The consistency in the sensory analysis of coffees using Q-graders

Article *in* European Food Research and Technology · March 2017

DOI: 10.1007/s00217-017-2863-9

Eur Food Res Technol
DOI 10.1007/s00217-017-2863-9

ORIGINAL PAPER

The consistency in the sensory analysis of coffees using Q-graders

Lucas Louzada Pereira[1] · Wilton Soares Cardoso[2] · Rogério Carvalho Guarçoni[3] · Aymbiré Francisco Almeida da Fonseca[4] · Taís Rizzo Moreira[5] · Carla Schwengber ten Caten[6]

Received: 30 November 2016 / Revised: 2 February 2017 / Accepted: 19 February 2017

Abstract The process of sensory evaluation is widely used in the coffee classification worldwide. This evaluation is popularly known as "Cupping Test" and constitutes the principal methodology to assess the final quality of the coffee drink. This study observed the effect of the interaction among Q-graders, as well as the coffee tasting process in the morning and in the afternoon. The study was conducted with trained tasters with Q-Grader certificates. The methodology followed the analysis protocol guidelines of the Specialty Coffee Association of America, SCAA, with the participation of two testing groups, each one with two Q-graders for the analysis. *T* Tests, followed by the Pearson's correlation and the analysis of hierarchical grouping were used for the data analysis. The results indicate that the tasters have full capacity of evaluation, although there are variances in relation to the perceptions of the attributes that define the best coffees. The results also indicate that the effect of the shift, when isolated from interaction, interfere neither in the evaluation of the batches, nor the performance of the tasters; however, when associated with interaction (chatting), the analysis indicate distortions, highlighting the need of improving the techniques of sensory evaluation.

Keywords Arabica coffee · Descriptive analysis · Sensory analysis · Statistics, Q-Graders

✉ Lucas Louzada Pereira
lucaslozada@hotmail.com

✉ Carla Schwengber ten Caten
tencaten@producao.ufrgs.br

Wilton Soares Cardoso
wilton.cardoso@ifes.edu.br

Rogério Carvalho Guarçoni
rogerio.guarconi@incaper.es.gov.br

Aymbiré Francisco Almeida da Fonseca
aymbire.fonseca@embrapa.br

Taís Rizzo Moreira
taisr.moreira@hotmail.com

1 Departamento de Engenharia de Produção e Transportes, Universidade Federal do Rio Grande do Sul - UFRGS, Avenida Osvaldo Aranha, 99, Porto Alegre 90035-190, RS, Brazil

2 Departamento de Ciência e Tecnologia de Alimentos, Instituto Federal do Espírito Santo - Ifes, Rua Elizabeth Minete Perim - Bairro São Rafael, Venda Nova do Imigrante 29375-000, Espírito Santo, Brazil

3 Departamento de Estatística, Instituto Capixaba de Assistência Técnica, Pesquisa e Extensão (INCAPER), Rua Afonso Sarlo, 160, Vitória 29052-010, ES, Brazil

4 Pesquisador, Empresa Brasileira de Pesquisa Agropecuária, Embrapa Café, Rua Afonso Sarlo, 160, Vitória 29052-010, ES, Brazil

5 Departamento de Ciências Florestais, Universidade Federal do Espírito Santo (UFES), Avenida Gov. Lindemberg, 316, Jerônimo Monteiro 29550-000, ES, Brazil

6 Departamento de Engenharia de Produção e Transportes, Universidade Federal do Rio Grande do Sul - UFRGS, Avenida Osvaldo Aranha, 99, Porto Alegre 90035-190, RS, Brazil

Published online: 09 March 2017

Introduction

Coffee is a beverage greatly appreciated in many countries around the world due to the fact that it is a natural product, with varied aromas and pleasant tastes. Among the most commonly grown species is the Arabica coffee (*Coffea arábica* L.), which presents highly valued sensory characteristics, providing a beverage with high commercial value [7].

Within the coffee industry, the tasting procedures are used for negotiating the commodity with attributes such as the quality of the beverage, which is described by the tasters, using their personal opinion and the experience in tasting achieved throughout the years [6]. Although the tasting process is widely used, for DiDonfrancesco et al. [3], it does not constitute the best method for the evaluation of the quality of the coffee, due to a variety of factors that interfere in the tasting process.

The taste and aroma nuances are what make the task of sensory evaluation and analysis of coffee very complex, as during the roasting process they may surpass the quantity of more than 800 aromatic compounds [1, 2; Bhumiratana et al. 2]. For Dzung [5], one of the main problems in the use of an expert in sensory evaluation is that the qualification of “expert tasters” (Q-graders) is not clearly defined. According to ISO 856-2 (1994), the experience is not the main criterion of a specialist, one must also be trained and have a high sensory sensitivity.

According to Alvarado and Linnemann [1], the “taster” is a judge who performs a sensory evaluation, and responsible for evaluating the quality of the coffee. Consequently, this agent influences the evaluation according their perceptions, leaving, on many occasions, the determination of the final quality of the coffee in their hands, thus, interfering in the price of the product.

