Investigating Consumer Food Choice Behavior: An Application Combining Sensory Evaluation and Experimental Auctions

R. Karina Gallardo[1], Yeon A Hong[2], Marcial Silva Jaimes[3], and Johanna Flores Orozco[3]

Abstract

We investigate what piece of information collected utilizing sensory evaluation tools exhibits better predictive capacity on the willingness to pay, is it information from preferences for a sensory quality attribute using hedonic scales or information on perceived intensity for the same attribute using intensity scales? We also estimated if extrinsic or intrinsic quality exerts a similar impact on consumer's willingness to pay. We conducted a sensory evaluation along with experimental auctions using three different apple varieties with college students in Metropolitan Lima, Peru. Findings from this study show that information collected on preference liking for apple quality attributes has a better explanatory capability for willingness to pay, compared to information on consumers' perceived intensity for the same attribute. The explanatory capability was measured using measures of goodness-of-fit. We also prove that willingness to pay was driven both by the apple variety induced intrinsic quality attributes and the extrinsic cues of the variety. Results add to the existing body of literature aiming to improve the understanding of consumer food choice behavior.

Keywords: willingness to pay, food choice behavior, intensity stimuli, hedonic

---

[1] School of Economic Sciences, Puyallup Research and Extension Center, Center for Precision Agriculture and Automated Systems, Washington State University, 2606 W. Pioneer, Puyallup, WA 98371.
[2] School of Economic Sciences, Washington State University, Hulbert Hall 201A, P.O. Box 646210, Pullman, WA 99164 (phone: (509) 339-4616; email: yeona.hong@wsu.edu).
[3] Facultad de Industrias Alimentarias. Universidad Nacional Agraria La Molina. Av. La Molina S/n. La Molina, Peru.

Introduction

Investigating consumers' food choices is a complex task to the extent that there seems not to be a consensus across disciplines on the best approach to study. A branch of marketing studies postulates that food choice behavior follows a structured process that could be described by different stages including problem recognition, information search, evaluation of alternatives, purchase decision, product consumption, and post purchase behavior (Kotler and Keller, 2012; Grunert, 2005). Whereas a branch of studies on economics and psychology, advocate for a different perspective based on simple heuristics, that is, consumers select or eliminate products based on a few salient attributes rather than using a systematic structured procedure (Rabin, 1998; Simon, 1957). In light of these contrasting perspectives, a popular alternative to improve the understanding of food choice behavior is to combine disciplines, such as sensory science and applied economics. In fact, there are numerous studies that follow such approach: Lange et al, 2002; Lund et al., 2006; Stefani et al., 2006; Combris et al., 2009, Mueller et al., 2010; Gallardo et al., 2011; Zhang et al., 2011; Dinis et al., 2011; Bi et al., 2011; Costanigro et al., 2014. In general, these studies seek to analyze the role of intrinsic and extrinsic food quality attributes on consumer's preferences and valuation for a food product. In these studies, consumers are asked to evaluate the external and internal sensory quality attributes and rate or rank the product and/or its sensory quality attributes in function of their preferences. Next, consumers participate in experiments that would reveal the impact of their preferences on their well-being (Combris et al., 2009). Usually this impact is measured by hypothetical-type questions using stated choice scenarios in a questionnaire format and/or incentive compatible experimental auctions. The importance of evaluating both external and internal sensory quality attributes stems on the postulate that internal quality attributes cannot be experienced at the time of purchase so consumers rely on the external cues and past experiences of the internal attributes. It is believed that past experiences are not stellar in predicting repeated purchases if there is not a rigorous recollection or if these experiences were not consistent (Bi et al., 2011). The aforementioned factors coupled with non-sensory factors –such as convenience, societal values, production technology, personal health, branding, and so on- aim to provide a complete depiction of food choice behavior (Jaeger, 2006).

