British Food Journal

"Quality does not sell itself": Divergence between "objective" product quality and preference for coffee in naïve consumers

Davide Giacalone Toke Reinholt Fosgaard Ida Steen Morten Münchow

BFJ
118,10

2462

Received 28 March 2016
Revised 10 June 2016
Accepted 11 June 2016

"Quality does not sell itself"

Divergence between "objective" product quality and preference for coffee in naïve consumers

Davide Giacalone
Department of Technology and Innovation, Faculty of Engineering, University of Southern Denmark, Odense, Denmark
Toke Reinholt Fosgaard
Department of Food and Resource Economics, University of Copenhagen, Frederiksberg, Denmark, and
Ida Steen and Morten Münchow
CoffeeMind Aps, Valby, Denmark

Abstract

Purpose – Food quality is a multi-dimensional concept comprising both objective and subjective components. Quality as defined from an industry perspective usually relies on different instrumental assessment and on ratings of "experts" which may not necessarily align with consumers' perception of quality. The purpose of this paper is to deal with consumers perceptions of intrinsic quality in coffee from a sensory scientific and behavioral economic perspective.

Design/methodology/approach – In a blind taste test ($n = 205$), naïve consumers tasted two cups of coffee and decided which they preferred. The two coffees varied greatly in their "objective" quality (based on expert grading) and retail value. Consumers were then revealed that one of the cups contained a coffee that was much more expensive than the other, and that they could get a free cup of their preferred coffee if they could correctly identify the most expensive one.

Findings – The results showed that preferences were equally distributed among the high- and low-quality samples, and that consumers did not perform better than chance level in the identification task. These results suggest that current grading systems used in the industry may be poorly correlated with the way consumers actually experience coffee, and thus that quality inference in the marketplace is more likely influenced by external cues (e.g. brand, label and price) than to intrinsic product quality. Nevertheless, the results also show that consumers who correctly answered the identification task were also significantly more likely to prefer the high-quality sample. This tentatively suggests that better sensory expertise is correlated with a preference for higher quality, though future studies are needed to confirm the correctedness of this interpretation.

Originality/value – This work highlights the difficulty of objectively defining food quality, and the limited usefulness of experts' ratings widely used in the industry. Managerial implications of these findings, as well as implications for consumer policy, are discussed.

Keywords Perception, Preference, Coffee, Behavioural economics, Consumer studies

Paper type Research paper

1. Introduction

1.1 Consumers' perception of food quality

Food and beverage products available to consumers are offered in increasingly many variations, with many factors guiding consumers' navigation through this huge landscape of possibilities (Scheibehenne *et al.*, 2010). External factors such as price and brand are often used as cues by consumers for differentiating between competing products, inferring quality and other attributes, and drive purchase decisions (Deliza and Macfie, 1996; Dickson and Sawyer, 1990; Verdú-Jover *et al.*, 2004).

British Food Journal
Vol. 118 No. 10, 2016
pp. 2462-2474

0007-070X
DOI 10.1108/BFJ-03-2016-0127

The concept of food quality, as well as its measurement and implications for consumer behavior, has received considerable attention in previous research (e.g. Grunert, 2005). There is general agreement that food quality has both an objective and a subjective component. The former denote product-related factors such as the physico-chemical, sensory and nutritional characteristics of foods and beverages, whereas the later refers to the subjective perception of the final consumers (Cardello, 1995a, b; Lawless, 1995; Sáenz-Navajas *et al.*, 2013).

Consumers generally ascribe quality in food and beverages to four basic concepts: sensory quality (i.e. taste and other sensory characteristics), healthiness, convenience, and – for some consumers and some product categories – process characteristics (e.g. organic, GMO-free, etc.) (Brunsø *et al.*, 2002; Grunert, 2005). In the food industry, most discussions of food quality would revolve around physical and chemical parameters that can be measured instrumentally. This is in part because food professionals are trained to adhere to fixed standards (e.g. in order to achieve consistent production and/or match relevant regulations), and in part because such objective quality aspects are much easier to measure than subjective consumers' perceptions, which are often situation-specific and changing across time (Lawless, 1995).

Nevertheless, the relationship between the two is crucial for the economic importance of quality: only when objective quality is recognized as such by consumers (and hence adds value for them), will it constitute a competitive advantage for food producers (Grunert, 2005). High-quality products will lead to higher sales only if consumers perceived the quality to be high enough to pay the higher price commanded by these products, in other words, they must have a good value for money (Grunert, 2005; Zeithaml, 1998). Unfortunately, there is rarely a one-to-one correspondence between objective and subjective quality, and indeed consumers' quality perception often weights more heavily on their individual expectations and past experience than on intrinsic product characteristics.

