Concurrent vs. retrospective temporal data collection: Attack-evolution-finish as a simplification of Temporal Dominance of Sensations?

Michel Visalli, Benjamin Mahieu, Arnaud Thomas, Pascal Schlich

Title

Concurrent vs. retrospective temporal data collection: Attack-Evolution-Finish as a simplification of Temporal Dominance of Sensations?

Auteurs

Visalli[1], M., Mahieu[1], B., Thomas[2], A., Schlich[1], P.

[1]Centre des Sciences du Goût et de l'Alimentation, CNRS, INRAE, Univ. Bourgogne Franche-Comté, F-21000 Dijon, France.
[2]SensoStat, Dijon, France

Keywords

Consumer study. Temporal data. Consumer behaviour. Method comparison.

Abstract

As tasting is a dynamic process, temporal data are collected simultaneously with tasting. Indeed, most newly reported studies involving consumers have been conducted using the Temporal Dominance of Sensations (TDS) or Temporal Check That Apply methods. Concurrent data collection reduces potential bias such as memory, but it attaches great importance to the moment at which the sensations are cited. Thus, it results in measurement of durations possibly made imprecise due to heterogeneity in consumers' behaviour, which could affect conclusions. A new retrospective method inspired from Temporal Order of Sensations, Pick 3 and Rank and the 3 phases of wine evaluation is introduced in this article. Based on a concept close to dominance, the Attack-Evolution-Finish (AEF) method allows consumers to select one descriptor each at the beginning, at the middle and at the end of the tasting. The results obtained with two different panels carrying out both the AEF and TDS tasks on dark chocolates are compared. The conclusions are very similar in terms of product discrimination. The retrospective task removes the consumers' heterogeneity in terms of the number of citations, delays and durations and thus requires no data transformation before analyses. In view of these results, rather than promoting a new method, the article raises questions about the level of detail to look at in temporal sensory data.

1. Introduction

Perception is a time-dependent process that has been investigated in sensory science for the last 60 years using different methodologies (Kemp et al., 2017). Intensity-based methods requiring a trained panel, such as Time-Tntensity (Lee & Pangborn, 1986), progressive profiling (Jack et al., 1994) or sequential profiling (Methven et al., 2010), are still used, but the current trend is in favour of rapid methods. Indeed, sensory analysis tends to work increasingly with a panel of consumers in the natural contexts of consumption (Jaeger et al., 2017; Jaeger & Porcherot, 2017; Meiselman, 2013). Among the temporal methods, Temporal Dominance of Sensations (TDS) (Pineau et al., 2009) and Temporal Check All That Apply (TCATA) (Castura et al., 2016) have already been successfully used with consumers (Ares et al., 2016, 2015; Dinnella, Masi, Zoboli, & Monteleone, 2012; Hutchings, Foster, Grigor, Bronlund, & Morgenstern, 2014; Jaeger et al., 2018, 2017; Rodrigues et al., 2016; Thomas, Visalli, Cordelle, & Schlich, 2015; Visalli, Lange, Mallet, Cordelle, & Schlich, 2016). Indeed, TDS requires little or no training (Albert et al., 2012; Di Monaco et al., 2014), even if most studies include a briefing phase (Albert et al., 2012; Hutchings et al., 2014; Rodrigues et al., 2016; Thomas et al., 2015), which is also suggested for TCATA (Jaeger et al., 2017). TDS and TCATA consist of measuring in continuous time and concurrently tasting the evolution of a predefined list of descriptors by clicking on dominant or applicable descriptors at any given time. Even if they are based on different concepts (Meyners, 2020), these methods are frequently compared. If they are usually in general agreement, TCATA has tended to pick up more

differences than TDS (Ares et al., 2015; Berget et al., 2020; Esmerino et al., 2017; Kawasaki et al., 2019; Nguyen et al., 2018). However, these differences are subtle, and even if most of them may be due to the task (dominance vs. applicability), the level of precision and replicability of these methods is not well documented.

Outside the field of sensory analysis, even if real-time data capture has been judged "applicable and preferable" for measuring changes over time by psychologists (Stone et al., 1999), it does not eliminate other potential sources of bias in self-reports (Schwarz, 2012). Indeed, focusing on differences over time is a demanding task for consumers, which can be a potential source of differences in panellists' behaviour (Varela et al., 2018), as previously observed by Pineau et al. (2012). Conceptually simpler methods that do not involve continuous duration. The Temporal Order of Sensations (TOS) (Pecore, Rathjen-Nowak & Tamminen, 2011; Torres-Moreno, Hort, & Tarrega, 2016), a method based on the concept of key descriptor, was developed with the specific objective to capture a particular attribute of interest. In TOS, panellists select from a list the first 3 attributes they perceive during the tasting. Another method called "Pick 3 and Rank" (P3R) (Vandeputte, Romans, Lenfant, & Pineau, 2011) was used to measure the temporality between several bites corresponding to a full portion. P3R consists in retrospectively picking then ranking the 3 most important descriptors perceived during a bite. Neither TOS nor P3R does take the whole duration of the tasting experience into account. This limitation could be leveraged by forcing a description in 3 phases to take into account the entire tasting experience, as is common in the world of wine with the attack-evolution-finish sequence (Grainger, 2009; Harrington, 2008; Osterland, 2012; Spence & Wang, 2018). For wine professionals, the attack phase is the initial impression that the wine makes on the palate. The evolution phase (also called mid-palate or middle range perception) focuses on the flavour profile. The finish is the final phase, where the aftertaste comes into play.