The information obtained from sensory science is unique as it enables measurements of food sensory characteristics as perceived by humans, different from other sources of information such as chemicals or instruments that are also used to characterize food (Navarro da Silva, 2013). Sensory science uses sensory evaluation as the primary method of analysis (Tuorila and Monteleone, 2009). Sensory evaluation is "a scientific method that evokes, measures, analyzes, and interprets responses to products as perceived through the senses of sight, smell, touch, taste, and hearing" (Lawless and Heymann, 1999). Moreover, sensory science is considered as the intersection of other disciplines, including behavioral sciences, biology, nutrition, and health (Tuorila and Monteleone, 2009).

Among the multiple areas covered by sensory science, evaluation of consumer preferences is an important one. Typically, consumer preferences are measured by the use of hedonic scales. The 9-point hedonic scale: 1= dislike extremely, 9= like extremely, is the most internationally accepted and widely used. This scale was developed in 1947 at the Quartermaster Food and Container Institute for the U.S. Armed Forces. With the scale, word descriptors are used along with numbers facilitating the interpretation of the mean values of the responses in terms of the degree of like/dislike (Lim, 2011). This scale is easy to implement and interpret both by respondents and researchers. However, it has limitations such as the high vulnerability to ceiling effects due to the small number of available categories and the general tendency of subjects to avoid using extreme categories (Lim, 2011).

Besides collecting information on preferences, scales used in sensory science enable collecting information on the chemical stimuli that sensory quality attributes trigger on panelists (Lim, 2011). The rationality of these scales is based on the idea that there is a direct relationship between perceived intensity and stimulus. Such relationship has been long studied in psychophysics, to the extent that current methods in psychophysics are able to capture the range of perceived intensities from threshold to maximum and capture with increased accuracy comparisons of perceived intensities across individuals (Bartoshuk, 2000). There are several ways to measure perceived intensity, for example, 9-point scales (or similar) with word descriptors and magnitude scales that measure the ratio of intensities perceived for one same sensory quality attribute. One main disadvantage of intensity scales is that there is "no provision for anchoring the judgments of individuals subjects to a common ruler" (Lim, 2011). In other words, there are no means to prove that a rating of "9" means the same to all panelists (Lim, 2011).

Given the different pieces of information on consumers' perceptions collected via sensory evaluation techniques: preferences versus perceived intensity of sensory quality attributes, one questions how such information relates with consumers' willingness to pay for a food product. In other words, what piece of information would have a better predictive capacity of consumers' willingness to pay, is it how much each sensory quality attribute is liked? Or is it how intense each sensory quality attribute is perceived? The primary goal of this study is to respond to these inquiries. To achieve this goal we estimate two sets of regressions, one having liking ratings and the other having perceived intensity in the set of explanatory variables. Then, we test which set of regression have a greater explanatory capacity using measures of goodness of fit. We also estimate how coefficients from either set of regressions compare, using non-parametric tests. A second goal of this study is to infer if variety induced sensory quality attributes -the intrinsic quality attributes that could be measured only when using sensory evaluation- or the variety itself -the extrinsic quality attributes- exerts a greater impact on the willingness to pay for a food product. Note that the goal of this paper is not to offer recommendations on general consumers' preferences and willingness to pay for apples, but to test the performance of different sensory evaluation scales when explaining willingness to pay behavior; and to test is the what set of quality attributes (extrinsic or intrinsic) exert a greater impact on the willingness to pay for a food product. We used apples because it is a familiar product to most if not all individuals.

To estimate the willingness to pay, we used a Vickrey second price auction. This type of preference elicitation methodology has the advantage of being incentive compatible. This means that participants face consequences after their bidding behavior, as they are presented incentives to assess and reveal their preferences as truthful as possible (Lusk and Shogren, 2007). The Vickrey second price consists in that every participant submits a bid, or her willingness to pay for the product being auctioned. The participant who submitted the highest bid would win the auction, that is, this participant will actually buy the product being auctioned. The price the winner would pay is the second highest price (selling or market price) (Lusk and Shogren, 2007). The advantage of the Vickery second-price auction, -over other auction formats- is that it is relatively simple to explain to participants, it creates an endogenous market-clearing price, making sure participants are involved in an active market environment exposed to market feedback (Lusk and Shogren 2007).