The latter point is important especially when considering the sensory quality of foods, because taste preferences are primarily learned upon repeated exposure (e.g. Birch, 1999; Köster and Mojet, 2007). For sensory quality, empirical evidence points at disconnection between consumer preferences and evaluation of quality by food professionals, that is, consumers often like products that experts consider to be of poor quality (Delgado and Guinard, 2011; Sáenz-Navajas *et al.*, 2013, 2015). The likely explanation is that average consumers may not have access or experience with finer, high-end examples of specific products, and hence may lack the necessary frame of reference to judge food quality (Lawless, 1995). Accordingly, evidence suggests that, given such experience, consumers learn to recognize the appeal of higher-end products and acquire new preferences for these (Giacalone *et al.*, 2014; Köster and Mojet, 2007; Lawless, 1995; Sáenz-Navajas *et al.*, 2013).

1.2 Motivation for the present research

Food quality has been an important issue for researchers in many disciplines, and indeed there are several papers addressing consumers' perception of quality and how it relates to "objective" quality in several important product categories such as wine (Sáenz-Navajas *et al.*, 2013, 2015; Verdu-Jover *et al.*, 2004), cheese (Grunert *et al.*, 2000; Hersleth *et al.*, 2004), and olive oil (Delgado and Guinard, 2011). However, there is a surprising paucity of research on this issue for several other commercially important product categories.

The focal product category in this paper is coffee. Coffee lends itself very well to experimental investigations of quality perception, for several reasons. It is the most

important beverage commodity by trade volume on the global market, and the tenth-biggest food and beverages commodity overall (Foodprocessing-Technology.com, 2014). Importantly coffee is sold at very diverse quality and price points, and therefore it creates an interesting case to study how well people can experience, and appreciate, quality differences. Furthermore, coffee is consumed by many, with almost every person having tried it at some point. An interesting element of assessing coffee quality is that it to a large extent is not objective (Feria-Morales, 2002), and hence an interesting question is whether consumers have the necessary ability to recognize quality differences in coffees.

Objectively assessing coffee quality is not a trivial task and even major bodies like the Specialty Coffee Associations of Europe (SCAE) and America (SCAA) or the International Coffee Organization lack an exact definition of, e.g., what constitutes a "specialty" or "gourmet" vs "commodity" quality. Instead, quality is practically assessed at the green coffee stage, roasting, brewing, and tasting. Additionally, retail price is also often used as a proxy indicator for quality. Common indicators associated with high quality (HQ) coffee at different stages of the value chain are reported in Table I.

From a sensory perspective, coffee is a complex matrix comprising a wide range of flavor active components, making it difficult to evaluate its quality by instrumental methods alone. Accordingly, sensory evaluation, usually conducted by experts, is a very important part of coffee quality evaluation. Coffee experts, sometimes called "cuppers," are widely employed in the coffee trade and most quality grading systems are based on their opinion (Di Donfrancesco *et al.*, 2014; Feria-Morales, 2002). "Experts" are defined in this context as individuals who have a long standing experience with a product and whose expertise is called upon for quality control and product development purposes (Lawless, 1984). Let us note here that this definition of expert is rooted in the historical tradition of quality grading by, e.g., government inspectors, and is contrast to the use of properly trained sensory evaluation panels, where several assessors are trained to identify and quantify specific sensory attributes (Lawless and Heymann, 2010).

There is a dearth of research conducted on whether experts' ratings correlate with proper sensory evaluation results, with two exceptions: a study by Di Donfrancesco *et al.* (2014) comparing results from "cupping" (sensory evaluation by coffee experts) and descriptive sensory analysis, and a review paper by Feria-Morales (2002) on green coffee quality. Both papers indicate that there is little overlap between experts' ratings and sensory analysis, and that these two approaches are not interchangeable. There are several reasons as to why this might be the case, ranging from the relative narrow set of attributes considered in expert grading systems, to poor sensory practices (e.g. coffee tasting by experts is not always been done in blind) that may bias the results

Table I.
Some common indicators for evaluation of coffee quality at different stage of the product value chain

Stage	Quality indicator	Reference
Green coffee	Absence of defects (broken beans, infected beans, size in itself and size variation)	Wintgens (2009)
Roasting	Slow roast, 12-15 minutes, rather than flash (3-4 minutes) roast	Huschke (2007)
Brewing	Brewing at 20% extraction of 60 g/liter rather than 30% extraction on 30 g/liter	Clarke and Vitzthum (2001)
Tasting	Sensory (analytical and/or hedonic) evaluation can be performed to evaluate quality. SCAA's cupping form, a hedonic scoring system, is the most widely used tool in the coffee industry	Di Donfrancesco *et al.* (2014)
Retail	Price could be used as an indicator	Daviron and Ponte (2005)

(Feria-Morales, 2002), and finally to the fact that many popular cupping protocols contain hedonic attributes and, therefore, rather than being an objective quality measure with universal validity, they may be in fact reflect the preferences of the coffee professionals.

More importantly, whether "quality" defined from a coffee industry perspective (e.g. experts' ratings) actually aligns with consumers' perception of quality is hitherto completely unknown. To address this gap, we report here on a simple experiment devised to investigate the ability of naïve consumers to differentiate between two coffees widely different in terms of (objective) quality; their preference for either low or HQ coffee; and the relationships between the former two.

