance, in a study of professional wine judges, Hodgson (2008) found that in only 30 out of 65 tasting panels, the wine was the only significant factor in determining the judges' quality ratings, whereas in the remaining 35, the judges' bias played an important role. In an experimental investigation with non-professional wine enthusiasts as participants, Weil (2007) served three glasses to the participants, of which two were identical. He then asked participants to identify the wine that differs from the doubleton. Only in 51% of all cases did the tasters correctly identify the pair of completely identical wines. While this is significantly higher than the 33.3% that would be expected in the case of a purely random choice process, it still leaves about half of the comparisons where the matching process was unsuccessful.

Despite the ever-increasing importance of wine experts in the wine market (Masset, Weisskopf, and Cossutta, 2015; Cardebat and Livat, 2016), recent studies suggest that these individuals do not necessarily have a better sensory/perceptual capacity than wine novices. They distinguish themselves primarily by their ability to express their feelings using their conceptual/semantic wine knowledge (Spence and Wang, 2019). However, their verbal descriptions of a wine's sensory qualities may not necessarily be useful for non-experts. In the study by Weil (2007), out of the participants who correctly identified the doubleton of identical wines in the series of three wines served, only 49.4% could successfully match this wine with the correct option in two alternative tasting notes by wine experts, a result that is not significantly different from what could be expected in a random process.

Rodrigues and Parr (2019) illustrate the role of the taster's cultural background in wine appreciation. Their article shows that the quality of a wine is not absolute; it depends on a frame of reference and the way the taster reacts to a variety of stimuli. For example, Niimi et al. (2017) show that describing the wine positively affects the emotions induced by its tasting (see also Niimi, Danner, and Bastian (2019) for a brief review of the link between wine and emotions). Parr (2019, p. 230) notes that "wine is as cerebral as it is sensual." As a result, it is not surprising to see that tasters, whether novices or experts, are sensitive to a set of biases induced by the color of the wine, the weight of the bottle, or its price.

Price as a signal of product quality has been extensively studied (e.g., Bagwell and Riordan, 1991). In the specific context of the wine market, Schnabel and

Storchmann (2010) found support for a positive correlation between price signals and wine quality, in particular for higher-priced wines, in a large sample of German wines. In a similar vein, through an experimental design, Mastrobuoni, Peracchi, and Tetenov (2014) found a signaling effect of price that was positive when going from a low to a medium price, but not significant when proceeding from a medium to a high price. Goldstein (2019) shows that providing information about price biases tasters' evaluations. He identifies an asymmetric effect, with negative information (i.e., presenting the wine as cheap, the so-called nocebo effect) affecting the expectations and therefore the evaluation of tasters more strongly than positive information (the so-called placebo effect). Goldstein's article complements the research of Plasmann et al. (2008), who show that price affects tasters' ratings but also their pleasure (measured via activity in the cortex). Actually, even when intrinsic cues are available (through sensory experience), extrinsic cues (e.g., color, prices, or expert ratings) seem to remain the key drivers of quality evaluation (Veale and Quester, 2008). Aqueveque (2018) shows, however, that experts use extrinsic cues less extensively than novices to infer quality. He further demonstrates the presence of a Dunning-Kruger (1999) effect, with less (more) competent tasters tending to overestimate (underestimate) their knowledge.

Our study examines whether the expectations of tasters can be influenced to make them like certain wines more than others. We proceed in two steps. First, we test participants' ability to notice that several wines served blind to them are identical. Serving identical wines is essential for ensuring that expressed preferences cannot be justified by actual differences in quality or style between wines. Next, we examine whether it is possible to direct tasters' preferences (expressed via quality ratings and willingness-to-pay) toward particular wines by altering their expectations. To do so, we influence tasters' expectations via extrinsic cues related to price, expert rating, and prestige of the wine (Ashton, 2017).

Three experiments were conducted. In the first experiment, we served three wines to the participants, and we then shared information about the ratings of those wines with them. We then served two additional wines without giving any further information to the participants. These two wines were identical to two of the three wines tasted before. The results are surprisingly clear-cut. No one identified that some of the wines were identical. This result is consistent with the study by Goldstein (2019), in which none of the participants identified that the two wines they were served were actually the same. More strikingly, participants also expressed strong preferences for the last two wines.

The second experiment began with a first round in which all participants were served an identical wine A. In the second round, the participants were split into two groups. The first group was served wine B and the second group wine C. The wine was then presented to everyone as a single wine B and the rating for wine B was shared. In the next round, the two wines were inversely served to the two groups of participants. The wine was then presented as a single wine C and the rating for wine C was shared. Objectively, wine C was more expensive and clearly better scored than wine B. In the following two rounds, the same logic was repeated, but this time, the information shared about the two wines concerned their price rather than their rating. Results suggest that one can influence tasters' evaluations. It must, however,

be noted that giving the wrong information is enough to modify expectations and alter the ranking of the wines, but it is not enough to completely erase the differences between two wines.