Some methodologies, such as the test protocols of the Specialty Coffee Association of America (SCAA) and the Brazilian Specialty Coffee Association (BSCA), determine procedures for the sensory evaluation of specialty coffees, and these are the ones commonly adopted in Brazil and in the world. This classification is based on the coffee cupping test, in other words, on taste, by means of tasters, being, on many occasions, variable from one region to another and does not characterize the real physical–chemical constituents that determine the best quality of the beverage [10].

Authors, such as Feria-Morales [6], Roos [12] and Bhumiratana et al. [2], emphasize the necessity of using physical–chemical analyses in the classification of this process as a way to reduce the impact of the sensory subjectivity. However, in the coffee production chain, the use of sensory analysis with Q-Graders has been adopted during the last years for quality determination. The sensory protocols recommend a calm environment during the analysis, but it is of tacit knowledge that Q-Graders comment on their results and judgments during the tasting process, considering the taste complexity of different nuances of the coffees. Nevertheless, the sensory analyses have been a very important tool in the characterization of the different attributes of beverages within the coffee chain [Oliveira et al. 13].

This way, the interference of factors such as shift (time) and interaction (chatting) are not routinely discussed in the literature, as factors influencing the quality of the sensory process itself, since the empirical perception during the Q-Graders work has indicated relative wear between shifts and environments with noise (chat between Q-Graders).

With the aim of providing further information about the quality of this tasting process and about the tasters’ performance, this research was carried out in order to evaluate coffees, tasters, and their judgments considering the effect of interaction and shift could have on the use of *SCAA* protocol to the specialty coffee tasting in a quality contest.

Materials and methods

Samples

For the completion of this study, we used 20 samples of specialty coffee, selected by a jury of four (04) tasters, all Q-Grader’s. The 20 coffee samples were selected to compose the final round of the quality contest in Venda Nova do Imigrante, a town located in the Southwestern mountainous micro-region of Espírito Santo State. The tests were carried out during 2015 and 2016.

The 20 samples were prepared in the coffee sensory analysis laboratory of the Federal Institute of Espírito Santo in Venda Nova do Imigrante, respecting the methodology of the Specialty Coffee Association of America (SCAA). The roasts were performed using the roaster Laborato TGP2 and monitored with Agtron-SCAA discs set. The roasting point of these samples was located between the colors determined by the discs #65 and #55 SCAA [11].

The roasting was carried out 24 h in advance, and the grinding respected the 8 (eight) resting hours following the roasting. All the samples were roasted between 9 and 11 min, and after the roasting and the cooling, they remained sealed, in accordance with the sensory analysis methodology established by the *SCAA*. The *SCAA* protocol [11] was used for the coffee evaluation.

Sample preparation for tasting

The 20 coffee samples were ground with a Bunn G3 electric grinder, with a medium granulometry (usually applied in cuppings in Brazil). Each lot was tasted with 5 cups, and the optimum concentration of 8.25 grams of ground coffee in 150 ml of water was adopted, in accordance with the midpoint of the optimal equilibrium graph to obtain the Golden Cup [11]. The point of water infusion occurred after the water reached the temperature of 92–95 °C. The tasters began their evaluations when the temperature of the cups reached 55 °C, respecting the 4-min time for the tasting after the infusion [11].

Evaluation method of the samples

The quality of a coffee lot, once evaluated through the SCAA [11] method, is expressed by a centesimal numeric scale. It is expected that coffees which obtained high scores should be clearly better than coffees which received lower notes, thereby demonstrating the consistency of the method of sensory analysis. The tasting form provides the possibility of evaluating eleven (11) important attributes to the coffee: Fragrance/Aroma, Uniformity, Absence of defects (clean cup), Sweetness, Flavor, Acidity, Body, Aftertaste, Balance, Defects, and Overall Evaluation. Highly positive results arise from the perception of a balanced group formed by the evaluated attributes.

The defects of the beverage result in unexpressive results due to unpleasant interferences in the flavor. The overall evaluation is based on the sensory memory that a taster has, always referring to coffees from the same origin and nature. The results of this sensory evaluation are established based on a scale of sixteen (16) units that represent the levels of quality with intervals of 0.25 (one-fourth of a point) among the numeric values set between “6” and “9”. For those coffees considered good, the range was between 6.00 and 6.75; very good, 7.00 to 7.75; excellent, 8.00 and 8.75; and exceptional, between 9.00 and 9.75 points.

Groups of sensory evaluation

Two groups of sensory evaluation were structured, and each group was composed of two judges (tasters), totaling four (04) coffee tasters.

The sensory evaluation was performed in two different shifts, morning and afternoon. In the morning, 10 coffee samples were placed on table one (01) and coded for two Q-graders, repeating the same procedure on table two (02). Morning analyses were conducted without any communication between Q-graders. In other words, the judges could not chat at all before, during and after the process of sensory evaluation. They started at 9:00 am and were finished at 10:15 am.

In the afternoon, the Q-graders returned to the laboratory to perform the same tests, and the same samples were kept for the respective Q-graders; however, the samples were randomized so that none of them knew the origin of the coffees. The positions of the tables were also reversed, to avoid familiarity of the Q-graders with the environment. The sensory evaluation began at 1:30 pm and was concluded at 2:30 p.m, and the same chat control employed in the morning was employed in the afternoon shift.