Our chosen benchmark was a blind taste test, in order to allow consumers ($n = 205$) to evaluate the two coffees independently of extrinsic factors and based primarily on their sensory perception (Lawless and Heymann, 2010). Since price considerations were not present in the current investigation, and the quality differences between the two products were substantial, it was conjectured that the vast majority of consumers would be able to detect taste differences, detect which one is the highest quality, and prefer the highest quality. In order to guide the results of this trial to focus on the quality parameter of the coffee, the samples were chosen to vary in quality within the same basic sensory categories. Coffee from the low quality (LQ) range is, from a flavor perspective, always characterized by a substantial bitterness and "dark flavors" like chocolate, earth and nuts. Coffees from the very high end of the quality spectrum often have very pronounced acidic and fruity notes that are unfamiliar to the average consumer and would likely be rejected from our intended sample population for the reasons outlined in the previous section. Therefore, to circumvent this problem, we carefully chose the high-quality coffee to have similar flavor characteristics (chocolate, nuts and low acidity) but just HQ ingredients and preparation processes.

Although price was not part of what consumers responded to, we decided to introduce an incentive to increase the validity of the results, namely consumers could walk away with a cup of the coffee they preferred if they could correctly identified which one was the more expensive of the two. This idea was inspired by a basic tenet in behavioral economics that choices and preferences expressed during a test have a higher validity (*vis-à-vis* simply declaring them) if participants have some self-relevant consequences once the experiment is over. Combining features of economic experiments and sensory test, although not yet commonplace, is widely thought to improve the validity of product tests with consumers (Combris *et al.*, 2009; Mueller *et al.*, 2010).

To the best of our knowledge, this is the first to study perceptions of coffee quality in a behavioral economics framework, with the possible exception of one study by Sörqvist *et al.* (2013), who also used experiments to study how consumer perception of coffee could be biased by providing external information (in that case, by adding an eco-label to one of two otherwise identical coffee samples). In contrast, we address the ability to detect quality differences correctly, and measure preferences for quality.

To further increase the external validity of the results, the experiment was conducted in a natural location (see Section 2.3) closer to the model of "living labs" than to the central location testing in which most consumer acceptance are conducted (Giacalone *et al.*, 2013; Wendin *et al.*, 2015).

2. Methods

2.1 Samples

The samples used in this study were two coffee beverages selected to differ widely in terms of quality, price points, and experts ratings of sensory quality based on SCAA's

hedonic scoring system (the most widely used tool in the industry to evaluate coffee quality, and most notably in the Cup of Excellence competitions). This rating system gives a composite score which takes into account the quality of individual sensory characteristics of coffee (namely, Fragrance/Aroma, Flavor, Aftertaste, Acidity, Body, Balance, Uniformity, Clean Cup, Sweetness) as evaluated by a certified coffee expert, as well as an evaluation on the presence of defects and an overall quality evaluation. The ratings are done on a scale from 6 to 10 (where 6 = good and 10 = outstanding) and containing quarter point increments between numeric values within this interval (making it a 16-point scale). This procedure results in a composite score takes values between 0 and 100, where scores above 80 generally denote HQ or gourmet coffee. Both samples were evaluated according to this protocol by the third author, who is a certified expert cupper, prior to the experiment. For details on this cupping protocol (see Specialty Coffee Association of America, 2015).

The LQ sample was a pre-ground coffee, rated at 65 in the SCAA hedonic scoring system. From a coffee professional perspective, this coffee was under dosed and over extracted which is typical for LQ coffee. By contrast, the HQ sample had a combined score of 83 on the on the SCAA scale, and was whole beans until just before the brewing process for maximum preservation of all volatile aromas. The brew was trimmed to comply with the Golden Cup Standard (TDS: 1.51, extraction degree: 21.3 percent) which are designed to uphold the highest standard for coffee excellence. The LQ coffee was, as it is usual for LQ coffee, pre-ground and it was not possible to adjust the brew to fall within the Gold Cup Standard[1]. The fact that this coffee was a pre-ground supports in itself its status as LQ coffee as this means that several desirable coffee aromas will evaporate before being trapped in the brew.

Finally, the large difference in quality is also evident in the commercial value of these two coffees: at the time when the study was conducted, the LQ coffee was priced at €6/kg, whereas the HQ coffee was priced at €22/kg.

The coffees were brewed at a nearby café (Café Europa) prior to the experiment. The brewing method was drip coffee brewed on a BUNN ICBA brewer (Bewley, UK). The coffee was stored in Bunn thermoses, which ensure a consistent serving quality within a 2 hours window, and served at a temperature of approximately 70°C.

2.2 Participants

A convenience sample of consumers ($n = 205$; 50.2 percent women) took part in the study. All consumers were at least 18 years, with a relative prevalence of younger consumers (18-30: 43.4 percent; 31-50: 33.7 percent; 51+: 22.9 percent). All participants self-identified as coffee drinkers and gave verbal voluntary consent to participating in the study.