This article introduces a new retrospective temporal method called Attack-Evolution-Finish (AEF) inspired from TOS, P3R and the tasting in 3 phases. AEF proposes to retrospectively split the tasting in 3 periods and then to select the most important descriptor during each period. As one can ask if such a method would result in a loss of information, the article compares the results from AEF and TDS studies on dark chocolates, the selection of the important descriptors being related to the concept of dominance. Regarding the results, the article discusses several aspects of sensory temporal data collection and their respective benefits and defects.

2. Material and methods

2.1. Protocol

2.1.1. Samples

The 5 dark chocolates provided by Barry Callebaut and labelled as C54, C65, C68, C70 and C73 were composed of 54%, 65%, 68%, 70% and 73% cocoa, respectively. Each chocolate was given to the consumers in a sealed transparent plastic container (height 3 cm, diameter 4 cm) labelled with a 3-digit code and containing 4 "callets" (pucks of chocolates formulated for melting rather than baking) of 0.5 g each that had to be completely consumed in a single intake. The samples were presented at ambient temperature in a sequential monadic order according to a Williams Latin square.

2.1.2. Consumers

The study took place at the Centre for Taste and Feeding Behaviour, Dijon, France. Two panels of consumers who were regular consumers of dark chocolates (at least once every two weeks) were recruited from a population registered in the ChemoSens Platform's PanelSens database. This database has been declared to the relevant authority (Commission Nationale Informatique et Libertés—CNIL—no. d'autorisation 1148039). The consumers were financially rewarded for their participation in the study. Panel 1 was composed of 68 consumers (36 men and 32 women aged 19 to 63 years old, the average age of men being 41 and the average age of women 40), and panel 2 was composed of 61 consumers (35 men and 26 women aged 19 to 61 years old, the average age of men being 42 and the average age of women 41).

2.1.3. Descriptors

The descriptors were chosen according to the expertise of Barry Callebaut and the experience of previous studies. The same list of descriptors was proposed in both the AEF and TDS sessions (the original French terms are mentioned in brackets): Astringent (Astringent), Bitter (Amer), Cocoa (Cacao), Dry (Sec), Fat (Gras), Floral (Floral), Fruity (Fruité), Melting (Fondant), Sour (Acide), Sticky (Collant), Sweet (Sucré), and Woody (Boisé). The descriptors were presented in a random order on the screen but this order was constant for each consumer.

2.1.4. Sessions

The 2 panels carried out both the TDS and AEF sessions; panel 1 started with the TDS method, and panel 2 started with the AEF method. Forty-eight hours separated the two sessions for each panel.

[INSERT FIGURE 1 HERE]

Figure 1: TDS (left) and AEF (right) measurement screens (translated from French).

TDS protocol and instructions

The consumers were briefed in groups just before the session. The concept of dominance was defined as "the sensation that catches the attention", and an example of a TDS screen was presented. No specific explanation about the descriptors was given, but consumers could ask questions. The sessions took place in individual booths running TimeSens© software 2.0 (INRA, Dijon, France). The consumers were reminded of the instructions on the first screen of the session as follows: "You will describe each chocolate by clicking at any moment on the sensation that catches your attention. A sensation can be clicked several times or never. There are no constraints on the number of sensations clicked. You will have to click on START at the same time you put the chocolate in your mouth and on STOP when you no longer perceive anything". Before tasting each sample, the attribute list was displayed on the screen as follows: "Here is the list of sensations available: astringent, bitter, cocoa, dry, fat, floral, fruity, melting, sour, sticky, sweet, woody. On the next screen, please remember the location of each of these sensations before you click on START". On each TDS measurement screen (Figure 1, left), the following instruction was displayed: "Now, put the 4 callets of chocolate in your mouth and immediately click on START, then click on the sensations that catch your attention. Once you do not perceive any more sensations, click on STOP". After each sample, the consumers had to rinse their mouth with water during a forced 30-second break.

AEF protocol and instructions

The experimenters assumed that the task was self-explanatory. Therefore, contrary to the TDS session, the consumers were not briefed and instead directly took their places in individual booths running TimeSens© software 2.0. The instructions were presented on the first screen of the session as follows: "You are going to taste 5 chocolates. After each tasting, we will ask you to choose from a list the 3 sensations that you perceived during the tasting, in the chronological order in which you perceived them. Here is the list of sensations available: astringent, bitter, cocoa, dry, fat, floral, fruity, melting, sour, sticky, sweet, woody". An example was provided on the second screen as follows: "Example: At first, I perceived this chocolate *sour*, then after a few moments I perceived it *fruity*, and at the end of the tasting I perceived it *sweet*. You can use the same sensation several times; for example: At first, I perceived this chocolate *sour*, then after a few moments I perceived it *sour*, and at the end of the tasting I perceived it *sweet*". For each sample, measurement was separated into two screens. The first screen measured the duration of the tasting as follows: "Put the 4 callets in your mouth and taste them. Focus on the chronological order of the 3 perceived sensations! When the tasting is finished, click on the 'NEXT' button to indicate the perceived sensations". A minimum time of 10 seconds was set for this screen. The second screen displayed the list of descriptors (Figure 1, right) and asked the following: "What sensations did you perceive during the tasting, in chronological order? (Click on the drop-down lists to answer). At first, I perceived this chocolate..., then after a few moments I perceived it..., and at the end of the tasting I perceived it...". In the rest of this article, the first sensation chosen will be referred to as "attack" (A), the second as "evolution" (E) and the third as "finish" (F). After each sample, the consumers had to rinse their mouth with water during a forced 30-second break.

Questionnaire

After each session of either AEF or TDS, a questionnaire adapted from the one used to compare TDS and TCATA (Ares et al., 2015) was displayed to assess the difficulty of the tasks.

The items of the questionnaire (see table 3) were measured on a Likert scale (Likert, 1932) using the following labels: "strongly agree", "agree", "neither agree nor disagree", "disagree", and "strongly disagree".