Methods and procedures

Data collection

The experimental auctions and sensory evaluation were conducted in June 2015 at the facilities of the Universidad Nacional Agraria La Molina in Lima, Peru. One hundred students were recruited two weeks in advance by flyers posted around campus. To participate in the study, individuals had to have eaten apples in the last three months and be in charge of the grocery shopping at home. Using student pools is often questioned. In principle, recruiting college students was more convenient and less costly than recruiting standard household individuals. Besides, the purpose of this study was to compare how liking and perceived intensity of attributes affected willingness to pay, not to derive conclusions about general consumer preferences towards a specific product. Nalley et al. (2006) argue that when deriving consumer preferences is not the central motivation of the study, students perform similarly to other groups in economic experiments.

All apple samples were procured from the same local grocery store. The experiment was conducted in two different sessions, each hosting 50 participants. In each session, individuals were requested to evaluate the three apple samples visually and by tasting; each apple sample was identified with letters D, N, or S. Then participants were asked to respond to a questionnaire describing the intensity and how much they like the visual quality attributes of each sample. Appearance attributes included perceived presence of external defects and size. After evaluating appearance attributes, researchers cut each apple sample given to each participant in two halves. To objectively assess apple size, participants were requested to measure the transversal diameter of each apple with a ruler and write that number as a response to the size question in the questionnaire. Next, panelists were asked to taste each apple sample. For this, the moderator gave a brief explanation of each quality attribute included in the study, for example, what is/how to measure crispness, firmness, sweetness, and acidity. Panelists were given instructions to rinse their palates with water in between tasting each sample to neutralize their taste buds. Next, panelists responded to the questionnaire in which they were required to rate how much they liked the following apple attributes using a 9-point scale (1 = dislike extremely, … , 9 = like extremely): crispness, firmness, sweetness, and acidity. They were also requested to rate the perceived intensity of each of the attributes using a 9-point scale (1 = not intense, …, 9 = extremely intense). Once most participants signaled they had finished responding to the questionnaire, they were requested to submit a bid in nuevos soles per kilo, in two repetitive rounds. Nuevo sol is the Peruvian currency; as of June 18, 2015, $1 was equivalent to 3.16 nuevos soles (Peru, Central Reserve Bank 2015). Following the second price auction, bids were organized in ascending order, and the first and second highest bid were identified along with the panelists submitting such bids. Researchers kept records of the winning bids and did not reveal to participants. To identify the winner of the auction, a binding sample and bid was selected randomly. Once winning sample and panelist were identified, the winning panelist bought 1 kilogram of apples and paid the second highest bid submitted in the session.

Econometric model

Censored bids are common in experimental auctions (Lusk and Shogren, 2007). Results from a censoring test on our bid data indicated that 6% of bids observations were censored. See Figure 1. In addition, likelihood ratio tests were used to test the appropriateness of the Tobit model compared to an OLS and Cragg's double hurdle model. Test results rejected OLS and Cragg's double hurdle in favor of the Tobit model. Results of the likelihood ratio tests to justify the use of the Tobit specification are available upon request to authors.

Coefficient estimates for the Tobit model were estimated by maximizing the likelihood function (LF) that follows (Greene, 2008),

$$LF = \prod_{i=1}^{100} \left(\frac{1}{\sigma}\phi\left(\frac{Bid_i - X_i\beta}{\sigma}\right)\right)^{UC_i} \Phi\left(\frac{-X_i\beta}{\sigma}\right)^{LC_i} \quad (1)$$

where LF is the likelihood function, Bid_i is the bid for panelist *i* (*i*=1, …, 100), X_i is the intensity or likeness rating for each quality attribute (e.g., appearance/ presence of defects, size, crispness, firmness, sweetness and acidity) as perceived by panelist i, β is the coefficient estimate the intensity or likeness rating for each quality attribute, UC_i is the indicator variable for the uncensored bid observations and LC_i is the indicator variable for the left censored bid observations, σ is the square root of the variance of the error terms, ϕ is the standard normal density function, and Φ is the cumulative standard normal distribution function. The censored marginal effects were calculated by,