2.3 Experimental procedures

The experiment was carried out at one of Copenhagen's major train stations – Nørreport station – on a single day. A stand with a "free coffee" sign was placed in a busy walking area to attract by-passers. As participants approached, one experimenter introduced them to the broad topic of the experiment, and noted their age and gender. Participants received two small samples (20 ml) of black coffee, one HQ and one LQ, and were asked to taste them and indicate which of the samples they preferred. They were handed the two coffee cups (30 ml black plastic cups. To the consumer the cups appeared unlabeled but a white dot in the bottom of the cup helped us keep track of the samples) simultaneously and were free to taste them in any order. They were also

explained that afterwards they would be asked a question with a clearly right or wrong answer and that, if they could answer correctly, they could walk away with a full 250 ml to-go cup of their preferred coffee sample. Up until this point, no information about the coffee samples was disclosed to the participants, allowing us to get an unbiased measure of their coffee preferences solely based on the flavor of the coffee.

After consumers tasted both samples and indicated which one they preferred, one of the experimenter revealed to them that one of the coffees they had just tasted was substantially more expensive than the other. They were then asked to identify which one they thought was the most expensive and, if they answered correctly, they received a full cup of the coffee they had just indicated as their preferred one. Most participants completed the experiment in less than 2 minutes.

Ethical approval was not sought for this study because it is not required according to the relevant national regulations[2].

3. Results

The data from our two primary measures (quality identification and coffee preference) were cross-tabulated in a 2×2 matrix dividing consumers into four groups:

(1) consumers who correctly identified the HQ sample and preferred it;

(2) consumers who correctly identified the HQ sample but preferred the LQ sample;

(3) consumers who failed to identify the HQ sample but preferred it; and

(4) consumers who failed to identify the HQ sample and preferred the LQ sample.

Recall that, because of the magnitude of the difference between the two coffee samples, we expected option 1 to show the highest frequency. That is, we expected that most consumers would be able to correctly tell which of the coffee was of the highest quality, and that among the consumers who can identify the HQ coffee, most of them would also prefer it to the LQ coffee. However, the actual data showed a quite different picture. As showed in Table II, the distribution of the participants is rather even across all groups. Only about half of the participants (52.7 percent) correctly identified the HQ sample as being the more expensive coffee. A binomial probability test indicated that consumers did not perform better than the level we would expect from merely guessing ($p = 0.485$, two-sided). This indicates that consumers did not generally detect a quality difference between the two coffee samples, in spite of the large differences in intrinsic quality.

A similar response distribution was observed regarding preferences, with only slightly less than half of the participants indicating a preference for the HQ coffee (Table II). Interestingly, the proportion of participants who could not correctly identify

Table II. Cross-tabulation reporting percentages of consumers ($n = 205$) that were able/unable to detect and that preferred high vs low quality coffee

	Preference		
	High quality	Low quality	Total
Quality detection			
Correct	27.8% (57)	24.9% (51)	52.7% (108)
Not correct	18.5% (38)	28.8% (59)	47.3% (97)
Total	46.3% (95)	53.7% (110)	100% (205)

Note: Numbers in parenthesis indicate the number of observations per cell

the HQ sample and preferred the HQ sample (38/97 = 39.2 percent) was lower than the corresponding proportion among the consumers who correctly identified the HQ sample (57/108 = 52.8 percent). The difference was marginally significant according to a Pearson's χ^2 ($\chi^2_{(1)} = 3.8$, $p = 0.051$), which would suggest that participants who were better at evaluating coffee quality had a tendency to prefer the HQ sample. To test this hypothesis, we carried out a probit regression model explaining consumers' preference for the HQ sample on the basis of their performance in the quality identification task (correct answer: Yes/No). The results, reported in Table III, indicate that indeed the correct identification significantly increased to the likelihood of preference for HQ. However, we suggest caution in the interpretation of this result as it might be related to the experimental conditions (see Section 4.1). The overall effect size (as measured by the pseudo R^2) was very limited, suggesting that several factors determine consumers' preference, in addition to product quality *per se*. The model also included gender and age as explanatory variables; however, the results showed that neither of these variables was significantly related to preference (Table III).

Taken overall, the results show that preference for quality and the ability to detect quality are far from clear-cut phenomena. The participants were not able to reliably identify the HQ sample, and preferences were equally distributed across the two samples. Importantly, the half who indeed can detect quality was not the same as the half who preferred HQ, and many of those who could identify the HQ sample actually preferred the LQ.

4. Discussion

Contrary to our expectations, our field experiment showed that, by and large, naïve coffee consumers were not able to detect quality differences in a blind test between two samples widely different in intrinsic quality (as defined from an industry perspective). Furthermore, many of the consumers who correctly identified the HQ sample actually preferred the LQ sample. Since price consideration were absent in this experiment, we conclude that perceived quality in isolation is not a strong driver of consumers' preferences – a conclusion supported by the limited proportion of variance explained by the probit regression model.

These results confirm the difficulty of objectively defining coffee quality, and the limited usefulness of experts' ratings previously highlighted in the literature (Di Donfrancesco *et al.*, 2014; Feria-Morales, 2002). Our work on coffee is aligned with previous results which show that product quality perceptions of experts and consumers may often diverge, as demonstrated in previous work for other important food and beverage categories, such as wine (Sáenz-Navajas *et al.*, 2013).