An additional open-ended question was asked about their overall opinion about the experiment.

After the second session (AEF for panel 1, TDS for panel 2), the following question concerning the relative comparison of AEF and TDS was asked on a five-point scale: "Compared to the task in the first session, did today's task seem to you to be 'much easier', 'easier', 'at the same level of difficulty', 'more difficult', or 'much more difficult?'" An open-ended question about remarks oriented towards the comparison of the two methods was also asked.

2.2. Data analysis

For the purpose of comparison, when necessary, the TDS sequences were split into 3 sequences of equal sizes. Correspondences with standardized times were established as follows: A = {0 - 0.33}, E = {0.34 - 0.66}, F = {0.67 – 1}).

AEF data were structured in a table with 5 columns "Consumer", "Product", "Descriptor", "Period (A/E/F)" and "Citation (0/1)" and 15480 rows (129 consumers x 5 products x 12 descriptors x 3 periods). For a given consumer and a given product, the sum of citations is equal to 3.

Statistical analyses were performed using R 3.5.0 software (R Core team, 2017) and TimeSens© 2.0 (INRA, Dijon, France).

2.2.1. Consumers' behaviour in TDS

The distributions used to assess heterogeneity in the consumers' behaviour in TDS were those defined in Lepage et al. (2014) and Visalli et al. (2016), namely, the distribution of the number of descriptors per TDS sequence, the distribution of the number of citations per TDS sequence, the distribution of the times of first citation per TDS sequence, and the distribution of the tasting duration per TDS sequence. The averages per consumer of these 4 indices were also considered.

The Pearson coefficients of correlation have been computed and tested against 0 for the following pairs of vectors: number of attributes / number of citations, number of citations / total durations, first times of citation / total durations.

The differences between TDS and AEF means in tasting durations (from START to STOP) were evaluated using an ANOVA model, i.e., duration = subject + product + method + 2-way interactions, with subject and subject interactions being random factors. It was followed by a Tukey LSD post-hoc test, with alpha = 0.05.

2.2.2. Sequentiality of sensations

TDS curves (Pineau et al., 2009) were plotted using TimeSens software, the times being standardized between 0 (time of citation of the first attribute) and 1 (time of click on STOP). The significance lines were drawn with alpha = 0.10, based on a binomial proportion test and the chance level being defined as 1/number of descriptors (12). To facilitate the visual correspondence between the TDS and AEF periods, vertical segments were added at t=0.33 (end of period A) and t=0.66 (end of period E).

For AEF, the proportions of dominances were represented as 3 side-by-side barplots, one for each period. To facilitate comparison with TDS, the significance lines were drawn in the same way as those for TDS.

The conclusions based on the TDS curves and AEF barplots are summarized in a table containing the significant attributes. As significance lines have been contested for the TDS curves (Meyners & Castura, 2019), the attributes were considered significantly dominant when their 90% simultaneous multinomial confidence interval (Goodman, 1965)

lower bounds were greater than (1/number of descriptors), as suggested by Meyners & Castura (2018). For AEF, the proportions of each attribute were considered for each period. For TDS, the proportions of each attribute at its maximum peak inside each period were considered. The function "MultinomCI" of the package DescTools was used.

2.2.3. Unidimensional analyses

For each TDS sequence, durations by descriptor were computed for the whole sequence and by period (A, E, F). Then, they were analysed using an ANOVA model, i.e., "duration = subject + product + error" (Galmarini et al., 2017).

For each AEF sequence, the total number of citations (0 to 3) by descriptor was computed. Then, a Poisson log-linear model for count data was fitted, i.e., "total number of citations = subject + product + error". Overdispersion (ratio "residual deviance / degrees of freedom") has been checked less than 1 using function "dispersiontest" of package AER. Residuals were checked using randomized quantile residuals by Dunn and Smyth (1996) with function "simulateResiduals" of package DHARMa. For each period and descriptor, a binomial model for binary data was fitted, i.e., "citation (0/1) = subject + product + error". Both models were fitted (function "glm", parameter family ="poisson" or "binomial") using a generalized linear model (GLM, Agresti, 2013). Then, analyses of deviance for generalized linear model fits were performed (function "Anova" of package car) using a likehood ratio (LR) test and assumed to be chi-square distributed. The null hypothesis of this test was that the count data are unrelated to the Product factor. It is to be noticed that, even if the GLMs should be preferred over linear models (LM), standard LM tests are robust and can have good type I error control, so they can also be used for counts (Warton, 2016). Thus, the same analyses have been conducted with LMs and the conclusions (not presented) were the same with slight differences in p-values. When they were significant (alpha=0.10), ANOVAs and analyses of deviance were followed by a Tukey post-hoc test (alpha=0.10). The functions "lsmeans" from the package lsmeans was used.

2.2.4. Multidimensional analyses

Canonical Variate Analysis (CVA) was used to represent the product map of the TDS durations (Galmarini et al., 2017) over all periods, with a level of confidence set to 90% for the binormal distribution of the product ellipses. Discrimination between product pairs was established using a Hotelling T^2 test on all dimensions. The trajectory Principal Component Analysis (PCA) of the TDS durations (Lenfant et al., 2009) at 3 points corresponding to the periods of AEF was also plotted to evaluate the within-product evolution. TimeSens software was used to produce theses graphics.