Sensory evaluation group on second day

On the second day, the order was reversed; the tasters in group 01 went to group 02, so that all 20 samples were tasted by all the judges in 2 days of the experiment.

Besides the exchange of jury, the tasters began the tasting at opposite ends of the tables; when one taster began the test from left to right, the other taster carried out the test from right to left, so that none of the samples was tasted last. The movement was controlled inside the test bench only on the first day.

At this day, the Q-graders could chat with each other, during and after the sensory analysis, the latest being a practice commonly used in sensory analysis processes with tasters around the world. Communication between them as pairs was not curtailed. In other words, the Q-graders were free to engage in conversation among their peers about the quality and sensory attributes of coffees in the morning and afternoon.

Experimental design

The two experiments were carried out under entirely randomized outlines with 20 repetitions (coffees), being the first with the purpose of evaluating the influence of the tasters’ interaction (chatting) in the quality of the descriptive sensory analysis. The second purpose was to observe the shift (time), in order to evaluate a best possible time for the tasters to conduct sensory analysis.

The averages of the characteristics were compared using the paired *t* test (before and after) at 5% probability, followed by Pearson correlation analyses (with the objective to verify the consistency of the correlations in the environment with and without chat in the two days of experiment), being the significance tested by the *t* test.

For the similarity evaluation among the coffees, a matrix was created with the averages of the variables and then a dendrogram was built using the Average Euclidean distance in order to measure the distance between two points and the Full Connectivity Hierarchical Grouping method. The purpose is to maximize the homogeneity of objects within groups, at the same time that the heterogeneity among the groups is maximized [8]. For the statistical analyses, the SPSS 19 was used.

Results and discussions

With the scores of all the attributes from the two groups of tasters, we obtained different statistics, being possible to compare the assumptions initially raised, prior to the experiments.

Evaluation of the effect of shift on the tasters

By means of the *t* test, it comes to the obtained results about the judgments of the 20 coffee samples in Table 1, of the analysis panel carried out in the morning and in the afternoon of the first day. This analysis allows evaluating the effect of shift (time) and interaction (environment without chat) on the tasters' scores on the different attributes given that each taster tasted approximately 50 cups of coffee (10 samples).

For the levels of the attributes evaluated on Table 1, it can be observed that no significant differences were perceived among the averages for all the characteristics evaluated, in other words, shift (time) did not interfere in the sensory analyses during the first testing block. The tasters got to repeat the same scores for all the attributes.

The data confirm that the process of sensory evaluation, when applied with well-trained and capable tasters, with the concentration and proper attention is precise and efficient in the coffee sensory evaluation process.

In order to study the effect of shift (time) in the relation among the attributes, the Pearson's correlation was carried out, since in a sensory perception of this type of tastings, each attribute directly affects the other, within the protocol. Tables 2 and 3 illustrate the Pearson's correlations among the attributes evaluated by the tasters in the 20 coffee samples.

Table 1 Averages and standard deviations for the characteristics: fragrance/aroma, uniformity, absence of defects, sweetness, flavor, acidity, body, aftertaste, balance, overall and total evaluation, before (morning) and after (afternoon)

Characteristics	Before (morning)		After (afternoon)	
	Average[a]	Standard deviation[b]	Average[a]	Standard deviation[b]
Fragrance/Aroma	7.5125 a	0.6952	7.8375 a	0.7663
Uniformity	9.7500 a	0.9104	9.7000 a	0.9787
Absence of defects	9.8000 a	0.8944	9.7000 a	0.9787
Sweetness	9.8000 a	0.8944	9.7000 a	0.9787
Flavor	7.7562 a	0.7018	7.4375 a	0.8316
Acidity	7.6500 a	0.7193	7.3875 a	0.7640
Body	7.5687 a	0.6804	7.5375 a	0.7653
Aftertaste	7.5062 a	0.6852	7.5938 a	0.6999
Balance	7.3312 a	0.5866	7.6750 a	0.7145
Overall	7.5125 a	0.7081	7.6250 a	0.7333
Total	82.1875 a	6.3541	82.1937 a	7.2955

[a]The averages of the attributes measured before and after followed by a letter in the line do not differ by the t test at 5% probability

[b]The standard deviation before and after

In the analysis of Tables 2 and 3, from the Pearson's correlations, uniformity of the relations among attributes with positive and significant correlations can be perceived. As the composition rule of the contest uses the additive model, in which values for each attribute are added in order to obtain the total value for a combination of attributes, positive correlations must be expected to achieve the maximum value [Hair et al. 8].

Several attributes showed high correlation values, especially aftertaste, balance and sweetness. In Dzung's work [5], which evaluated the correlation among the attributes of a coffee sensory analysis; those of greatest correlation were flavor, balance and aftertaste. The results of this study confirmed only the high correlation for the balance variable; Alvarado and Linnemann [1] also observed inconsistencies in the processes of sensory analysis in relation to attributes, indicating that the sensory analysis process has a certain subjectivity.

In order to represent the samples, by groups of quality, given the range of specialty coffees, dendrograms were used at the end of each shift.

The dendrogram in Fig. 1, representing the analysis carried out in the morning, suggests the existence of two homogeneous groups: group A, formed by coffee 6, and group B, formed by the other coffees. Group B can still be divided in 2 sub-groups, obtaining, this way, sub-group B1, formed by 7, 18, 11, and 13, the best coffees in terms of scores, and the other coffees compose sub-group B2.