The third experiment reveals that one can orient the preferences of the tasters toward a particular wine. In this experiment, an identical wine was served in three successive rounds. It was first served blind from a decanter. Then it was served openly from a bottle of moderately priced Ghemme. Finally, it was again served openly from a bottle of highly-priced Barolo. Results show that presenting the same wine in a more prestigious bottle leads to a better evaluation and higher willingness-to-pay. Our results further illustrate that those who self-evaluate as "good tasters" adapt their ratings and prices more strongly to the information given to them. This can be explained in two ways: (i) they express more assertive opinions because they feel more experienced, and (ii) they may also feel the need to demonstrate that they actually know wine by aligning their expressed opinion with what they think should be correct (independently of their own sensory perception). This suggests that one of the differences between novices and more skilled consumers is that the latter express stronger opinions.

II. Research design

This section first clarifies our research agenda and the questions we examine through the experiments. The context in which the experiments took place is then presented. Finally, the experiments are discussed in detail, and their links to the research questions are established.

A. Research agenda

Tasting and rating a wine is a technical and complicated exercise. The brain plays an important role, and various cognitive biases have been demonstrated to affect even professional tasters. Expectations play a key role. The information available to tasters contributes to their expectations and consequently has an impact on their evaluations. In this study, we address the following three research questions:

  I. What is the effect of wine tasters' expectations on their evaluation (quality rating and willingness-to-pay) of a wine?
 II. Is it possible to reduce differences in the evaluations of objectively different wines by modifying wine tasters' expectations?
III. Is it possible to induce differences in the evaluation of objectively identical wines by modifying wine tasters' expectations?

The first question is directly inspired by Siegrist and Cousin (2009). They demonstrate that information on wine ratings influences the expectations of tasters and, thereby, their ratings and willingness-to-pay. The purpose of this question is, therefore, to examine whether we obtain results consistent with those documented in the literature, before tackling more original questions.

The second question involves the modification of tasters' expectations via information shared about the wines. Some participants receive incorrect information about

the rating or the price of the wines they taste. The objective is to see if this incorrect information can be sufficient to smooth out the differences in evaluations between different wines. Specifically, the goal is to assess if modified expectations can cause a clearly better wine to end up being rated similarly (or worse) to a lesser wine. This question extends the analysis of Plassmann et al. (2008), who show that tasters prefer a wine that is supposed to be more expensive as compared to another if both wines are actually identical.

The third question is the logical conclusion of the two previous ones. It aims at examining if tasters' expectations can be modified in such a way as to predetermine the wine they evaluate best. This is a completely new question, but the experiment (described later) incorporates elements already studied, notably in Wang and Spence (2019), who examine the role of wine color by coloring white wine into rosé. All participants, from beginners to experts, had more difficulty assessing the fake rosé. But interestingly, experts appeared to be more influenced by the color. The authors argue that this could be explained by the fact that experts consider the color to be a piece of information about the wine. In an earlier study, Parr, White, and Heatherbell (2003) showed that experts are better at identifying "masked" red wines than social drinkers, but they nevertheless remain affected by the coloring.

For the three questions and in the three experiments, we also test for differences in behavior between tasters who evaluate themselves as good or very good tasters and others who consider themselves rather novices.

B. Context

The three experiments took place during an elective course in "Wine Economics" taught at a European hospitality business school. All the participants had previously taken courses on wine knowledge, enology, and wine service. This means that the profile of the participants is rather homogeneous. However, some of the participants belong to wine committees and participate more frequently in tastings. This makes it possible to examine whether wine knowledge affects tasters' reactions to a change in their expectations. A questionnaire used during the first session of the course (week 1) is used to collect this information.

The Wine Economics course runs for five weeks (two three-hour sessions per week). It typically includes two tastings in addition to the experiment. These tastings are conducted with guest speakers and focus on specific wine regions. During these tastings, participants are invited to share their appreciation of the wines through a standardized questionnaire. The same questionnaire is used in the experiments. The experiments always take place during the last session of the course (week five). This means that the students are already familiar with the questionnaire and its contents. This reduces the risk of misinterpretation.

In order to prevent students from engaging in strategic behaviors (e.g., trying to identify if there is a trick), the experiments did not take place every semester. The first one took place in 2018, the second in 2021, and the third in 2022. Moreover, in order not to give students the impression that the tasting of the semester may include tricks, it is organized in the exact same way as the other tastings that take place earlier during the course: students are asked to taste silently and respond

individually, but they are allowed to leave the class if necessary or move between tasting rounds if they wish. As explained next, this may result in attrition, but this is the price to pay if one wants to make sure that participants answer in an unbiased way.

Finally, the experiments take place in a tasting laboratory with individual booths, creating excellent overall conditions for tasting and data collection.

C. Experiments

Table 1 presents an overview of the three experiments. They share several characteristics. In all three cases, five wines were tasted. Moreover, in each experiment, several of the wines tasted were identical.