As AEF data were count data, correspondence analysis (CA) of the contingency table product x descriptor (5 rows: chocolates, 12 columns: descriptors) was used to represent the product map of citations of descriptors over all periods of AEF. The individual profiles were projected as supplementary elements, and then the covariance matrix related to these projected points was calculated and used to draw confidence ellipses under a binormal assumption (Saporta & Hatabian, 1986), with a level of confidence set to 90%. Discrimination between product pairs was established using a Hotelling test on all dimensions. Trajectory CA (Castura et al., 2016) of the contingency table product/period * descriptor (15 rows: 5 chocolates x 3 periods, 12 columns: descriptors) was used to represent the within-product evolution over the 3 periods. The function "CA" of the package FactoMineR was used.

2.2.5. Answers to questionnaire

The Likert scale labels were transformed into scores between 1 (strongly disagree) and 5 (strongly agree), which were averaged over the consumers. Then means were compared with a two-tailed one sample t-test against a known mean of 3 (corresponding to "neither agree or disagree"). The relative difficulty scale was coded between -2 (TDS much easier) and 2 (AEF much easier). An ANOVA model, i.e., "difficulty = method + panel + error", was then performed. The answers to the open-ended questions were qualitatively analysed.

3. Results

3.1. Consumers' behaviour in TDS

[INSERT FIGURE 2 HERE]

Figure 2: Histograms of consumer behaviour observed during the TDS sequences. A: Distribution of the number of descriptors used in a sequence, B: Distribution of the number of dominant sensations (number of clicks), C: Distribution of the times of first dominant sensations, D: Distribution of the total durations of the evaluation.

Figure 2A shows the distribution of the number of descriptors observed during the TDS sequences. The numbers of descriptors varied from 1 to 10. Two thirds of the sequences included 3 to 6 distinct descriptors, thus slightly or somewhat less than one half of the proposed descriptors. Figure 2B shows the distribution of the number of dominant sensations observed during the TDS sequences. The number of dominant sensations varied from 2 to 23. A representative sequence included 4 to 8 dominant sensations. Compared to statistics of Figure 2A, it means that one or two descriptors in average are used twice in a given sequence (in AEF, 6% of sequences had twice the same descriptors, less than 0.2% had three times the same descriptors). Figure 2C shows the distribution of the times of first dominant sensations observed during the TDS sequences. These times varied from 1 to 30 seconds. Approximately 60% of the consumers cited their first attribute before 8 s, and about 15% reported it after 12 s. Figure 2D shows the distribution of durations observed during the TDS sequences. These durations varied from 5 to 125 seconds with a coefficient of variation of 46%. Generally, the 4 distributions were characterized by a positive skewness (long right tails), and a large heterogeneity (coefficients of variations of 32, 51, 83 and 56% respectively).

The coefficient of correlation between the number of attributes and the number of citations was significant ($r=0.74$, $p<0.001$). The coefficient between the number of citations and the total durations was also significant ($r=0.42$, $p<0.001$) but with a lower coefficient. It means that consumers having cited a large number of descriptors have the longer durations. The coefficient between the first times of citation and the total durations ($r=0.20$, $p=0.02$) was also significant, but with the lower coefficient. It denotes a tendency for the consumers having starting the earlier to have shorter evaluation durations while consumers having a late start have longer evaluation durations.

The mean duration of the tasting (not represented) associated with the AEF task was 30.2 s with a standard deviation of 22 s (CV=73%).

The difference of durations between the two methods ($F=297.32$, $p<0.001$) was significant, the mean duration being longer in TDS (46.2 s) than in AEF (30.2 s). The differences of durations between the 5 products were also significant ($F=2.84$, $p=0.023$), the post-hoc test showing that C54 was perceived longer than C65. The interaction "product x method" was not significant, meaning that the difference between the products were not significantly different with the two methods.

3.2. Sequentiality of sensations

[INSERT FIGURE 3 HERE]

Figure 3: TDS curves (left) and AEF barplots (right) of the proportion of dominances for C54, C65, C68, C70, and C73 (from top to bottom). The grey mask corresponds to the region below significance level (as defined in TDS). The descriptors significant in sense of the binomial test are summarized below each pair of figures. The descriptors also significant in sense of the multinomial test are in bold and followed by a letter in superscript if the test was significant for one method only (A for AEF, T for TDS).

Figure 3 represents the TDS curves and AEF barplots for the 5 chocolates. Overall, whatever the method, except for C54, the proportions of dominance were not very high (below 30%). The levels of the attribute that reached the highest dominance rate observed in this study inside each period were comparable. Without going into details, the product temporal profiles were relatively similar between the two methods in the sense that the main dominances (those being largely above significance) were the same. The multinomial test was less liberal than the binomial test to

determine the dominant descriptors (21% and 43% of significances were lost using the multinomial test in TDS and AEF respectively). In TDS (AEF), for C54, C65, C68, C70, and C73, the binomial test showed 8 (5), 13 (9), 14 (10), 12 (9) and 14 (9) significant attributes (a total of 61 in TDS and 42 in AEF), while the multinomial test showed 7 (4), 9 (5), 12 (7), 9 (5) and 11 (4) (a total of 48 in TDS and 24 in AEF). Whatever the test, TDS systematically showed more dominant descriptors compared to AEF (45% and 100% more with the binomial and the multinomial tests respectively), but these additional significant descriptors were almost all below 15%; therefore, it is suggested that AEF captures dominances established with certainty and TDS adds a number of potential dominances. It seems that the percentage of Cocoa in the chocolate was not the main driver of perceived dominances for Bitter and Astringent, the descriptors having been cited more dominant at a panel level for C68 and C70 than for C73. One can note that Floral was never significant whatever the method and the test. Based on the binomial, but not on the multinomial test, Fruity reached significance in TDS only and in one product only. It is suggested that Floral and Fruity, which might be applicable in black chocolates, may not be adequate to use with consumers in a TDS or AEF paradigm.