In this dendrogram of Fig. 1, it is important to highlight that coffee no. 6 was differentiated by all tasters, forming an isolated cluster, as being the one presenting defects during all the coffee cupping, hence indicating the possibility of phenolized beverage (described by the notes on the testing protocol) during the sensory evaluation process; the beverage classified as "riada" has the characteristic of slight iodoform or phenolic acid taste [14]. The Q-graders found the coffee defective and discounted the points according to the protocol of the SCAA. These results corroborate those of Alvarado and Linnemann [1].

In the dendrogram of Fig. 2, on the afternoon period, from the first day of testing, the existence of two homogeneous groups can also be observed: group A—formed by coffee 9, and group B—composed by the other coffees. The group can be divided in two sub-groups. In sub-group B1, there are the samples 6, 7, 11, 13, and 18, the others compose sub-group B2.

Although the *t* test at 5% probability had indicated that the effect of shift had not interfered in the tasting process, the previous dendrogram indicates a disparity among

Table 2 Pearson's correlation coefficients among fragrance, uniformity, absence of defects, sweetness, flavor, acidity, body, aftertaste, balance, overall and total evaluation, of the coffees with sensory analyses carried out in the morning

Variables	Uniformity	Absence of defects	Sweetness	Flavor	Acidity	Body	Aftertaste	Balance	Overall	Total
Fragrance	0.5456**	0.5120*	0.5120*	0.8375**	0.8199**	0.7612**	0.8318**	0.8320**	0.8382**	0.8585**
Uniformity		0.9695**	0.9695**	0.5791**	0.4822*	0.5177**	0.5088*	0.5327**	0.4949*	0.8091**
Absence of defects			0.9999**	0.5890**	0.5399**	0.5427**	0.5174**	0.5341**	0.5027*	0.8219**
Sweetness				0.5890**	0.5399**	0.5427**	0.5174**	0.5341**	0.5027*	0.8219**
Flavor					0.8467**	0.8739**	0.7233**	0.8355**	0.8139**	0.8862**
Acidity						0.8381**	0.7838**	0.7140**	0.7856**	0.8453**
Body							0.8633**	0.8382**	0.8839**	0.8777**
Aftertaste								0.8701**	0.9574**	0.8655**
Balance									0.9119**	0.8684**
Overall										0.8765**

*e

**Significant at 5 and 1% probability respectively

Table 3 Pearson's correlation coefficients among fragrance/aroma, uniformity, absence of defects, sweetness, flavor, acidity, body, aftertaste, balance, overall and total evaluation, of the coffees with sensory analyses carried out in the afternoon

Variables	Uniformity	Absence of defects	Sweetness	Flavor	Acidity	Body	Aftertaste	Balance	Overall	Total
Fragance/Aroma	0.7034**	0.7034**	0.7034**	0.8400**	0.8632**	0.9111**	0.9007**	0.9017**	0.8662**	0.9317**
Uniformity		0.9999**	0.9999**	0.5254**	0.5860**	0.5779**	0.6194**	0.6811**	0.6416**	0.8489**
Absence of defects			0.9999**	0.5254**	0.5860**	0.5779**	0.6194**	0.6811**	0.6416**	0.8489**
Sweetness				0.5254**	0.5860**	0.5779**	0.6194**	0.6811**	0.6416**	0.8489**
Flavor					0.9215**	0.9355**	0.9134**	0.8899**	0.8953**	0.8731**
Acidity						0.9246**	0.8558**	0.8484**	0.8395**	0.8829**
Body							0.9158**	0.9318**	0.9143**	0.9077**
Aftertaste								0.9702**	0.9420**	0.9194**
Balance									0.9652**	0.9449**
Overall										0.9205**

*e

**Significant at 5 and 1% probability respectively

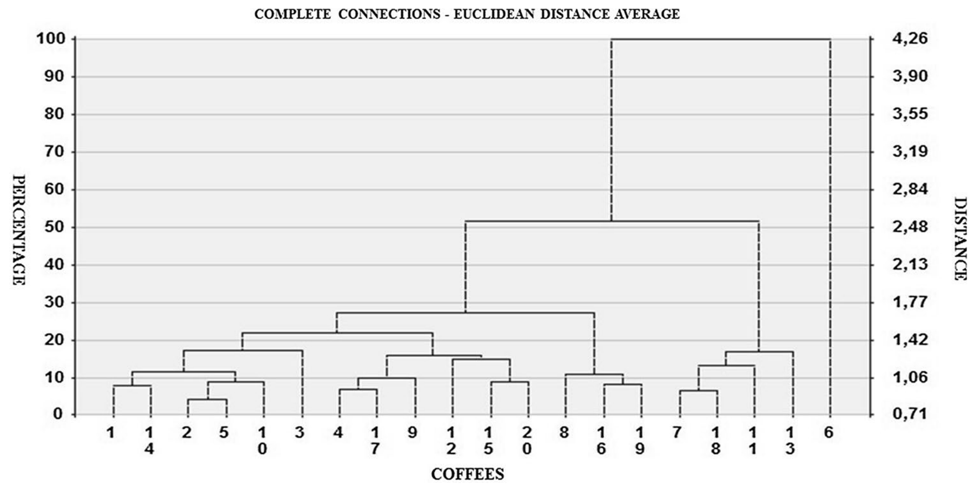

Fig. 1 Dendrogram among coffees obtained based on the sensory analyses carried out in the morning