To address the first research question, the first experiment consists of tasting five wines, but only three are actually different (two wines are served twice). All wines were tasted blind. For the first three wines (round 1: wine A, round 2: wine B, and round 3: wine C), no information was given to the participants. The Wine Advocate's ratings of the first three wines were provided to participants after they had been tasted. This information has, of course, an effect on tasters' expectations. The wines presented in rounds 4 and 5 (wine B and wine A) were the same as those presented in rounds 2 and 1, respectively. The purpose of this experiment is to examine the role of expectations on participants' ratings and willingness-to-pay.

To address the second research question, five wines were served in the second experiment. The class is divided into two groups without this information being given to the participants. In the first round, all participants receive the same wine (A). In the second round, the first group receives wine B, while the second group receives wine C. The rating of wine B is also communicated to all participants. Thus, half of the participants have correct information (those who were served wine B), and the other half have incorrect information (those who were served wine C). In the third round, wine C is served to the first group, while the other group receives wine B. The rating of wine C is communicated. Rounds 4 and 5 are similar to rounds 2 and 3, and the same wines are served again in the same order. However, the information given is about the price of the wines rather than their

Table 1. Design of the experiments

Panel A: Experiment 1						
	Round 1	Round 2	Round 3	After Round 3	Round 4	Round 5
Wine served	A	B	C		B	A
Information given	—	—	—	TWA ratings of A, B, and C	—	—
Deception involved	No	No	No		No	No

Wine A: Domaine Courbis, Cornas, Les Eygats 2011 (TWA rating: 93)
Wine B: Domaine de la Charbonnière, Châteauneuf-du-Pape, Les Hautes Brusquières 2010 (TWA rating: 97)
Wine C: Domaine de l'Horizon, Roussillon, Rouge 2008 (TWA rating: 93)

Panel B: Experiment 2					
	Round 1	Round 2	Round 3	Round 4	Round 5
Wine served	A	B (group 1) C (group 2)	C (group 1) B (group 2)	B (group 1) C (group 2)	C (group 1) B (group 2)
Information given	—	Rating of B	Rating of C	Price of B	Price of C
Deception involved	No	Yes (*)	Yes (*)	Yes (*)	Yes (*)

(*): group 2 only
Wine A: Clos du Caillou, Côtes du Rhône, La Réserve 2015 (TWA rating: 91, Price: CHF 35)
Wine B: Domaine de Ferrand, Châteauneuf-du-Pape 2015 (TWA: 90, CHF 50)
Wine C: Domaine de la Janasse, Châteauneuf-du-Pape, Chaupin 2015 (TWA: 96, CHF 70)

Panel C: Experiment 3						
	Round 1	Round 2	Round 3	After Round 3	Round 4	Round 5
Wine served	A	B	C		C1	C2
Information given	—	—	—	Label of A, B, and C	Bottle and label of C1	Bottle and label of C2
Deception involved	Yes (*)	Yes (*)	No		Yes (**)	Yes (**)

(*): white wine colored in red
(**): the bottle presented does not correspond to the wine that is effectively served
Wine A: Valentina Andrei, Valais, Roussane Marsanne 2021
Wine B: Sybille Kuntz, Mosel, Riesling trocken 2021
Wine C: Mas Jullien, Terrasses du Larzac, Autour de Jonquieres 2016
Wine C1: Mas Jullien, Terrasses du Larzac, Autour de Jonquieres 2016 (served in a bottle of Cantalupo, Ghemme, Collis Breclemae 2011)
Wine C2: Mas Jullien, Terrasses du Larzac, Autour de Jonquieres 2016 (served in a bottle of G. Mascarello, Barolo, Monprivato 2013)

rating. It should be noted that wine C has a clearly better score and is more expensive than wine B. The purpose of this experiment is to examine the extent to which building similar expectations may help reduce gaps in evaluation between two different wines.

The third experiment follows a different framework as it involves two different white wines and one single red wine. The participants first taste two white wines colored red (rounds 1 and 2). Then the three identical red wines are served one after another, blind from a decanter (round 3), from a bottle of a 40-euro Ghemme (an appellation in Northern Piedmont) (round 4), and from a bottle of a 150-euro Barolo (round 5). The three bottles from which the actual wine came are strictly

identical and were opened at the same time. In order to reduce the risk that participants may suspect the last three wines to be identical, the names of the first three wines are revealed once they have been tasted (i.e., right after round 3), and the participants are told that the last two wines are meant to conclude the semester on a festive note and that they would therefore not be drunk blind.