$$\frac{\partial E[Bid_i|X_i]}{\partial X_i} = \beta\Phi\left(\frac{X_i\beta}{\sigma}\right) \quad (2)$$

where Bid_i is the bid for panelist *i* (*i*=1, …, 100), X_i is the intensity or likeness rating for each quality attribute as perceived by panelist i, β is the coefficient estimate the intensity or likeness rating for each quality attribute, σ is the square root of the variance of the error terms, and Φ is the cumulative standard normal distribution function. Coefficient estimates and marginal effects were calculated in SAS® v.9.2.

Two regressions were estimated based on the empirical specification (1), one including ratings for liking sensory quality attributes and the other including ratings for intensity perceived for sensory quality attributes. To measure which model explained better variations in bids, we used the following criteria of goodness of fit: (i) McFadden likelihood ratio index, (ii) the Akaike Information Criterion (AIC), and (iii) the Schwarz Criterion (SC).

We conducted the non-parametric Mann-Whitney test to compare the ordinal ranking of the coefficient estimates between models including liking and intensity as explanatory variables. For example, if the coefficient estimate for crispness liking was the largest among all coefficients in the model that used liking as explanatory variables and if the coefficient estimate for crispness intensity was also the largest among all coefficients in the regression using intensity as explanatory variables. To further compare how liking and intensity coefficient estimates differ, we conducted a pairwise comparison to compare marginal effects from liking and intensity ratings on bids.

Results

Summary statistics of the liking and perceived intensity for each sensory quality attribute is presented in Table 1. To assess external appearance liking, we asked panelists to rate in the 9-point scale how much they liked the external appearance of apple samples. To assess the "intensity" perceived of external appearance we asked panelists how they perceived the extent of external defects on the apple fruit, in a 9-point scale (1=no defects, 9=abundant defects that I would not buy). In relation to size, the question asked how much panelists liked the fruit size, and when asking for intensity, the actual fruit diameter was used. Pairwise comparisons indicated statistically significant differences across liking and intensity ratings for each sensory quality attribute. Additionally, we conducted a correlation test to assert if liking and intensity ratings were positively correlated. We found a positive correlation between liking and intensity for attributes crispness, sweetness, and acidity. For firmness, the correlation coefficient between liking and intensity was negative and the correlation between fruit size and liking score for size was not statistically significant.

Comparing bids submitted for each apple variety tasted, panelists offered higher bids for variety ‘Royal Gala’, followed by ‘Delicia’ and ‘Fuji” (see Table 2). Pairwise comparisons across bids for each apple variety signal that average bids for ‘Delicia’ were \$0.186 /kg lower than ‘Royal Gala’ and \$0.103 /kg higher than ‘Fuji’. Bids for the ‘Royal Gala’ variety were \$0.289 /kg higher than ‘Fuji’.

Coefficient estimates for the Tobit model are presented in Table 3. To infer if liking or intensity ratings explained better variations on bids, we estimated two regressions, one included liking ratings and the other regression included intensity ratings for each sensory quality attribute in the set of explanatory variables. Results from the McFadden likelihood ratio index, Akaike Information Criterion, and Schwarz Criterion favored the models including liking over the models including intensity ratings (see Table 3). This gives interesting cues as to the rational process followed by panelists, who are willing to pay higher bids for the apple sample they like the most.

To investigate if the variety induced sensory quality attributes or the varietal differences across samples exerted a greater impact on bids, we conducted three sets of regressions, (i) full model including sensory quality attribute and binary indicators for variety, (ii) restricted model including sensory quality attribute variables without binary indicators for variety, and (iii) restricted model including only binary indicators for variety. Likelihood ratio tests to compare the full versus the restricted model led us to reject the restricted model in favor of the full model when liking ratings were included as explanatory variables (likelihood ratio statistic was 6.77, 95% critical chi-square value with 2 degrees of freedom was 5.99). Similarly when intensity ratings were included as explanatory variables (estimated likelihood ratio statistic was 8.35) we rejected the restricted model in favor of the full model.