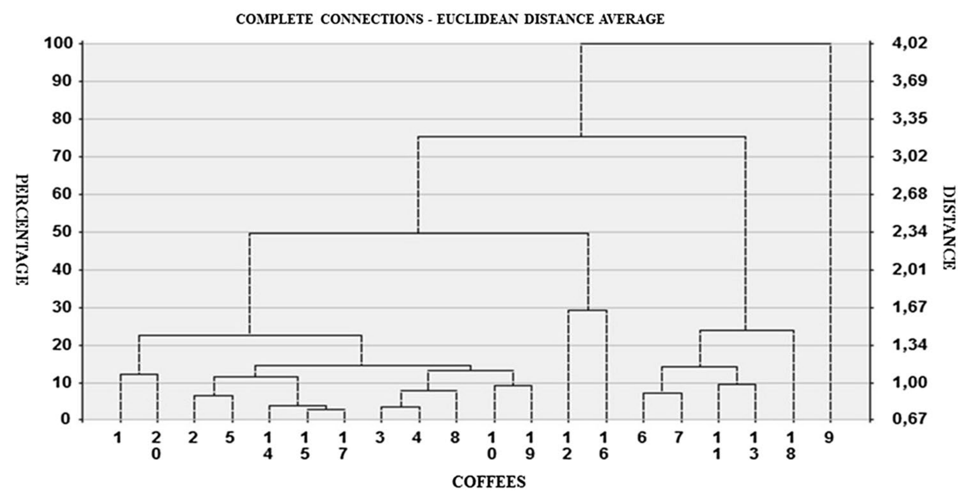

Fig. 2 Dendrogram among coffees obtained based on the sensory analyses carried out in the afternoon

the tasters. In the morning, coffee number 6 evidenced defect (dirty or phenolic cup), however, in the afternoon, this condition was evidenced in coffee sample 9, isolated from the other groups, indicating the same perception of evaluation in no. 6 sample in the morning.

In addition, the Q-graders did not score coffee number 9 with the presence of a defect. At this point, when the subgroups B1 of the morning dendrogram (Fig. 1) and the dendrogram of the afternoon (Fig. 2) were evaluated, homogeneous subgroups can be perceived only with the difference of coffee no 6, indicating inconsistency in the judgments only in this sample of coffee. However, it is not possible to state that the error is only in the evaluation of the Q-graders, the complexity in the formation of aroma compounds and coffee flavor can often disrupt the judgment given the chemical interactions that occur with the roasted samples.

Table 4 Variables that exercised the most the relative influence in the construction of the dendrogram on the first day of testing

Influence on the variables between morning and afternoon (first day of evidence)			
Morning	Percentage	Afternoon	Percentage
Uniformity	23.7	Uniformity	36.3
Acidity	22.6	Acidity	30.0
Fragrance/aroma	9.47	Fragrance/aroma	10.5

Concerning the relative contribution of the variables for the construction of the dendrograms of Figs. 1 and 2, there is Table 4.

Again, there is consistency in the judgments, in relation to the attributes, where the same variables that most influence the notes of the coffees are repeated by the judges in the morning and in the afternoon. These results corroborate those of Alavardo and Linnemann [1], Civille and Oftedal [4], DiDonfrancesco et al. [3], regarding the consistency and use of judges in sensory analysis.

Evaluation of the interactive environment among tasters

During the second day, the second analyses jury was carried out. Two coffee testing moments were set, following the same principle of the first day. However, on the second day, it was determined that the tasters were free to chat during the coffee tasting and consequently the movement order for the test of the lots was not controlled, with the number of tasters per jury as the only control (two tasters in each panel, respectively).

In the results of the *t* test about the average of the scores of the attributes presented in Table 5, it can be observed that there are significant differences among the averages for the fragrance/aroma and body characteristics before and after. In other words, the fact of existing interaction among the tasters, with chatting, comments and noise interfered on the sensory analysis.

The distortions in the process of sensory analysis corroborate with the results of Feria-Morales [6], indicating that the environment with chat is disturbing in the process of sensory analysis. This point deserves more attention on the part of scientific investigations, since the process of sensory analysis of coffee ends up being influenced by these interactions in the environment.

In the relation study among the attributes, through the correlations among the variables, it obtained on the second day, the Tables 6 and 7, with a result of the Pearson's correlations among the attributes evaluated by the tasters in the 20 coffee samples.

Analyzing Table 6 in an isolated manner, we find diverse non-significant correlations; this implies that some characteristics do not have effect on the final coffee score. For example, consider the uniformity characteristic that did not influence any other variable and at the same time is not able to alter the final score. Since there is an additive characteristic for the final score, and in comparison, with the significant correlations of Table 2, the judgments of the second day, in an interactive environment, made the sensory analysis flawed.