III. Empirical analysis

A. Effect of expectations on the evaluation (Experiment 1)

The results of Experiment 1 are reported in Table 2. Panel A1 shows the ratings of the various wines. More than 80% of the respondents preferred wines 4 and 5 to wines 2 and 1 (which were actually the same). The average scores increase from 70.7 and 76.1 to 83.8 (+13.1) and 86.7 (+10.6), respectively. These differences are significant at the 99% level. Moreover, the median difference in scores is very close to the mean. All this suggests that despite the small sample, the differences are both statistically and economically significant. The prices (Panel B1) are totally consistent with the ratings, with one nuance. Giving information on the ratings of wines 1 to 3 reduces the standard deviation of the ratings but increases that of the prices. This can probably be explained by two elements: the information given to the participants pertains to ratings of the wines, and the upward revisions of respondents' expectations are mechanically bounded for ratings (maximum 100 points) but not for prices.

One potential issue lies in the fact that respondents probably do not all use the same scale to evaluate the ratings and prices of the various wines. Some are more generous, others less. Some maintain small differences between the best and worst wines, while others use a wider range. Finally, some may be extreme when a wine matches their preferences perfectly or not at all. In short, the distribution of ratings and prices certainly deviates from normal and is potentially very different from one respondent to another. This problem is reinforced by the small sample size. One solution is to use the five-wine ranking. Thus, for each respondent, the original rating (and price) is re-expressed as a ranking from 1 (= highest) to 5 (= lowest). The results reported in Panels A2 and B2 are based on the rankings. We can see that the results remain very stable, which again suggests that they are robust.

In order to deepen the analysis, we further examined whether variables associated with individual participants could help explain their reactions. We consider the coherence of participants (i.e., whether their variation in score is coherent with the variation in price), gender, and their wine knowledge self-assessment. As the sample is small, the results of multivariate regressions lack statistical significance. However, a simple t-test shows that "good tasters" tend to adapt their ratings and prices more strongly after information about the quality and price of wines 1 to 3 has been shared with them. The results are statistically significant for the pair wine 2 vs. wine 4 (differences in ratings: t-stat of −2.17 and p-value of 0.02; differences in prices: t-stat of −1.99 and p-value of 0.04). For the pair wine 1 vs. wine 5, the difference is too small to be significant. This difference between the two pairs may be due to the fact that the information was given just before wine 4, and therefore some participants potentially adapted their expectations more strongly for wine 4 than for wine 5.

Table 2. Results of experiment 1

	Wine 1	Wine 2	Wine 3	Wine 4	Wine 5	Delta 4-2	Delta 5-1
Panel A1: Original scores							
Average	70.71	76.10	62.19	86.74	83.81	10.65***	13.1***
Median	75.00	80.00	75.00	90.00	89.00	10.00	15.00
Std. deviation	18.09	13.64	27.20	10.74	13.19	11.57	12.72
>0						84%	81%
Panel B1: Original log(prices)							
Average	3.31	3.67	3.39	4.09	3.99	0.42***	0.68***
Median	3.40	3.81	3.56	4.14	4.01	0.37	0.56
Std. deviation	0.44	0.42	0.61	0.51	0.51	0.46	0.50
>0						84%	97%
Panel A2: Rankings scores							
Average	3.82	3.27	4.18	1.55	2.18	−1.73***	−1.65***
Median	4.00	3.00	5.00	1.00	2.00	−2.00	−2.00
Std. deviation	0.92	1.02	1.17	0.76	1.00	1.41	1.53
<0						84%	81%
Panel B2: Rankings log(prices)							
Average	4.45	3.06	3.94	1.55	2.00	−1.52***	−2.45***
Median	4.50	3.00	4.00	1.00	2.00	−1.50	−3.00
Std. deviation	0.61	1.05	1.05	0.72	0.85	1.41	1.06
<0						84%	97%

Note: *, **, and *** denote significance at the 90%, 95%, and 99% level, respectively.

B. Modifying expectations to reduce differences in evaluation among different wines (Experiment 2)

Figure 1 shows how participants compare wines B (rated 90 points by TWA and sold for about CHF 50) and C (96 points, CHF 70) according to two factors: (1) whether they are in the group that received correct information or not; (2) whether the information concerns the rating of the wine or its price.

The left panel of Figure 1 reports the results when information about the ratings of the wines is provided to the participants. Half of them (group 1) taste wine B in round 2 and wine C in round 3 and get the correct information. The other half (group 2) tastes the wines in reverse order and thus gets the wrong information. The wine served in round 3 obtains a much better ranking both in terms of quality and willingness-to-pay. The result is similar for both groups, which is quite remarkable since the two wines are served in reverse order for group 2 as compared to group 1. When group 1 receives the highest-rated and most expensive wine, group 2 receives the lowest-rated and least expensive wine, but the information given is the same for everyone: "here is a wine with a TWA rating of 96 points." It must be noted that group 2 reacts less strongly to the information given and that the difference in ranking between the wines served in rounds 2 and 3 is less pronounced for this group, but this group nevertheless shows a preference for the latter wine even though this wine is actually less good (or rather less well rated) and less expensive than the previous one.