3.3. Unidimensional analyses

Table 1: ANOVA of durations or citations by descriptor, method and period.
The columns 4 to 6 report the p-values for the product effect for TDS (F statistic, df: 4) ~~or~~ and AEF (LR statistic) for each period, bold values being significant with alpha = 0.10.
The columns 7 to 10 reports the number of product pairs discriminated (NPD) by period (9 possible comparisons).

Attribute	Protocol	p-values All periods	p-values Attack	p-values Evolution	p-values Finish	NPD All periods	NPD Attack	NPD Evolution	NPD Finish
Astringent	TDS	0.106	0.415	**<0.001**	0.107	2	-	4	-
	AEF	**0.027**	0.588	**0.01**	**0.064**	1	-	0	1
Bitter	TDS	**<0.001**	**<0.001**	**<0.001**	**<0.001**	8	6	5	4
	AEF	**<0.001**	**<0.001**	**<0.001**	**<0.001**	7	5	6	5
Cocoa	TDS	0.286	**<0.001**	**0.096**	**0.049**	3	3	1	1
	AEF	**<0.001**	**0.004**	**0.051**	0.127	3	2	1	-
Dry	TDS	**0.001**	**<0.001**	**<0.001**	0.305	7	8	4	-
	AEF	**<0.001**	**<0.001**	**0.01**	0.98	7	8	3	-
Fat	TDS	0.496	0.131	0.331	**0.099**	-	-	-	0
	AEF	0.102	**0.001**	0.502	0.12	-	3	-	-
Floral	TDS	0.123	**0.013**	**0.03**	0.128	-	3	1	-
	AEF	0.180	0.31	0.377	0.659	-	-	-	-
Fruity	TDS	**0.032**	0.621	**<0.001**	**0.034**	2	-	4	1
	AEF	**0.044**	**0.054**	**0.066**	**0.025**	1	0	1	1
Melting	TDS	**0.049**	0.72	**0.001**	**<0.001**	4	-	3	4
	AEF	**0.001**	**0.069**	**0.003**	0.184	3	0	2	-
Sour	TDS	**0.002**	**<0.001**	**<0.001**	**<0.001**	6	4	4	4
	AEF	**<0.001**	**0.005**	**<0.001**	**0.009**	4	1	3	2
Sticky	TDS	**0.013**	**0.002**	**<0.001**	0.125	4	3	5	-
	AEF	**0.015**	**0.001**	0.406	**0.002**	4	0	-	2
Sweet	TDS	**<0.001**	**<0.001**	**<0.001**	**<0.001**	5	6	4	4
	AEF	**<0.001**	**<0.001**	**<0.001**	**<0.001**	7	3	5	6
Woody	TDS	**0.044**	0.105	0.134	**0.008**	3	-	-	2
	AEF	**<0.001**	**0.011**	**0.043**	**0.014**	3	1	1	1

Table 1 shows that all periods combined and with alpha = 0.10, the TDS discriminates products for 8 descriptors and the AEF discriminates products for 10 descriptors. All attributes except Fat and Floral (plus Cocoa and Astringent for TDS) had significant differences in durations (TDS) or citations (AEF). The tests performed all periods combined were independent from those performed by period. A non-significant statistic all periods combined does not necessarily imply that the statistics by period would also be non-significant. Thus, in periods A, E and F, the TDS discriminates products for 7, 10 and 8 descriptors and the AEF discriminates products for 10, 9 and 7 descriptors. The number of descriptors discriminating the products was higher in period E and F for TDS, while it was higher all periods combined and in period A for AEF. Taking everything into account, AEF seems to discriminate slightly more products than TDS. The conclusions of the pairwise comparison tests by period were in accordance overall, but when TDS discriminates products with a given attribute, more product pairs were separated (44, 33, 35 and 20 in TDS versus 40, 23, 22 and 18 in AEF for all periods then periods A, E and F). This happened notably for Astringent (in period E), Floral (in periods A

and E), Fruity (in period E), Melting (in period F), Sour (all periods) and Sticky (in periods A and E). However, one should keep in mind that this table compares durations of dominance to frequencies of citations, two different concepts.

3.4. Multidimensional analyses

[INSERT FIGURE 4 HERE]

Figure 4: Multidimensional maps, axes 1 and 2. A - CVA of the TDS durations for all periods combined, with 90% confidence ellipses. B – Trajectory PCA of the TDS durations over the 3 periods of AEF. C – CA of the AEF sensations for all periods combined, with 90% confidence ellipses. D – Trajectory CA of the AEF sensations over the 3 periods of AEF.

All periods combined and considering all axes (figures 4A and 4C), the between-product discriminations were the same in TDS and AEF, and all product pairs were discriminated with an α-risk less than 0.1%. 91.1% of the total variance was explained on axes 1 and 2 of CVA of durations (figure 4A), and the test for the determination of the number of significant axes (not presented here) recommends to keep 3 axes. The 2 first axes of the CA of citations (figure 4C) of citations explained 94.4% of variance. Whatever the map, the first axis mainly separated the products in 3 poles, basically C54 (perceived Sweet and Dry for a longer duration/by more consumers than the others), C68/C70 (perceived Astringent, Bitter, Woody and Sour for a longer duration/by more consumers) and C65/C73 (perceived Cocoa, Fruity and Melting for a longer duration/by more consumers).

The within-product evolution (figures 4B and 4D), was very similar in TDS and AEF. The PCA of durations (figure 4B) explained 81.6% of the variance on the 2 first axes. The axis 1 separates the products in the same way as the CVA, the axis 2 showing the evolutions of all products except C65 from Dry to Sweet (C54) or to Sticky, Astringent and Bitter (C68, C70 and C73). The CA of citations (figure 4D) explained 74.8% of the variance and can be interpreted in the same way as the PCA.