This non-significant correlations were in part generated by the absence of defects, sweetness, and flavor

Table 5 Averages and standard deviations of the fragrance/aroma, uniformity, absence of defects, sweetness, flavor, acidity, body, aftertaste, balance, overall and overall evaluation, before and after

Characteristics	Before (morning)		After (afternoon)	
	Average[a]	Standard deviation[b]	Average[a]	Standard deviation[b]
Fragrance/aroma	7.9125 a	0.5635	7.6625 b	0.6702
Uniformity	10.0000 a	0.0000	9.6075 a	1.0346
Absence of defects	10.0000 a	0.0000	9.7000 a	0.9787
Sweetness	10.0000 a	0.0000	9.8000 a	0.8944
Flavor	7.7812 a	0.4796	7.6188 a	0.7134
Acidity	7.6125 a	0.5332	7.4000 a	0.6621
Body	7.7438 a	0.4204	7.4500 b	0.5769
Aftertaste	7.7250 a	0.5465	7.5812 a	0.6154
Balance	7.6813 a	0.4614	7.6188 a	0.6864
Overall	7.7250 a	0.5465	7.5687 a	0.7041
Total	84.1813 a	3.1624	82.0075 a	6.1315

[a]The averages of the measured characters before and after, followed by the same letter in the line, do not differ by the t test at 5% probability

[b]The standard deviation before and after

Table 6 Pearson's correlation coefficients among fragrance/aroma, uniformity, absence of defects, sweetness, flavor, acidity, body, aftertaste, balance, overall and overall evaluation of the coffees with sensory analyses carried out in the morning

Variables	Uniformity	Absence of defects	Sweetness	Flavor	Acidity	Body	Aftertaste	Balance	Overall	Total
Fragrance/aroma	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.8473**	0.8445**	0.7334**	0.8469**	0.8547**	0.7855**	0.9535**
Uniformity		0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]
Absence of defects			0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]
Sweetness				0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]	0.0000[ns]
Flavor					0.9276**	0.6820**	0.7843**	0.6530**	0.5898**	0.8825**
Acidity						0.7442**	0.7861**	0.6313**	0.5604**	0.8836**
Body							0.6900**	0.7479**	0.6256**	0.8291**
Aftertaste								0.7787**	0.8265**	0.9235**
Balance									0.8895**	0.8915**
Overall										0.8526**

NS not significant

*e

**Significant at 5 and 1% probability, respectively

Table 7 Pearson's correlation coefficients among fragrance/aroma, uniformity, absence of defects, sweetness, flavor, acidity, body, aftertaste, balance, overall and overall evaluation of the coffees with sensory analyses carried out in the afternoon

Variables	Uniformity	Absence of defects	Sweetness	Flavor	Acidity	Body	Aftertaste	Balance	Overall	Total
Fragrance/aroma	0.2628^{ns}	0.4393*	0.4961*	0.8639**	0.7909**	0.8792**	0.8476**	0.7960**	0.8150**	0.8326**
Uniformity		0.9171**	0.8206**	0.1894^{ns}	0.2172^{ns}	0.3280^{ns}	0.3544^{ns}	0.4850*	0.4471*	0.6811**
Absence of defects			0.8898**	0.3929*	0.3573^{ns}	0.4847*	0.5232**	0.6042**	0.6043**	0.8116**
Sweetness				0.4516*	0.4088*	0.4895*	0.5092*	0.5551**	0.5244**	0.7968**
Flavor					0.9369**	0.8503**	0.7823**	0.6817**	0.7179**	0.7898**
Acidity						0.8860**	0.7737**	0.6788**	0.6998**	0.7741**
Body							0.8921**	0.8692**	0.8651**	0.8751**
Aftertaste								0.9240**	0.9011**	0.8761**
Balance									0.9571**	0.8953**
Overall										0.8904**

NS not significant

*e

**Significant at 5 and 1% probability respectively

characteristics, which received scores of 10 from tasters in all coffees. It is possible to suppose that the change of the environment, with proportionate interactions by comments and small talks lessened the quality of the judgments on the levels of the attributes due to lack of concentration.

Table 7 illustrates the correlations in the afternoon, and there is the presence of non-significant correlations, in other words, the environment influences the tasters. The studies of Feria-Morales [6], Roos [12], Bhumiratana et al. [2], have indicated the need to use other techniques besides sensory analysis to reduce subjectivity in the process of evaluating coffee quality through Tasters.

In relation to the classification of the 20 coffees evaluated on the second day, under the interactive environment (chatting). This results in dendrogram A of Fig. 3 for the morning and dendrogram B for the afternoon.

The first dendrogram (A) suggests the existence of two homogeneous groups: group A—formed by coffees 3, 4, 5, 12, 17, and 20, and the rest in group B. In the sub-groups of the groups, it is found sub-group B1, composed by 7, 13, 11, and 18, and sub-group B2, formed by the other coffees.