The right panel of Figure 1 reports the results when the information provided pertains to the price of the wines tasted. Results are less marked as compared to the ones discussed previously. The group that receives the correct information continues to rate more favorably the wine that is actually the most expensive. The group that receives false information gives a lower rating to the supposedly more expensive
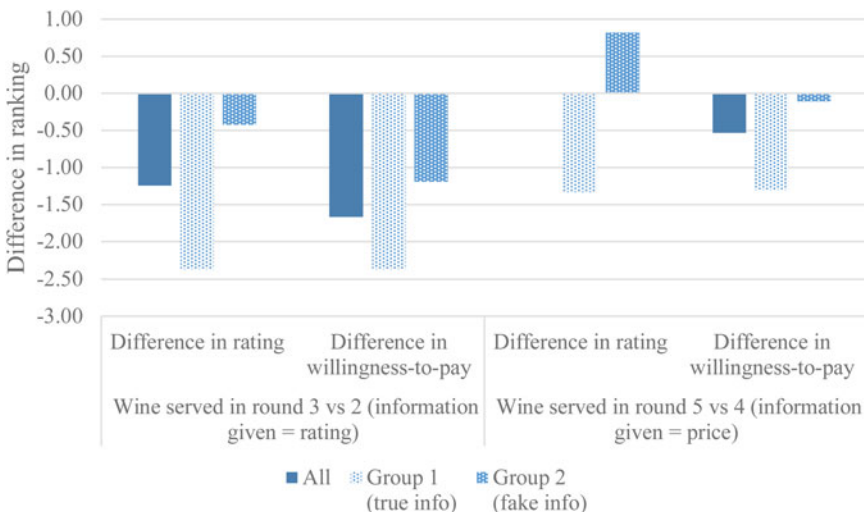

Figure 1. Summary results of experiment 2.

Table 3. Regression analysis of experiment 2

	Wine C vs. Wine B: Difference in ratings		Wine C vs. Wine B: Difference in willingness-to-pay	
	(I)	(II)	(III)	(IV)
Panel A: Regression results				
Intercept	−0.82 (0.66)	−1.24 (0.79)	−1.25*** (0.44)	−1.54*** (0.53)
Wrong information (WI)	1.31** (0.64)	1.87** (0.86)	2.43*** (0.44)	2.8*** (0.58)
Information on price (IP)	−0.39 (0.61)	−0.35 (0.61)	−0.31 (0.42)	−0.25 (0.42)
Self-assessment (SA)	−1.64** (0.64)	−0.94 (0.97)	−1.1** (0.44)	−0.6 (0.67)
Interaction SA × WI		−1.27 (1.3)		−0.86 (0.89)
Nobs.	35	35	34	34
R2	0.33	0.35	0.61	0.62
Panel B: Implications (based on regression coefficients)				
Reaction (difference between wine C and wine B) of all tasters:				
True information on rating	−0.82	−1.24	−1.25	−1.54
True information on price	−1.21	−1.59	−1.56	−1.79
Wrong information on rating	0.49	0.63	1.18	1.26
Wrong information on price	0.10	0.28	0.87	1.01
Reaction (difference between wine C and wine B) of "good" tasters:				
True information on rating	−2.47	−2.18	−2.35	−2.14
True information on price	−2.86	−2.53	−2.66	−2.39
Wrong information on rating	−1.15	−1.58	0.08	−0.21
Wrong information on price	−1.55	−1.92	−0.23	−0.46

Notes: Panel B shows the implications of the results (coefficients) from Panel A in terms of reaction (i.e., change in rating or price) of the tasters following the release of information about the rating or the price of a wine. Standard errors are in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% level, respectively.

wine (which is actually less expensive) but is still, surprisingly, willing to pay a little more for it. This weaker reaction and the gap between rating and willingness-to-pay for group 2 can certainly be explained by the fact that the prices announced for both wines are in a rather similar range (CHF 50 and CHF 70), whereas their ratings are more dissimilar (90 points, i.e., a very good wine, versus 96 points, i.e., a truly extraordinary wine).

Table 3 reports the results of a regression, the purpose of which is to better understand and assess the statistical significance of the participants reactions to the information shared with them. The difference in ratings and willingness-to-pay between wines B and C (dependent variable) is regressed on a set of dummy variables to control if the information shared with the participants is wrong ("WI" in the table), if it pertains to the price of the wine (IP), and if the participant evaluates himself/herself

as a good taster (SA). In order to control for the fact that a good taster may react differently as compared to a taster with less expertise, we also consider a specification in which the wrong information dummy and the self-assessment variables interact together (SA × WI). In order to ease the interpretation of the results, the table is structured in two panels: Panel A reports the results from the regression, and Panel B analyzes their implications.

The intercept shows that participants give generally higher ratings and are ready to pay more for wine C (which is the supposedly better and more expensive wine) than wine B. Respondents who received false information, however, reacted in a markedly different manner. They tend to prefer wine B to wine C. The corresponding dummy coefficient is statistically significant in all four regressions. Panel B further shows that the difference in ratings and willingness-to-pay between the two wines remains smaller as compared to the participants who received correct information. This means that altering tasters' expectations by giving them false information is sufficient to reverse their ranking (as compared to the group who received true information), but not to completely smooth out the differences between the two wines. Sharing information about the price of a wine rather than its rating leads to a slightly stronger difference between wines B and C, but the corresponding coefficient is not significant.