3.5. Answers to questionnaire

Table 2: Mean answers to the questionnaire and their 95% confidence intervals

	TDS	AEF
1. The oral explanations were useful (TDS) / I wish I had oral explanations (AEF).	4.35 ± 0.18	2.53 ± 0.17
2. The explanations displayed on the screen about how to evaluate chocolates were useful (TDS) / sufficient (AEF).	4.59 ± 0.16	4.53 ± 0.13
3. I understood how to evaluate the chocolates.	4.87 ± 0.09	4.75 ± 0.11
4. The list of sensations was exhaustive.	4.09 ± 0.16	3.80 ± 0.14
5. The sensations were sufficiently explanatory.	4.34 ± 0.15	3.93 ± 0.14
6. I wished I could select more than 3 sensations.		2.62 ± 0.16
7A. It was easy to identify the sensations that caught my attention during the tasting (TDS). 7B. It was easy to identify the sensations perceived at the beginning of the tasting. 7C. It was easy to identify the sensations perceived at the middle of the tasting. 7D. It was easy to identify the sensations perceived at the end of the tasting.	3.88 ± 0.17	 3.25 ± 0.16 3.98 ± 0.14 3.83 ± 0.13
8. It was easy to quickly click on a sensation when it caught my attention.	4.50 ± 0.10	
9. It was easy to identify when to click STOP.	3.94 ± 0.18	
10. The order in which I listed the sensations was important.	4.59 ± 0.16	4.10 ± 0.13
11. I could list the same feeling several times.	4.81 ± 0.14	4.22 ± 0.15
12. The questionnaire interface was easy to use.	4.85 ± 0.08	4.75 ± 0.07
13. The task that was asked of me was easy.	4.51 ± 0.11	4.40 ± 0.13

Table 2 shows that all the answers were in the direction of positive agreement (easiness, usefulness). All means were significantly different from 3 (neither agree nor disagree) with $p < 0.001$ except for questions 1 and 6 (AEF). It means that the oral explanations were declared useful for TDS and not for AEF and that the consumers did not declare to be limited by the restriction imposed on the selection of only 3 sensations in AEF.

The average relative difficulty score obtained from the comparison of the two methods by the same consumers was -0.15, significantly lower than 0 (F=6.3, p=0.02), which means that TDS was judged to be slightly easier than AEF. The panel effect was also significant (F=5.1, p=0.03), with the last method that was used being judged easier.

In the open-ended comments, the methods were declared "simple" (6 consumers). Favourable TDS comments emphasized that "having the list of sensations in front of the eyes and choosing them at the very moment of feeling is easier, more spontaneous" (5) and that "it is easier to choose the sensations right after placing the chocolate into your mouth or during the tasting" (7). Positive comments about AEF reported the "easier choice after the tasting, no time pressure" (6), "easier because less choices to do, quick, concise" (4). Negative comments about AEF concerned the difficulty of "picking only 3 sensations" (5), "differentiating the periods" (3) or "memorizing the order" (5). Negative comments about TDS were about the "lack of clarity of the explanations" (2) or "a need for more instructions, a warm-up, a training" (2). Regardless of the method, several consumers also found it "too bad for not being able to evaluate more flavours" (2), that "some descriptors are useless" (3), that "some descriptors are missing" (2), or they regretted not being able to "express a free opinion" (1).

4. Discussion

Overall differences in conclusions obtained with TDS and AEF

AEF measured temporality in products in a minimalistic way, as each sequence was composed of 3 descriptors. Even if this was not a limit evoked by the consumers in their answers to the questionnaire, the results showed that most of the TDS sequence included 4 to 8 dominant descriptors. By construction, TDS curves had more dominant descriptors than AEF. As one would expect from the methods, TDS can pick up short-lived differences in the moment and AEF picks up big remembered differences. However, the big picture is the same whatever the method, and the attributes that were dominant in TDS but not in AEF had low dominance rates that denoted a poor agreement. We suggest using the multinomial test for deciding whether these additional dominances should be considered ~~or not~~. For these reasons, regarding absolute product description, AEF could miss some sensations compared to TDS. It was confirmed by the results of unidimensional analyses (section 3.3), which show that TDS discriminated more product pairs. Yet, when comparing products performing a multidimensional analysis (section 3.4) based on Hotelling test (thus accounting for correlations between descriptors), TDS was not superior to AEF. The 2 methods were able to discriminate all product pairs.

Possible consequences of heterogeneity in durations and number of citations in TDS

Heterogeneity in product durations can have an impact on statistical analysis such as ANOVA or CVA. Indeed, if the products have different durations (as observed between C54 and C68), then the differences in the durations between the descriptors can be due to differences in the product durations. For example, in this study, the total duration was longer for C54 than for C68. This could result in small differences in ANOVA results by attribute when considering the durations or standardized durations (Galmarini et al., 2017).

Because of the observed heterogeneity of the times of first citations and total durations among consumers, left and right time standardization was required to represent the TDS curves. If data standardization erases differences in evaluation time by aligning individual perceptions, it can change individual patterns substantially enough (Meyners, 2018).

The mean values observed in this study for the number of descriptors and the number of citations were consistent with the mean values reported in the literature (Pineau et al., 2012). However, in this study as in others (Hutchings et al., 2014), again large heterogeneity was observed. A high number of citations per judge can in particular lower the panel agreement or make the proportions of the TDS curves unstable because of citation times that are more difficult to align.

The differences observed between the two methods are probably due to the task than to the heterogeneity in citations and durations. In AEF, heterogeneity (CV of 73%) also existed regarding the durations of the tastings. It was even more apparent than in TDS (CV of 46%), but only citations (3 per consumer) and proportions were considered in unidimensional and multidimensional analyses. Thus, it would be interesting to conduct a follow-up study to determine whether the conclusions of AEF would be more reproductible than those of TDS.