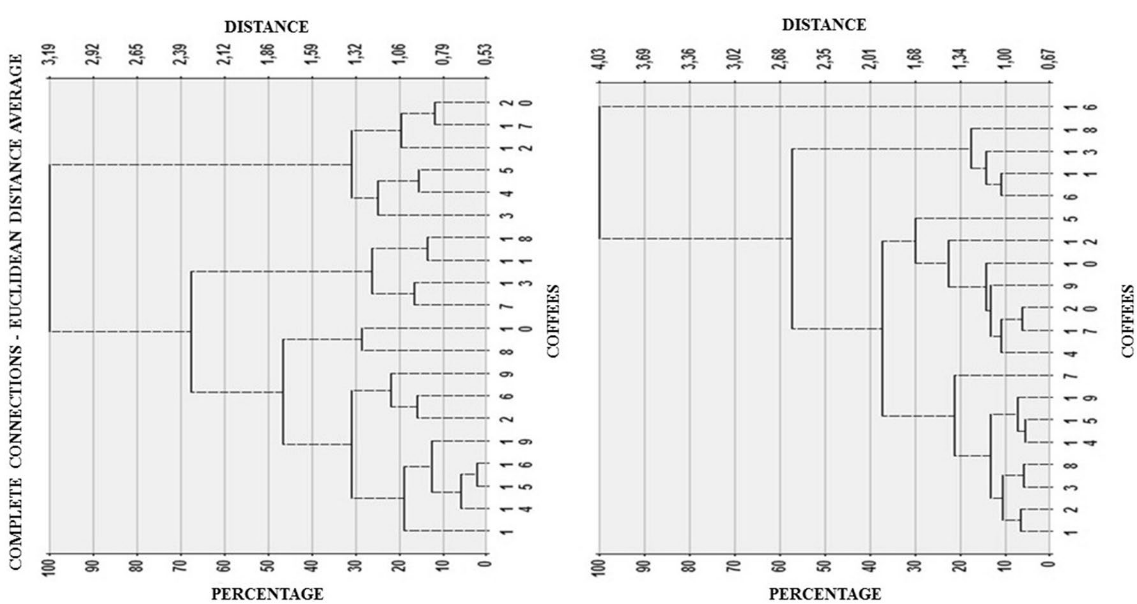

Fig. 3 Classification of the coffees, obtained from the sensory analyses carried out during the second day. Dendrogram A results of the analyses carried out in the morning. Dendrogram B results of the analyses carried out in the afternoon

For the afternoon period in the interactive environment (chatting), there is dendrogram B of Fig. 3, which suggests the existence of two homogeneous groups: group A—formed by coffee 16, and group B—composed by all the others coffees. Subdividing group B, we find sub-group B1, formed by 6, 11, 13, and 18, and sub-group B2, composed by the other coffees.

Based on this dendrogram, it is necessary to emphasize that group A—formed by coffees 7, 13, 11, and 18—has the best scores, all of these coffees with notes above 90 points, considered outstanding by the SCAA protocol. Group B—formed by coffees 3, 4, 5, 12, 17, and 20—contained the coffees with the worst results on general average, ranging from 75 to 79.75 points, considered Very Good by the SCAA protocol, and group C by the other coffees, leveled above Group B, but below group A, being framed as Excellent, cafes between 80 and 87 points.

When analyzing the relative contribution of the variables in the construction of the dendrograms of Fig. 3, the summary of the variables is presented on Table 8. On the second day, the variables were not the same as those on the first day; therefore, the resulting information is based on the following data:

The variables Balance, Overall and Body, from Table 8 (those that compose dendogram A), totally differ from the variables presented on Table 4, being one more issue to be considered, to evaluate how the analyses of the second day were highly influenced by the interactive environment proposed in the experiment.

SCAA protocol and the multivaried analyses

Based on the SCAA protocol, the best coffees would be determined by the average of the sums of the scores of the 10 attributes. Hence, coffee number 18 with 92.62 points of average, would be considered the best coffee among the 20 evaluated ones. However, this methodology does not evaluate the shift or the environment in which the sensory analyses are carried out.

Furthermore, in this methodology, there is no way to evaluate the correlation among the attributes, given that non-significant correlations may indicate altered scores, not only due to the environment. In addition, the results may generate the hypothesis that the non-significant correlations may arise from a taster's unskilled evaluation [9], a relevant fact in the measurement of consistency of Q-graders in sensory panels.

Thus, performing the analysis of the best coffee, among the 20 finalists through multivariate analysis, with application of clusters in dendrograms, it was possible to obtain homogeneity in three samples 11, 13, and 18 (group with the highest grades, 92.62; 90.25, 89.37, respectively), which recur in all clusters, finding the best coffees in attributes evaluated by the judges.

By means of this data, it is necessary to discuss whether a simple average of the sum of attributes would represent a champion coffee or a better quality one. Likewise, it must be considered that deeper and multivariate analyses (of statistic techniques), when evaluating aspects linked to the quality of the judgments and of the tasters, evidence a more consistent direction, even if not only one coffee is found but rather a group of superior ones.

Moreover, it is necessary to explore other tools that may help in the decision making in such a way that it is possible to draw more precise conclusions on the aspects relative to the coffee sensory analysis process.

The use of physical–chemical techniques could reverse what was seen in this study, in relation to coffee sample number 6, which, in dendrogram Fig. 1, appears in a separate group for presenting a score with defect (phenolized cup). This fact was not repeated in the other tests; coffee number 6 appears twice in the group with high scores, observed in Fig. 2 and dendrogram B and Fig. 3.