The coefficient associated with the taster's self-assessment of their expertise shows that respondents who consider themselves to be good connoisseurs seem to have a stronger preference for wine C as compared to wine B than the other respondents. If we take into account a possible interaction between the self-assessment and the "wrong information" variables, we find that good connoisseurs rated wine C higher than wine B independently of whether the information they received was correct or wrong. It thus seems that they are less sensitive to the release of (wrong) information. This might be due to the fact that they are more confident in their ability to accurately assess a wine. When it comes to their willingness-to-pay, the picture is a bit different: good connoisseurs who received wrong information are still willing to pay more for wine C, but the difference becomes much less substantial. All in all, this discussion suggests that more knowledgeable tasters do not hesitate to express stronger opinions. This observation is consistent with Experiment 1.

C. Modifying expectations to induce differences in evaluation among similar wines (Experiment 3)

Figure 2 shows the general results of Experiment 3. We can see that wines C1 (served in round 4) and C2 (round 5) are rated much higher than wine C (round 3), which is identical. The "collective reputation" effect (here Piedmont versus Languedoc) on tasters' expectations seems to be important. We can also observe a difference between wines C1 and C2, but it remains quite small. Both wines come from the same region, but the latter benefits from a higher individual reputation. Overall, presenting the same wine in a more prestigious bottle leads to a better evaluation and a higher willingness-to-pay. It is interesting to contrast these observations with the fact that, in practice, wine consumers do not necessarily display a preference for more expensive wines. Indeed, Goldstein et al. (2008, p. 7) show that "unless they are experts, [individuals] enjoy more expensive wines slightly less."
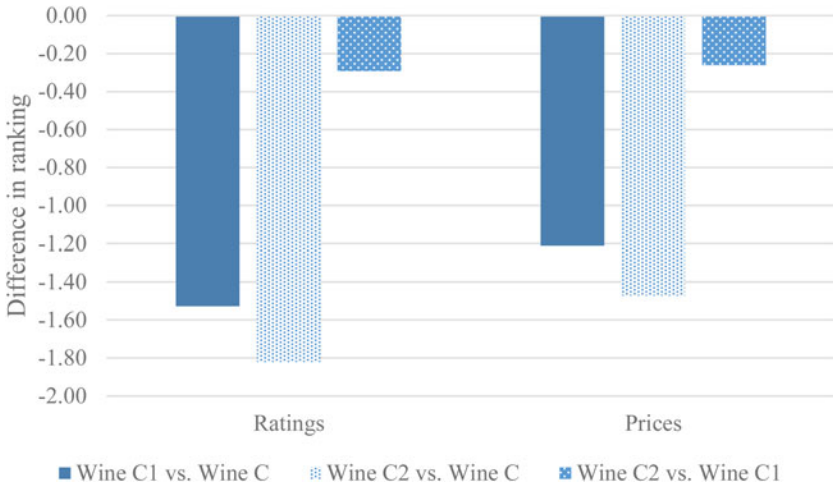
Figure 2. Summary results of experiment 3.

Table 4 shows the results of a more detailed analysis of the results of Experiment 3 with a regression. First, it should be noted that the R-squared is low. This is due to the fact that few variables are significant. The results again support the claim that respondents prefer wines C1 and C2 as compared to wine C. The difference in evaluation among the wines is slightly more pronounced (though the difference is not significant) when looking at ratings instead of willingness-to-pay. The coefficient associated with the self-assessment of expertise variable is clearly negative, but significant only in one specification. This suggests that participants who consider themselves good

Table 4. Regression analysis of experiment 3

	Wine C1 vs. C	Wine C2 vs. C	Wines C1 and C2 vs. C	
Intercept	−1.40*** (0.31)	−1.62*** (0.47)	−1.36***	−1.41***
Delta price	0.33 (0.43)	0.37 (0.65)	0.35 (0.39)	0.35 (0.39)
Self-assessment (SA)	−0.76 (0.57)	−1.31 (0.88)	−1.04* (0.52)	−0.76 (0.74)
Wine C2 vs C1 (dC)			−0.29 (0.39)	−0.2 (0.43)
Interaction SA x dC				−0.55 (1.05)
Nobs.	23	25	48	48
R2	0.10	0.10	0.11	0.11

Notes: Standard errors are in parentheses. *, **, and *** denote significance at the 90%, 95%, and 99% level, respectively.

tasters adjust their ratings and willingness-to-pay more strongly. One may further note that participants have a slightly more favorable rating for wine C2 relative to wine C1. The last specification suggests, however, that this difference in ratings is larger for good tasters. In other words, people who rate their knowledge as good find a more substantial difference between wines C1 and C2 than novices. This result is not statistically significant, so we analyze it with caution. Our interpretation is that connoisseurs feel more compelled to revise their evaluation upward when a visibly more expensive and, in principle, better wine is served to them. This observation requires further analysis, but it suggests that the experimental design works well and that it works at least as well on more experienced tasters as on less experienced tasters.