Concurrent vs. retrospective measurement

In both TDS and AEF, the stimulus was detected, interpreted, and then transcribed as a click on a descriptor in the list. In TDS, the action of clicking is hoped to be as close as possible to perception, but in fact, a minimum and individual-dependent delay is necessary as confirmed by the mean time of first citations. As it was unlikely that the first descriptor has been perceived after 8 seconds, this delay was probably due to a cognitive load in choosing the dominant descriptor or to consumers having pressed the START button before to put the sample in mouth. It should be interesting to check if the results from a trained panel would have been impacted in the same way. Besides, longer total durations have been observed in TDS compared to AEF, but it is hard to know whether TDS overestimates durations or AEF underestimates them. It is to be noticed than differences in the evaluation times between methods evaluating the same products have already been observed (Meyners, 2020), the durations being longer with TCATA and TDS by modality compared to TDS. In any case, the task has an impact on delays and durations that must be considered when coupling TDS measures with true real-time data such as Proton-transfer-reaction mass spectrometry (Mesurolle, Saint-Eve, Déléris, & Souchon, 2013) or facial recognition (Urbano, Mahieu, Thomas, Schlich, Visalli, 2018).

The answers to the questionnaire showed that TDS was considered in average slightly easier and more natural than AEF, but it should be noted that TDS has been very frequently used in Dijon, and in this study, 70% of the consumers had experience with the method in previous consumer tests. In addition, contrary to TDS, the consumers were not briefed before evaluating the chocolates with AEF. Answers to more specific items of the questionnaire did not allow going further in understanding the perceived differences between the two methods, probably because the consumers answered the questions in a manner that will be viewed favorably by the experimenter (bias of social desirability, Cerri, Thøgersen, & Testa, 2019). However, the (few) answers to the open-ended question about the overall opinion of the task suggested than some people felt more comfortable when having the list of attributes during the tasting, while others were stressed by the measurement during the tasting. In any case, the results of AEF are congruent with those observed in TDS and the temporality of the products was caught without relying on durations.

The retrospective action of clicking in the AEF method required a memorization effort that probably resulted in less spontaneous answers. The memory can be involved in different ways. First, the consumer can taste and take mental notes of the descriptors that describe the 3 periods, then pick these descriptors after the tasting because he remembered the descriptors. In this case, the consumer reports what he experienced with a delayed perception. Second, the consumer can taste and remember the experience then pick the descriptors later. In this case, the consumer reports what he remembered of the experience.

Continuous time vs. periods

In the TDS curves, the proportions sum to 1 at each time. This means that for a given descriptor, the proportions at times t and t+1 can be due to different consumers having a delayed perception or different perceptions (Cardot et al., 2019). Thus, the overall TDS panel overview can correspond to the perception of nobody. In AEF, as the proportions are computed with the whole panel in each period, it really corresponds to the panel overview, assuming each consumer has a similar conceptualization of the periods. However, it is more difficult to have several descriptors that pass the significance inside a given period, especially when there is a high agreement on a descriptor. This explains the difference observed between the significant number of descriptors in AEF barplots and TDS curves.

In TDS, to simplify the statistical analyses, time has sometimes retrospectively been split into 3 periods of uniform durations (Dinnella, Masi, Naes, & Monteleone, 2013; Lepage et al., 2014). This transformation makes TDS and AEF even closer conceptually. However, in AEF, the choice of the temporal periods corresponding to "attack", "evolution"

and "finish" is left to the consumer, which makes the frontiers of the periods subjective. Particularly, the evolution phase is conceptually different from the attack and finish phases. Indeed, the frontiers of the evolution phase are not clearly temporally defined, being in between attack (first sensation) and finish (last sensation). It would be interesting to compare the results obtained within each period of AEF to those of TDS curves having their frontiers automatically determined using semi-Markov chains (Lecuelle, Visalli, Cardot, & Schlich, 2018).

Citations vs. durations

In TDS, durations are computed as the differences between the times of citation of the new and the previous dominant descriptors. If periods of “no dominance” exist, they do not seem to interfere with the duration-based conclusions. As durations are quantitative measures, it is natural to use common QDA statistical analyses such as PCA or CVA. However, both TDS and AEF are methods based on attribute citations, and TDS could also be analysed with CA (Frost et al., 2018). CA takes into account sums of citations while PCA and CVA weight the citations by the durations of the attributes. While durations may seem more precise, in this study, CVA and CA maps discriminated all product pairs in the same way, suggesting than in this study durations did not provide additional information compared to citations.

Dominance in AEF and TDS

The ISO standard (ISO, 2016) recommends to define the dominance sensation in TDS as “the sensation that catches his/her attention at a given time, which does not mean that this sensation has to be very or the most intense in the product”. In practice, the concept of dominance is not clearly established and several definitions exist (Varela et al., 2018). Indeed, dominance may be diverse among individuals, and can be related to novelty, intensity, or cognition, but “it easily measures something useful” (Schlich, 2017). Considering this, in AEF, rather than giving an explicit definition for dominance, the forced selection of only one descriptor by period implicitly led the consumer to report his “dominant” sensation, whatever his representation of dominance was. Determining if the concept of dominance were really the same in TDS and AEF remains an open question, but it is out of the scope of this article. In all cases, consumers have faced a forced choice involving a degree of subjectivity. Thus, it should still be noted that if TDS can be biased by the halo-dumping effect (Varela et al., 2018), AEF can certainly be biased in the same way, especially with descriptors such as Sour and Astringent that are closely related.

Briefing or not?