In the full model, the liking ratings for crispness, sweetness, and binary variable for variety ‘Royal Gala’ had a positive effect on the bids submitted. This was similar to the model using intensity ratings in the set of explanatory variables: intensity ratings for crispness, sweetness, and binary variable for ‘Royal Gala’ were positive and statistically significant. That the coefficients for quality characteristics and variety were statistically significant, indicate that it is both the variety induced quality characteristics and the variety itself that affects bids. The three apple samples presented to panelists were three varieties with different external attributes: the ‘Delicia’ is elongated in shape and red color, ‘Royal Gala’ is red with cream and yellow stripes, and ‘Fuji’ is bicolored yellow and red. All three samples were presented with peels, hence it is possible that panelists recognized these varieties from their external appearance and recalled previous sensory experiences that influenced their preferences and bids. When not including the binary variables for varieties (the restricted model), liking rating for size, crispness and sweetness were positive and statistically significant. For the restricted model including intensity ratings in the set of explanatory variables, only the intensity rating for sweetness was positive and statistically significant.

Implications from these findings are twofold. First, models including liking ratings outperformed the models including intensity ratings in the set of explanatory variables. Liking scores for most attributes included in the model (i.e., appearance, crispness, sweetness, firmness, and acidity), except size, exhibited a statistical significant correlation with perceived intensity scores. However, liking and intensity did not have a similar predictive capability of willingness to pay, liking scores show a higher predictive capacity. The second implication is the importance of the apples external appearance (extrinsic quality) and the possibility that participants recognized and recalled past consumption experiences showing a stronger preference for the apple they recalled they liked the most. If the interest is centered in eliciting willingness to pay for intrinsic sensory quality attributes, it is recommended to present panelists peeled samples, so they could not recognize a priori the variety being evaluated and possibly influencing their preferences and willingness to pay.

Conclusions

Food choice behavior is complex. Combining disciplines such as sensory science and experimental economics is becoming a popular approach to improve the understanding of food choice behavior. In this study we combine both disciplines to investigate what piece of information exhibits better predictive capacity on the willingness to pay, is it information from preferences measured using hedonic liking scales or information on perceived intensity using intensity scales? We also estimated if extrinsic or intrinsic quality exerts a similar impact on consumer's willingness to pay. Results from this study, show that preference liking has a better explanatory capability for willingness to pay, compared to perceived intensity. The more they liked a sensory quality attribute, they more they are willing to pay for the food product. Whereas not the case regarding perceived intensity, the stronger they perceived an attribute not necessarily they are willing to pay higher, despite the correlation existing between liking scores and intensity scores. Another interesting finding is that willingness to pay was not only driven by variety induced intrinsic sensory quality attributes alone, but that extrinsic cues on the actual variety also influenced willingness to pay. Findings from this study add to the existing body of literature aiming to improve understanding of consumers' food choice behavior.

References

Bartoshuk, L.M. 2000. Comparing sensory experiences across individuals: Recent psychophysical advances illuminate genetic variation in taste perception. Chemical Senses, 25: 447-460.

Bi, X., L. House, Z. Gao, and F. Gmitter. 2011. Sensory evaluation and experimental auctions: measuring willingness to pay for specific sensory attributes. American Journal of Agricultural Economics, 94: 562-568.

Combris, P., P. Bazoche, E. Giraud-Héraud, and S. Issanchou. 2009. Food choices: What do we learn from combining sensory and economic experiments? Food Quality and Preference, 20: 550–557.

Costanigro, M., S. Kroll, D. Thilmany, and M. Bunning. 2014. Is it love for local/organic or hate for conventional? Asymmetric effects of information and taste on label preferences in an experimental auction. Food Quality and Preference, 31: 94–105.

Dinis, I., O. Simoes, and J. Moreira. 2011. Using sensory experiments to determine consumers' WTP for traditional apple varieties. Spanish Journal of Agricultural Research, 9: 351–362.