Everything indicates that coffee number 6 was erroneously disqualified and, in order to avoid this type of mistakes, at least to disqualify a coffee because of a defect (which directly impairs the producer). Physical- chemical analyses should be carried out for the verification and confirmation of defects even before the sensory or parallel analyses, with the objective of greater transparency to the method of sensory analysis.

Developing sensory classification systems is not a trivial task and must be done individually for each product, taking into consideration its peculiarities, just as the sensory evaluation must be done by trained and experienced tasters, in order to identify and understand the quality of a particular product Feira-Morales [6].

Table 8 Variables that exercised the most the relative influence on the constructions of the dendrogram in the second day of testing

Relative influence of the variables between morning and afternoon (second day of evidence)			
Morning	Percentage	Afternoon	Percentage
Balance	24.2	Uniformity	24.2
Overall	21.6	Fragrance/aroma	17.9
Body	21.1	Balance	13.7

Conclusion

Q-graders have full judgment capacity and, even though there may be distortions in the coffee sensory evaluation process, it can be concluded that there is homogeneity in relation to the coffee cupping.

The shift (time) on the first day of testing did not pursue a significant effect on the evaluation of the coffees, but during the second day, the interaction (chatting) among the tasters did interfere in the sensory evaluation process, indicating the act of chatting should be avoided during the process.

The protocols used do not take into consideration more robust statistical analyses, which may generate more reliable results so that it is possible to avoid coffees from being eliminated erroneously from contests, consequently, impairing the coffee growers.

The variables/attributes of the sensory analysis protocol demonstrate that uniformity, acidity, body, fragrance/aroma, balance and overall were the factors that contributed the most for the decision making in the choice of the best coffees.

Acknowledgements The authors thank the Federal Institute of Espírito Santo for supporting this research and also the translation and review of this article, as well as the Q-graders' participation, who dedicated themselves to the realization of this study. We also thank CNPq and SETEC for the availability of resources for research.

Compliance with ethical standards

All ethical standards regulated in Brazil were duly followed. This article comes from a research project, being in compliance with all ethical principles.

Conflict of interest None.

References

1. Alvarado RA, Linnemann AR (2010) The predictive value of a small consumer panel for coffee-cupper Judgment. Br Food J 112 No. 9:1023–1032
2. Bhumiratana N, Adhikari K, Chambers E (2011) Evolution of sensory aroma attributes from coffee beans to brewed coffee. LWT Food Sci Technol 44:2185e2192
3. Di Donfrancesco BD, Guzman NG, Chambers E Comparison of results from cupping and descriptive sensory analysis of Colombian brewed coffee. The Sensory Analysis Center, Kansas State University, Manhattan, KS 66502. Facultad de Ingeniería, Universidad Surcolombiana, Neiva, Huila, Colombia. Journal of Sensory Studies ISSN 0887–8250
4. Civille GV, Oftedal KN (2012) Sensory evaluation techniques—Make "good for you" taste "good". Physiol Behav 107 (2012) 598–605.
5. Dzung NH, Dzuan L The role of sensory evaluation in food quality control, food research and development: a case of coffee study. Disponível em: http://www4.hcmut.edu.vn/~dzung/Sensoryrole.pdf.
6. Feira-Morales AM (2002) Examining the case of green coffee to illustrate the limitations of grading systems/expert tasters in sensory evaluation for quality control. Food Qual Preference 13:355–367
7. Fernandes SM, Pereira RGFA, Pinto NAVD, Nery MC, Pádua FRM (2003) Constituintes químicos e teor de extrato aquoso de cafés arábica (Coffea Arábica l.) E CONILON (Coffea Canephora pierre) torrados. Ciênc Agrotec Lavras 27(5):1076–1081 (set./out)
8. Hair JF, Black WC, Badin BJ, Anderson RE, Tathan R (2009) Análise multivariada de dados. 6ª ed., Bookman, Porto Alegre
9. Hayakawa F, Kazami Y, Wakayama H, Oboshi R, Tanaka H, Maeda G, Hoshino C, Iwawaki H, Miyabayashi T (2010) Sensory lexicon of brewed coffee for japanese consumers, untrained coffee professionals and trained coffee tasters. J Sens Stud 25:917–939.
10. Molin RND, Reis AR, Junior EF, Braga GC, Scholz MBS (2008) Caracterização física e sensorial do café produzido nas condições topoclimáticas de Jesuitas, Paraná. Maringá 30(3): 353–358
11. SCAA. Specialty Coffee Association of American. Protocols (2013) Disponível em: http://www.scaa.org/PDF/resources/cupping-protocols.pdf.
12 Roos CF (2009) Sensory science at the humane machine interface. Trends Food Sci Technol 20:63e72
13 Oliveira PD, Borém FM, Isquierdo EP, Giomo G S, Lima RR, Cardoso RA (2013) Aspectos fisiológicos de grãos de café, processados e secados de diferentes métodos, associados à qualidade sensorial. Coffee Sci Lavras 8(2):211–220
14 Pinto NAVD, Fernandes SM, Pires TC, Pereira RGFA, Carvalho VD (2001) Avaliação dos polifenóis e açúcares em padrões de bebida do café torrado tipo expresso. Rev Bras de AGROCIÊNCIA 7(3):193–195 (set-dez)