Table III.
Probit regression ($n = 205$) table modeling the probability of preferring high quality based on correct identification of quality differences and gender

Dependent variable: quality preference (1: prefer HQ, 0: prefer LQ)			
Predictor	Coefficient	SE	Sig.
Constant	−0.31	0.17	0.04*
Correctly identified HQ (1: yes, 0: no)	+0.37	0.18	0.04*
Gender (1: female, 0: male)	−0.11	0.18	0.53ns
Age (1: < 50; 0: > 50)	+0.34	0.21	0.10ns

Notes: Robust standard errors are reported in parentheses next to the regression coefficients. McFadden Pseudo $R^2 = 0.023$. Significance levels: ns, not significant; *$p < 0.5$; **$p < 0.1$; ***$p < 0.001$

In a recent paper, Dijksterhuis notes that many professionals in food and FMGC companies hold the pervasive, and erroneous, belief that HQ products will automatically lead to high sales (Dijksterhuis, 2016). Our research shows that if such definition of quality only takes into account the product itself (i.e. without external cues such as price and brand), then this belief is unjustified and possibly misleading: in fact, it might be partly responsible for the fact that more than 75 percent of all food products launched in the market fail to fulfill their expected financial returns (Nielsen, 2015).

Overall, our findings suggest that people are not able to point out high and LQ coffees: their perception of coffee quality mostly looks like a random process. More to the point, our study suggests that even if one can detect quality and price concerns are not present, HQ may not be the obvious choice. Indeed, the most interesting group was probably the group that correctly identifies the HQ sample but prefers the LQ sample. We think there are at least two possible (non-mutually exclusive) explanations for this finding.

One possibility is that preference might be more related to typicality, rather than to HQ *per se*. The LQ sample is closer, in terms of flavor, to the coffee consumers are ordinarily exposed to. Since sensory preferences are mostly learned upon repeated exposure, the LQ sample might have been preferred because its sensory profile was closer to expectations than the more novel HQ sample. This implies that, although a consumer might be able to recognize the higher quality from a perceptual point of view, they might nonetheless prefer the more familiar of the two. This explanation is in line with the results of Giacalone *et al.* (2014). In a series of studies where consumers were exposed to a range of widely different beers, they found that expert/knowledgeable beer consumers rated the samples as significantly less novel and complex than novices consumers did, and yet no differences between these two groups were found with regards to overall liking for the beers. Unfortunately, it is not possible to conclude with confidence that this is the case on the basis of the results presented, since they are based on a convenience sample of consumers.

The other possible explanation is that there might have been differences in drinkability between the two coffee samples that influenced the preference of this group of consumers, irrespectively than the actual pleasantness of the sample. The concept of drinkability in a beverage is defined as the characteristic of not making one tired of drinking some or a lot of it. Although there is no convenient definition of drinkability, it is related to, e.g., sensory specific satiety, i.e., how much taking a sip of a beverage reduces the pleasure derived from the next sip. Assuming that the HQ was more rich and complex in flavor while the LQ might have been more bland, consumers (who knew the potential reward was a full cup of either one), might have chosen the one they thought would be more pleasant to drink until the last drop, rather than the one that was more flavorful.

On a related note, in the experimental conditions for this study the incentive for the consumers was that they could get a full cup to-go of the coffee they preferred. Clearly, drinking a beverage "on the go" is a form of consumption that does not encourage appreciation of the sensory properties of coffee, but rather its functional properties. As observed by Steenkamp (1990), sensory attributes of coffee will be particularly important in consumer choice if the experience sought is hedonic satisfaction, while other functional properties (e.g. keep oneself warm or awake) may be more important in other situation. Hence, some consumers may have preferred the LQ coffee as more "situationally appropriate" for consumption on the go than the more flavorful HQ coffee. This explanation is consistent with extant research in consumer perception of food and beverage demonstrating that situational appropriateness of products is poorly correlated with intrinsic product quality (Giacalone *et al.*, 2015; Giacalone and

Jaeger, 2016), and that different sensory profiles within one product category might be preferred depending on the context in which the consumption takes place (Di Monaco *et al.*, 2014; Lawless, 1995).

4.1 Limitations and directions for future research

This study has several limitations that should be acknowledged, as they leave some of the conclusions outlined above open to alternative explanations.

The first one is that this study is based on a convenience sample of consumers for which only limited background data were collected. In follow-up work, purposeful sampling should be used to look at difference between group segments defined in terms of product usage and expertise. For example, the hypothesized relationship between product exposure and preference should be explored in future research by recruiting consumers varying in consumption patterns (e.g. regular LQ vs regular HQ consumers). Likewise, since only two samples were tested, follow-up studies spanning a larger variation in terms of sensory profiles and quality differences would be needed to confirm the generalizability of the findings.