Gallardo, R. K., E. Kupferman, and A. Colonna. 2011. Willingness-to-pay for Optimal Anjou Pear Quality. HortScience, 46: 452–456.

Greene, W.H. 2008. Econometric Analysis, 6th Ed. Prentice Hall, New Jersey.

Grunert, K.G. 2005. Food quality and safety: Consumer perception and demand. European Review of Agricultural Economics, 32: 369-391.

Jaeger, S.R. 2006. Non-sensory factors in sensory science research. Food Quality and Preference, 17: 132-144.

Kotler, P. and K.L. Keller. 2012. Marketing Management, 14th ed. Upper Saddle River, NJ. Pearson Prentice Hall.

Lange, C., S. Issanchou, P. Combris. 2000. Expected versus experienced quality: Trade-off with price. Food Quality and Preference, 11: 289-297.

Lawless, H.T. and H. Heymann. 1999. Sensory evaluation of food: Principles and practices. Gaithersburg. Aspen.

Lim, J. 2011. Hedonic scaling: A review of methods and theory. Food Quality and Preference, 22: 733-747.

Lund, C.M., S.R. Jaeger, R.L. Amos, P. Brookfield, P., and F.R. Harker. 2006. Tradeoffs between emotional and sensory perceptions of freshness influence the price consumers will pay for apples: Results from an experimental market. Postharvest Biology and Technology: 41: 172–180.

Lusk, J. L. and J.F. Shogren. 2007. Experimental Auctions. Cambridge Journal of Economics.

Mueller, S., P. Osidacz, I.L. Francis, and L. Lockshin. 2010. Combining discrete choice and informed sensory testing in a two-round process: Can it predict wine market share? Food Quality and Preference, 21: 741–754.

Nalley, L.L., D. Hudson, and G. Parkhurst. 2006. Consistency of consumer valuation under different information sets: An experimental auction with sweet potatoes, 37: 56-67.

Navarro da Silva, A., R.C. dos Santos Navarro da Silva, M.A. Marques Ferreira, V.P. Rodrigues Minim, T.M. Texeira da Costa, R. Perez. 2013. Performance of hedonic scales in sensory acceptability of strawberry yogurt. Food Quality and Preference, 30: 9-21.

Peru, Central Reserve Bank. 2015. Information on exchange rates. http://www.bcrp.gob.pe/

Accessed June 2015.

Rabin, M. 1998. Psychology and Economics. Journal of Economic Literature, 36: 11-46.

Simon, H. 1957. A behavioral model of rational choice. In models of man, social and rational: Mathematical essays on rational human behavior on a social setting. New York: Wiley.

Stefani, G., D. Romano, and A. Cavicchi. 2006. Consumer expectations, liking and WTP for specialty foods: Do sensory characteristics tell the whole story? Food Quality and Preference, 17: 53–62.

Tuorila, H. and E. Monteleone. 2009. Sensory food science in the changing society: Opportunities, needs, and challenges. Trends in Food Science and Technology, 20: 54-62.

Zhang, H., R. K. Gallardo, J. McCluskey, and E. Kupferman. 2010. Consumers' Willingness to Pay for Treatment-Induced Quality Attributes in Anjou Pears. Journal of Agricultural and Resource Economics, 35:105–117.

Table 1. Summary statistics of liking and perception of intensity ratings. Correlation between liking and perception of intensity ratings for apples.

Sensory attribute	Average rating using a 1-9 scale 1=extremely dislike/not intense, …, 9=extremely like/intense		Paired t-test Liking vs. intensity	Pearson correlation Liking vs. intensity
	Liking	Intensity		
Appearance/ presence of defects	6.133 (1.828) [1]	3.370 (2.056)	2.763***[2]	-0.330***
Size	6.649 (1.533)	7.429 (0.486)	-0.785***	-0.095
Crispness	6.773 (1.749)	6.520 (1.837)	0.253***	0.725***
Firmness	6.483 (1.846)	5.241 (2.065)	1.239***	-0.146**
Sweetness	5.673 (2.093)	5.203 (2.074)	0.470***	0.769***
Acidity	5.337 (1.974)	3.967 (2.021)	1.370***	0.403***

[1]Standard deviation in parentheses.