D. Some considerations regarding the use of deception in our experiments

The use of deception in wine experiments is a recurrent topic of discussion. Indeed, these experiments often have elements related to psychology and economics. However, in economics, deception, and in particular the fact of misleading the participants on the payoffs of an experiment, is generally considered unacceptable. The main problem is that economists want participants' choices to be influenced only by the expected payoffs and not by strategic behavior induced by the possibility that an experiment may contain erroneous elements. In psychology, on the other hand, deception is frequently used—it is even often at the heart of experiments.

Our three experiments *all* contain some elements of deception. The first experiment does not rely on any misinformation. But the simple fact of presenting the same wine several times in a blind tasting can be considered implicit deception, or at least ambiguity. Cooper (2014, p. 111) notes that "Deception is generally considered a sin of commission rather than omission, so other experimental techniques that could arguably be classified as deception are considered acceptable. Examples include the use of deliberate ambiguity […]." We can see that the border between deception and non-deception, and between unacceptable and acceptable, is tenuous. This type of ambiguity is found in most wine experiments (see, e.g., Goldstein (2019)).

Experiment 1 contains ambiguity. Experiments 2 and 3 incorporate explicit deception with misinformation about the wines tasted. At this point, it is useful to note one of the unique properties of wine experiments: it is actually possible to control the whole process, to have wines tasted that are rigorously similar, and thus to ensure that differences in perception between two tastings are the result of the context and/or the information communicated and nothing else. In other words, the use of deception via the tasting of several identical wines allows for an ideal experimental framework in which any measurable effect is solely induced by the information given to the participants. These considerations lead to several questions: Would it have been possible to perform experiments 2 and 3 using only ambiguity and without explicit deception? Is it acceptable to have taken the step from ambiguity to explicit deception in the present context? What are the possible consequences of using deception, and how can they be minimized?

Cooper (2014, p. 113) elaborates on deception and discusses the extent to which it can be considered acceptable in wine economics. He lists four conditions that can

justify the use of deception: (1) the deception does not harm subjects beyond what is typical for an economic experiment without deception; (2) the study would be prohibitively difficult to conduct without deception; (3) subjects are adequately debriefed after the fact about the presence of deception; and (4) the value of the study is sufficiently high to merit the potential costs associated with the use of deception.

Condition 1 is met. Likewise, we debriefed the participants, so condition 3 is satisfied, as in Lewis and Zalan (2014). As for condition 2, it would have been difficult to carry out experiments 2 and 3 without deception. In experiment 2, price ranges could have been reported rather than a single value, as suggested by Cooper (2014). A price range between CHF 50 and CHF 90 could have been given, so the average value would have corresponded to the price as reported (CHF 70). As for the ratings, other wines would have had to be chosen because it is not possible to communicate an appropriate range with wines scored 90 and 96 (the maximum rating is capped at 100 points). It should be noted, however, that giving information about a range is likely to induce other considerations at the participant level and result in less accurate data. The size of the sample would certainly have to be increased considerably, which poses significant organizational and logistical problems. More generally, such an approach contains a high level of ambiguity, which shows, again, that the boundary between the absence and presence of deception is often tenuous in wine economics. As for experiment 3, it would have been almost impossible to carry out and maintain such precise control over the relationship between the information given to participants and their reactions. Overall, the second condition can therefore also be considered satisfied.

The last condition is more subjective since it deals with the value of the research and the possible consequences of using deception. The risk is that "once subjects doubt the veracity of the experimental instructions and materials, they can begin to form their own theories about how they will be paid" (Cooper, 2014, p. 112). In the context of wine tasting, this risk seems limited. Indeed, the exercise of blind tasting is usual and almost always contains an element of ambiguity. The risk can be further reduced if the participants are properly briefed before the tasting and debriefed afterward. Moreover, it is easy to call on other participants at each tasting without them exchanging experiences with participants from previous tastings. In the present case, the experiments took place in the context of a course that involves blind tastings and they were spread over several years. Therefore, there is little or no risk that participants will give biased answers in response to doubts about the information provided about the experiment. As Cooper (2014, p. 114) states about Lewis and Zalan (2014) and Mastrobuoni, Peracchi, and Tetenov (2014) (whose papers share several common elements with our study), "the value of these studies is ultimately in the eye of the beholder."