A second important limitation is related to the experimental choice to use price as a proxy for quality in our field experiment. This choice was motivated by the fact that it was deemed easier to probe consumers about expected price differences between, than to ask them directly to evaluate "quality". Although price is undoubtedly positively related with quality, it is only one facet of it and in real life product quality will also be influenced by other factors such as brand name, packaging and store name (Rao and Monroe, 1989). Therefore, future studies with more complex designs are advised to look at the effect of these factors, as well as their interactions, in order to estimate the importance of their individual contribution to determining quality perception of coffee. Labeling should be a particularly factor to include as a covariate in future studies, since recent findings point at an important role in determining perceptual experiences and choice for coffee (Sörqvist *et al.*, 2013; Van Loo *et al.*, 2015). Another interesting direction for future work would be to replicate the study in a different setting location, such as a high-end coffee estate, to look at possible context dependencies with regards to preferences for HQ coffee.

Last but not least, an important limitation to the present study is related to the task order during the field experiment and the finding that the consumers who correctly recognized the HQ coffee were more likely to express a preference for the HQ sample. Since consumers first expressed their preference and only then were asked about price differences, an additional explanation of this finding might be that consumers could be more likely to report that they prefer the taste of the more expensive beverage. Hence, the fact that 39.2 percent of participants preferred the HQ sample but did not say it was the most expensive, could indicate that only a minority of the participants diverged from the arguably socially desirable response. In turn, the 52.8 percent who said they preferred the HQ sample and identified it as the most expensive were simply responding in the rational way. Therefore, future studies looking more closely at the relationship between price and perceived quality are also necessary. A possibility to assess the impact of the social desirability bias could be, for example, to conduct repeated preference tests by providing (in lure) different price levels to the respondents, and to find out whether this leads to differential preference distributions.

5. Conclusions

This work has investigated consumers' perception of coffee quality (defined in terms of expert scores and retail price) and its relationship to preference from a behavioral

economics framework. The main results obtained were that consumers were generally not able to recognize HQ coffee in a blind test; preferences for high vs LQ coffee were equally distributed; but among those who correctly recognized the HQ coffee, there was a tendency to prefer HQ.

Taken overall, these results suggest that many coffee consumers may currently lack the sensory skills to recognize HQ coffee, and are therefore more likely to rely on external cues, such as price, packaging and advertising, that may or may not reflect the intrinsic quality of the product. Hence, from a consumer policy perspective, this study suggests that additional/better way to inform consumers about coffee quality are needed to protect their welfare. Importantly, our findings suggest that current sensory grading system based on experts' opinion to characterize quality (in the present case, the SCAA scale) are poor predictors of the way the product is experienced by consumers. From a sensory scientific perspective the main objection to the SCAA scoring form is that it contains hedonic attributes and, therefore, rather than being an objective quality measure with universal validity it is more comparable to a consumer test reflecting the preferences of the coffee professionals. Hence, the discrepancy between the preferences in our experiment and the SCAA scoring can be best interpreted as hedonic disagreement between consumers and coffee professionals. This is a serious limitation of the usefulness of these tools, and hence, future research should look into developing grading systems that are truly correlated with the way consumers actually appreciate such quality standards.

From a managerial perspective, our findings have clear implications for coffee producers, particularly in the specialty coffee market. To a large extent, our findings suggest that consumers may experience coffee as "just coffee", i.e., as a mere commodity with little potential for differentiation. This could pose a problem in terms of product adoption for HQ coffees: if consumers do not experience any sensory benefits from drinking HQ coffee, they might be reluctant to pay a premium price for such products. Moreover, HQ coffee might not live up to expectations generated by the premium price, and thus such products may never move beyond the trial purchase (as opposed to repeat purchase) phase. All in all, coffee producers would be better off by reducing their reliance on experts' ratings, and giving up the belief that quality (as defined from their perspective) will somehow sell itself.

Another interesting point for coffee professionals in the specialty coffee business whose mission it is to turn consumer preference from LQ to HQ is that, from these results, it seems that among the LQ preference group there are a least two different consumer types: those who do not have the perceptual skills to distinguish LQ from HQ and those who do. These two groups pose very different challenges and probably require very different strategy for moving their preferences toward the HQ.

Nevertheless, the results also indicated that consumers who are better at identifying quality differences are also more likely to prefer HQ coffee, suggesting that better sensory skills increase appreciation of quality (although future studies are needed to confirm this explanation). To the extent that this can be leveraged by producers, a promising way to go for the specialty coffee business would be to improve the level of "sensory literacy" in the consumer population. Several initiatives from the coffee specialty sector go into this direction: for example, the SCAA is currently engaged in several collaborative initiatives, such as the Coffee Flavor Wheel (Counter Culture Coffee, 2013) and the World Coffee Sensory Lexicon (World Coffee Research, 2016), aimed at establishing common vocabulary for communicating coffee flavor, both within the industry and to the final consumers. In other product categories

(most notably in wine), these tools have been widely instrumental in enhancing public understanding on food tasting and terminology and increase people's reliance on their own perception.