[2]* p<0.1 ** p<0.05 *** p<0.01

Table 2. Average for bids for ‘Delicia’, ‘Royal Gala’ and ‘Fuji’ apples. Pairwise comparison of bids across varieties.

‘Delicia’	‘Royal Gala’	‘Fuji’
Mean (St. deviation) in $/kg		
0.851	1.037	0.748
(0.376)	(0.460)	(0.422)
Pairwise comparison		
‘Delicia’ vs. ‘Royal Gala’	-0.186***	
‘Delicia’ vs. ‘Fuji’	0.103***	
‘Royal Gala’ vs. ‘Fuji’	0.289***	

Table 3. Coefficient estimates for the willingness to pay for quality characteristics for 'Delicia', 'Royal Gala' and 'Fuji' apples.

Variable	Full model		Model including sensory variables		Model excluding sensory variables
	Like	Intensity	Like	Intensity	
Intercept	0.170	-0.184	0.141	-0.101	0.741***[2]
	(0.155)[1]	(0.598)	(0.146)	(0.443)	(0.043)
Appearance/ defects	0.017	-0.013	0.014	-0.017	--
	(0.016)	(0.013)	(0.016)	(0.013)	
Size	0.026	0.081	0.034*	0.072	--
	(0.019)	(0.084)	(0.018)	(0.059)	
Crispness	0.037**	0.026*	0.038*	0.021	--
	(0.019)	(0.016)	(0.019)	(0.015)	
Firmness	-0.010	0.008	-0.016	0.009	--
	(0.018)	(0.012)	(0.017)	(0.012)	
Sweetness	0.030*	0.026*	0.048***	0.046***	--
	(0.017)	(0.015)	(0.015)	(0.013)	
Acidity	-0.005	0.018	-0.002	0.017	--
	(0.016)	(0.013)	(0.016)	(0.013)	
Variety "Delicia"	0.096	0.038	--	--	0.110*
	(0.072)	(0.106)			(0.061)
Variety "Royal Gala"	0.184***	0.188**	--	--	0.285***
	(0.070)	(0.075)			(0.061)
Sigma	0.420***	0.424***	0.425***	0.430***	0.434***
	(0.017)	(0.018)	(0.018)	0.018	(0.018)
Number of obs.	296.000	298.000	296.000	298.000	300.000
-2 log likelihood	331.431	338.645	338.197	346.993	354.838
AIC[3]	351.431	358.645	354.197	362.993	362.838
SC[3]	388.368	395.616	383.746	392.569	377.653
McFadden likelihood ratio	0.021		0.025		

[1]Standard errors in parentheses.

[2]* $p<0.1$ ** $p<0.05$ *** $p<0.01$.

[3] AIC is the Akaike information criterion. SC is the Schwarz criterion.

Table 4. Marginal Effects of quality characteristics on the willingness to pay for ‘Delicia’ and ‘Royal Gala’.

Variables	Full model		Paired t-test	Model including sensory variables		Paired t-test
	Like	Intensity	Liking vs. intensity	Like	Intensity	Liking vs. intensity
Appearance/ defects	0.016	-0.013	0.029***	0.014	-0.017	0.031***
Size	0.025	0.079	-0.054***	0.033	0.071	-0.038***
Crispness	0.036	0.025	0.011***	0.037	0.021	0.016***
Firmness	-0.010	0.008	-0.018***	-0.016	0.009	-0.025***
Sweetness	0.030	0.025	0.004***	0.047	0.044	0.003***
Acidity	-0.004	0.018	-0.022***	-0.002	0.017	-0.018***
Variety ‘Delicia’	0.094	0.037	0.057***	--	--	--
Variety ‘Royal Gala’	0.179	0.183	-0.089***	--	--	--

Table 5. Non-parametric comparison between models using like and intensity ratings when explaining variations on bids.