IV. Conclusions

Our study contributes to the existing literature along several lines. We show that tasters' expectations can be modified and that this often has a significant impact on their rating and willingness-to-pay. We further show that by modifying their expectations, one can alter the way participants rate the wines they taste. This leads to changes in

ranking, and these changes seem to be more important for tasters with a good level of wine knowledge than for novices. In general, those whose self-evaluation identified them as "good tasters" tend to adapt their ratings and prices more strongly to the information given. That is, one of the differences between novices and more skilled consumers is that the latter express stronger opinions. This can be explained in two ways: (i) they express more assertive opinions because they feel more experienced, and (ii) they may also feel the need to demonstrate that they actually know wine by aligning their opinion with what they think is correct.

Overall, our results suggest that it is possible and relatively easy to nudge wine consumers into preferring a predetermined wine. It would be good to deepen the analyses presented in this paper. In particular, repeating similar experiments with different and larger samples may make the results more robust. In addition, it might be interesting to examine in more detail whether certain demographic and cultural variables affect the way tasters revise their expectations and evaluation.

References

Aqueveque, C. (2018). Ignorant experts and erudite novices: Exploring the Dunning-Kruger effect in wine consumers. *Food Quality and Preference*, 65, 181–184.

Ashton, R. H. (2017). Dimensions of expertise in wine evaluation. *Journal of Wine Economics*, 12(1), 59–83.

Bagwell, K., and Riordan, M. H. (1991). High and declining prices signal product quality. *American Economic Review*, 81(1), 224–239.

Cardebat, J.-M., and Livat, F. (2016). Wine experts' rating: A matter of taste? *International Journal of Wine Business Research*, 28(1), 43–58.

Cooper, D. J. (2014). A note on deception in economic experiments. *Journal of Wine Economics*, 9(2), 111–114.

Goldstein, R. (2019). Half-blind tasting: A deception-free method for sizing placebo and nocebo responses to price and packaging attributes. *Journal of Wine Economics*, 14(3), 321–331.

Goldstein, R., Almenberg, J., Dreber, A., Emerson, J., Herschkowitsch, A., and Katz, J. (2008). Do more expensive wines taste better? Evidence from a large sample of blind tastings. *Journal of Wine Economics*, 3(1), 1–9.

Hodgson, R. T. (2008). An examination of judge reliability at a major U.S. wine competition. *Journal of Wine Economics*, 3(2), 105–113.

Kruger, J. M., and Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134.

Lewis, G., and Zalan, T. (2014). Strategic implications of the relationship between price and willingness to pay: Evidence from a wine-tasting experiment. *Journal of Wine Economics*, 9(2), 115–134.

Masset, P., Weisskopf, J.-P., and Cossutta, M. (2015). Wine tasters, ratings, and en primeur prices. *Journal of Wine Economics*, 10(1), 75–107.

Mastrobuoni, G., Peracchi, F., and Tetenov, A. (2014). Price as a signal of product quality: Some experimental evidence. *Journal of Wine Economics*, 9(2), 135–152.

Niimi, J., Danner, L., and Bastian, S. E. (2019). Wine leads us by our heart not our head: Emotions and the wine consumer. *Current Opinion in Food Science*, 27, 23–28.

Niimi, J., Danner, L., Luxing, L., Bossan, H., and Bastian, S. E. (2017). Wine consumers' subjective responses to wine mouthfeel and understanding of wine body. *Food Research International*, 99(1), 115–122.

Parr, W. V. (2019). Demystifying wine tasting: Cognitive psychology's contribution. *Food Research International*, 124, 230–233.

Parr, W. V., White, K. G., and Heatherbell, D. A. (2003). The nose knows: Influence of colour on perception of wine aroma. *Journal of Wine Research*, 14(2–3), 79–101.

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

Rodrigues, H., and Parr, W. (2019). Contribution of cross-cultural studies to understanding wine appreciation: A review. *Food Research International*, 115, 251–258.

Schnabel, H., and Storchmann, K. (2010). Prices as quality signals: Evidence from the wine market. *Journal of Agricultural & Food Industrial Organization*, 8, 1–23.

Siegrist, M., and Cousin, M.-E. (2009). Expectations influence sensory experience in a wine tasting. *Appetite*, 52(3), 762–765.

Spence, C. (2020). Wine psychology: Basic & applied. *Cognitive Research: Principles and Implications*, 5(22), 1–18.

Spence, C., and Wang, Q. J. (2019). Wine expertise: Perceptual learning in the chemical senses. *Current Opinion in Food Science*, 27, 49–56.

Storchmann, K. (2012). Wine economics. *Journal of Wine Economics*, 7(1), 1–33.

Veale, R., and Quester, P. (2008). Consumer sensory evaluations of wine quality: The respective influence of price and country of origin. *Journal of Wine Economics*, 3(1), 10–29.

Wang, Q. J., and Spence, C. (2019). Drinking through rosé-coloured glasses: Influence of wine colour on the perception of aroma and flavour in wine experts and novices. *Food Research International*, 126, 108678.

Weil, R. L. (2007). Debunking critics' wine words: Can amateurs distinguish the smell of asphalt from the taste of cherries? *Journal of Wine Economics*, 2(2), 136–144.