Notes

1. Brewing parameters are available from the authors upon request. The extraction percentage was calculated using a VST LAB Coffee II Refractometer (VST CoffeeTools™). The LQ coffee could not be handled by the Refractometer due to the very fine ground level of the pre-ground coffee (in the coffee jargon this is referred to as "over-extraction"). Upon tasting the coffee, we compensated for this problem by using less coffee ("under-dosing") and ended up with a dosage of 45 g/L.
2. According to the Danish National Committee on Health Research ethics, sensory evaluation studies (incl. consumer acceptance) do not need formal approval as long as the target food is consumed in quantities at or below safe use. Additionally, sensitive data (as defined by the Danish Data Protection Agency) about the participants were neither collected nor stored in this study.

References

Birch, L.L. (1999), "Development of food preferences", *Annual Review of Nutrition*, Vol. 19, pp. 41-62.

Brunsø, K., Fjord, T.H. and Grunert, K.G. (2002), "Consumers' food choice and quality perception", MAPP Working Paper No. 77, Aarhus School of Business, Aarhus.

Cardello, A.V. (1995a), "Consumer expectations and their role in food acceptance", in MacFie, H.J.H. and Thomson, D.M.H. (Eds), *Measurement of Food Preferences*, Springer, London, pp. 253-297.

Cardello, A.V. (1995b), "Food quality: relativity, context and consumer expectations", *Food Quality and Preference*, Vol. 6 No. 3, pp. 163-170.

Clarke, R. and Vitzthum, O.G. (2001), *Coffee: Recent Developments*, Wiley-Blackwell Publishing Ltd, Oxford.

Combris, P., Bazoche, P., Giraud-Héraud, E. and Issanchou, S. (2009), "Food choices: what do we learn from combining sensory and economic experiments?", *Food Quality and Preference*, Vol. 20 No. 8, pp. 550-557.

Counter Culture Coffee (2013), "Coffee taster's flavor wheel", available at: https://counterculturecoffee.com/learn/coffee-tasters-flavor-wheel (accessed January 12, 2016).

Daviron, B. and Ponte, S. (2005), *The Coffee Paradox: Global Markets, Commodity Trade and the Elusive Promise of Development*, Zed Books, London and New York, NY.

Delgado, C. and Guinard, J.X. (2011), "How do consumer hedonic ratings for extra virgin olive oil relate to quality ratings by experts and descriptive analysis ratings?", *Food Quality and Preference*, Vol. 22 No. 2, pp. 213-225.

Deliza, R. and Macfie, H.J.H. (1996), "The generation of sensory expectation by external cues and its effect on sensory perception and hedonic ratings: a review", *Journal of Sensory Studies*, Vol. 11 No. 2, pp. 103-128.

Di Donfrancesco, B., Gutierrez Guzman, N. and Chambers, E. (2014), "Comparison of results from cupping and descriptive sensory analysis of colombian brewed coffee", *Journal of Sensory Studies*, Vol. 29 No. 4, pp. 301-311.

Di Monaco, R., Giacalone, D., Pepe, O., Cavella, S. and Masi, P. (2014), "Effect of social interaction and meal accompaniments on acceptability of sourdough-prepared croissants: an exploratory study", *Food Research International*, Vol. 66, pp. 325-331.

Dickson, P.R. and Sawyer, A.G. (1990), "The price knowledge and search for supermarket shoppers", *Journal of Marketing*, Vol. 54 No. 3, pp. 42-53.

Dijksterhuis, G.B. (2016), "New product failure: five potential sources discussed", *Trends in Food Science & Technology*, Vol. 50, pp. 243-248.

Feria-Morales, A.M. (2002), "Examining the case of green coffee to illustrate the limitations of grading systems/expert tasters in sensory evaluation for quality control", *Food Quality and Preference*, Vol. 13 No. 6, pp. 355-367.

Foodprocessing-Technology.com (2014), "The ten most traded food and beverage commodities", (accessed January 29, 2016).

Giacalone, D. and Jaeger, S.R. (2016), "Better the devil you know? How product familiarity affects usage versatility of food and beverages", *Journal of Economic Psychology*, Vol. 55, pp. 120-138.

Giacalone, D., Bredie, W.L.P. and Frøst, M.B. (2013), "All-in-one test"(AI1): a rapid and easily applicable approach to consumer product testing", *Food Quality and Preference*, Vol. 27 No. 2, pp. 108-119.

Giacalone, D., Duerlund, M., Bøegh-Petersen, J., Bredie, W.L.P. and Frøst, M.B. (2014), "Stimulus collative properties and consumers' flavor preferences", *Appetite*, Vol. 77, pp. 20-30.

Giacalone, D., Frøst, M.B., Bredie, W.L.P., Pineau, B., Hunter, D.C., Paisley, A.G., Beresford, M.K. and Jaeger, S.R. (2015), "Situational appropriateness of beer is influenced by product familiarity", *Food Quality and Preference*, Vol. 39, pp. 16-27.

Grunert, K.G. (2005), "Food quality and safety: consumer perception and demand", *European Review of Agricultural Economics*, Vol. 32 No. 3, pp. 369-391.