	McFadden likelihood ratio index	Non-parametric Mann Whitney Two sided Pr > |z|
Like vs. intensity coefficients full model	0.021	0.791
Like vs. intensity coefficients restricted model (not including variety variables)	0.025	0.731

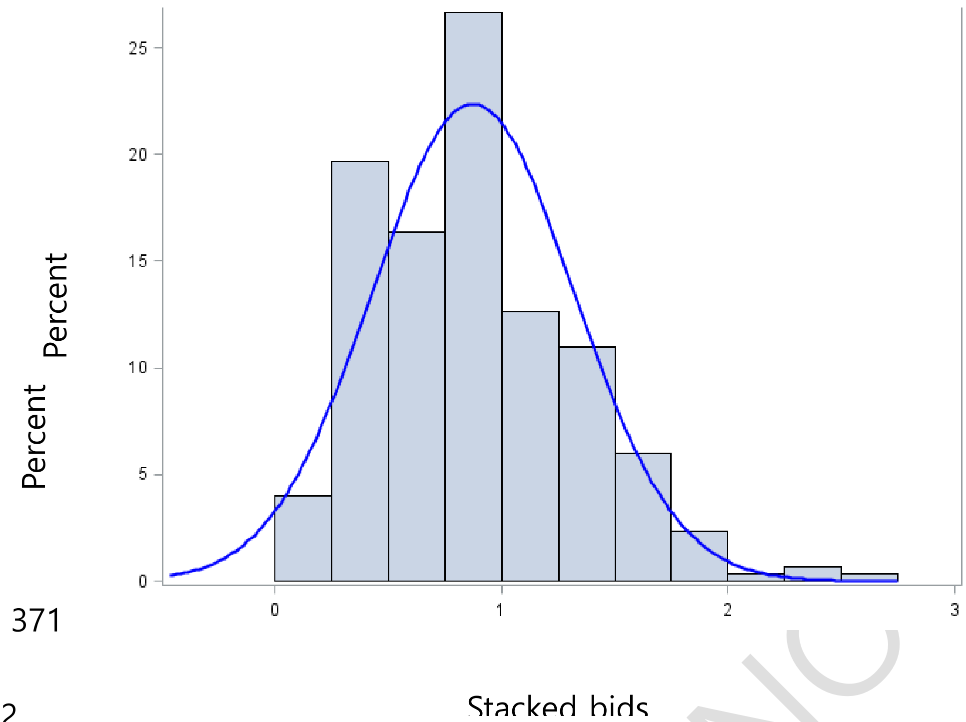

Figure 1. Histogram for stacked bids for all three apple varieties ‘Delicia’, ‘Royal Gala’, and ‘Fuji’.

Resumen

En este estudio investigamos qué información recopilada utilizando herramientas de evaluación sensorial muestra una mejor capacidad predictiva sobre la disposición a pagar, ¿es la información de las preferencias medidas usando escalas de afición hedónicas o información sobre la intensidad percibida usando escalas de intensidad? También estimamos si la calidad extrínseca o intrínseca ejerce un impacto similar en la disposición a pagar del consumidor. Realizamos un estudio de evaluación sensorial y subastas experimentales con tres variedades de manzanas en la que participaron estudiantes de una Universidad en Lima Metropolitana en Perú. Los resultados de este estudio demuestran que la información recopilada sobre la preferencia por un atributo de calidad sensorial tiene una mejor capacidad predictiva para la disposición a pagar, en comparación con la información sobre la percepción de la intensidad percibida del atributo de calidad sensorial. Además, demostramos que tanto atributos intrínsecos de calidad sensorial inducidos por la variedad de manzana y las señales extrínsecas sobre la variedad en si, tienen un impacto en la disposición a pagar. Los resultados se suman a la literatura existente que tiene como objetivo mejorar la comprensión de la conducta de los consumidores al comprar alimentos.

Palabras clave: disponibilidad a pagar, subastas experimentales, comportamiento del consumidor, escalas hedónicas, estímulos de intensidad.