Grunert, K.G., Bech-Larsen, T. and Bredahl, L. (2000), "Three issues in consumer quality perception and acceptance of dairy products", *International Dairy Journal*, Vol. 10 No. 8, pp. 575-584.

Hersleth, M., Illseng, M.A., Martens, M. and Næs, T. (2004), "Perception of cheese: a comparison of quality scoring, descriptive analysis, and consumer responses", *Journal of Food Quality*, Vol. 28 No. 4, pp. 333-349.

Huschke, R. (2007), *Industrial Coffee Refinement*, Sv Corporate Media.

Köster, E.P. and Mojet, J. (2007), "Theories of food choice development", in Frewer, L. and Van Trijp, H.C.M. (Eds), *Understanding Consumers of Food Products*, Woodhead Publishing, Cambridge, pp. 93-124.

Lawless, H.T. (1984), "Flavour description of white wine by 'expert' and non-expert wine consumers", *Journal of Food Science*, Vol. 149 No. 1, pp. 120-123.

Lawless, H.T. (1995), "Dimensions of sensory quality: a critique", *Food Quality and Preference*, Vol. 6 No. 3, pp. 191-199.

Lawless, H.T. and Heymann, H. (2010), *Sensory Evaluation of Food: Principles and Practices*, 2nd, ed. Springer, New York, NY.

Mueller, S., Osidacz, P., Leigh Francis, I. and Lockshin, L. (2010), "Combining discrete choice and informed sensory testing in a two-stage process: can it predict wine market share?", *Food Quality and Preference*, Vol. 21 No. 7, pp. 741-754.

Nielsen (2015), "Breakthrough innovation report", European edition, September, available at: www.dansk-dlf.dk/pageimages/files/Nielsen%20Breakthrough%20Innovation%20Report%202015%20European%20Edition_digital.pdf (accessed January 15, 2016).

Rao, A.R. and 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*, Vol. 26 No. 3, pp. 351-357.

Sáenz-Navajas, M., Ballester, J., Pêcher, C., Peyron, D. and Valentin, D. (2013), "Sensory drivers of intrinsic quality of red wines: effect of culture and level of expertise", *Food Research International*, Vol. 54 No. 2, pp. 1506-1518.

Sáenz-Navajas, M., Avizcuri, J.M., Ballester, J., Fernandez-Zurbano, P., Ferreira, V., Peyrone, D. and Valentin, D. (2015), "Sensory-active compounds influencing wine experts' and consumers' perception of red wine intrinsic quality", *LWT – Food Science and Technology*, Vol. 60 No. 1, pp. 400-411.

Scheibehenne, B., Greifeneder, R. and Todd, P.M. (2010), "Can there ever be too many options? A meta-analytic review of choice overload", *Journal of Consumer Research*, Vol. 37 No. 3, pp. 409-425.

Sörqvist, P., Hedblom, D., Holmgren, M., Haga, A., Langeborg, L., Nöstl, A. and Kågström, J. (2013), "Who needs cream and sugar when there is eco-labeling? Taste and willingness to pay for 'eco-friendly' coffee", *PLoS ONE*, Vol. 8 No. 12, pp. 1-9.

Specialty Coffee Association of America (2015), "SCAA protocols – cupping specialty coffee", available at: www.scaa.org/PDF/resources/cupping-protocols.pdf (accessed March 20, 2016).

Steenkamp, J.B.E.M. (1990), "Conceptual model of the quality perception process", *Journal of Business Research*, Vol. 21 No. 4, pp. 309-333.

Van Loo, E.J., Caputo, V., Nayga, R.M. Jr, Seo, H.S., Zhang, B. and Verbeke, W. (2015), "Sustainability labels on coffee: consumer preferences, willingness-to-pay and visual attention to attributes", *Ecological Economics*, Vol. 118, pp. 215-225.

Verdú Jover, A.J., Lloréns Montes, F.J. and del Fuentes Fuentes, M.M. (2004), "Measuring perceptions of quality in food products: the case of red wine", *Food Quality and Preference*, Vol. 15 No. 5, pp. 453-469.

Wendin, K., Annika, Å. and Anna, S. (2015), "Exploring differences between central located test and home use test in a living lab context", *International Journal of Consumer Studies*, Vol. 39 No. 3, pp. 230-238.

Wintgens, J.N. (2009), *Coffee: Growing, Processing, Sustainable Production. A Guidebook for Growers, Processors, Traders and Researchers*, 2nd ed., Wiley-VHC, Weinheim.

World Coffee Research (2016), *Sensory Lexicon*, 1st ed., World Coffee Research, available at: http://worldcoffeeresearch.org/images/pdfs/WCR_Sensory_Lexicon_Edition_1_2016.pdf (accessed March 1, 2016).

Zeithaml, V.A. (1998), "Consumer perceptions of price, quality and value: a means-end model and synthesis of evidence", *Journal of Marketing*, Vol. 52 No. 3, pp. 2-22.

Corresponding author
Davide Giacalone can be contacted at: dg@iti.sdu.dk