In TDS, the consumers have to understand the concept of dominance. They have to click on START and STOP buttons. They also have to declare in real-time what dominant sensations they perceived by consecutively clicking on several buttons. This makes the briefing phase recommended to avoid possible miscomprehension of the task, as previously reported in a Temporal Dominant of Emotions study performed at home (Peltier et al., 2019). The AEF task was quite simple: the consumers just had to select one descriptor in each of the 3 proposed drop-down lists. It does not mean that attribute selection was easy and that consumers did not ask themselves about the criterion of their choice, but the task itself was self-explicit, easy to explain with instructions on screens and thus did not require a briefing.

Recommendations about AEF

This study should be considered as a preliminary study and the method still needs to be investigated and replicated before being validated. AEF relies on a retrospective evaluation of products at 3 specific periods. In this, AEF is a paradigm shift from current temporal methods that collect data simultaneously with tasting. Thus, the method is clearly not suitable when durations are of interest, when an accurate measure of temporality is required or when subtle changes occur at very specific moments of the tasting. At the same time, AEF presents some benefits. AEF does not seem to require training or even briefing. As it does not rely on durations, AEF data does not need to be time-standardized, the heterogeneity due to the continuous time measurement is reduced and the statistical analysis of AEF is simple. There are several scenarios for which AEF may prove to be interesting. An obvious case is when the panel leader does not have software that allows to record durations. Indeed, the results of the method can be collected just using a paper and a pen. AEF seems suitable for capturing the big picture of temporality or for comparing several

products, especially for at-home studies, when no briefing phase is possible. AEF could also be more suitable than other temporal methods when there is a long list of descriptors (more than 10). In this study, AEF was used to measure dominances in an implicit way, without specifying the definition of dominance. This choice is debatable, and the ISO definition could have been used as well. The most important concept underlying AEF is the evaluation by period rather than in continuous time: thus, if dominance is judged too multidimensional, asking for the most intense sensation in each period could be preferred. As a perspective, AEF could even be extended to "multiple AEF" to record applicable descriptors during each period, making it comparable to TCATA, with the advantage of not needing the fading option (Ares et al., 2016; Rizo et al., 2020).

Precision of temporal methods and method comparisons

The authors think that introducing the time dimension adds new information to sensory data, but also a new source of uncontrolled variability. Thus, rather than promoting a new method, the presented results and elements of discussion question the granularity of temporal sensory data.

The first question to ask should be "What level of precision is expected for the time aspect of the measurement?". Meyners (2020) said that "relating to a piece of music, we are thinking of TDS at rather tracking changes in the melody, while TCATA explores the harmonies.". Using another analogy, AEF could be the trailer of a movie. Like a book cover, it doesn't tell all of it but it gives a few impressions of the story. Is it more interesting to learn that people tasting a chocolate perceive it Cocoa after 5 seconds then Bitter 2 seconds after, or that after consuming it they remember the attack mainly as Bitter? Again, what is more useful for industry depends on the objective but also on the product complexity. If the objective is a new product development or optimisation and if the manufacturing process enables changes that can have an impact at very specific time points, thus maybe the most precise tool using trained panels is required. If the products only differ by their attribute intensities, Time-Intensity or Multi-Attribute Time-Intensity (Kuesten et al., 2013) should be the right tool. If the differences between products are mainly driven by the presence or absence of several attributes, TCATA can be considered. If tracking major changes in perception is good enough, TDS can be a reasonable choice. If the objective is to get essential information to communicate about products (for example on back labels or sensory claims), then AEF with consumers can be an alternative.

Then, a second question could be "Is the method I want to use at the expected level of precision?". Based on this preliminary study, it is not clear whether additional noise is more important than additional information, as considering product temporality in only 3 periods does not seem to result in a severe loss of information. In other words, this could suggest that it was just about the right level of precision which can be expected of temporal data collected with consumers on this type of products. Of course, all the previous elements of this discussion should be kept in mind when comparing the results of two methods having observed differences of the same order of magnitude as those potentially due to imprecision. Investigating this expected level of precision would definitely be of great interest. It would require studying other criteria, such as repeatability or at least reproducibility, to either complete and validate or invalidate the conclusions based on this study. Meta-analysis based on several datasets or simulations could also help to draw general conclusions.

5. Conclusion

This article introduced AEF, a new retrospective method for temporal data collection inspired from Temporal Order of Sensations, Pick 3 and Rank, and wine evaluation in 3 phases. AEF was based on a concept close to dominance and allowed consumers to select 3 descriptors, one per period (attack, evolution, final), taking into account the entire tasting. As AEF data analysis does not consider durations, the consumers' heterogeneity in terms of delays and durations is no longer a problem, and no data transformation is required. As every consumer had to give only a sequence of 3 descriptors, no briefing was required, and the consumers' heterogeneity due to differences in the number of citations was also avoided. Retrospective measurement was proven feasible, and opened new perspectives. This makes AEF a method to consider for capturing the big picture of temporal descriptions, especially for at-home studies with consumers. Finally, the discussion encouraged the readers to think about the granularity of temporal

sensory data. Indeed, in this study, AEF and TDS discriminated the products in a very similar way. This suggests that considering temporality in only 3 periods could be the right level of precision which can be expected of temporal data collected with consumers, at least for this type of products.

Acknowledgments

Barry Callebaut for providing the chocolate samples.

Betty Hoffarth for recruiting the consumers.

Catherine Pedron for preparing the samples.

Anne-Laure Loiseau and Odile Brossard for hosting the consumers.

Christine Lange for participating in the implementation of the protocol.

The reviewers for their significant help in improving the manuscript.

The region Bourgogne-Franche-Comté and the company SensoStat for the funding of the Ph.D. of Benjamin Mahieu.

References

Agresti, A. (2013). *Categorical data analysis* (3rd ed.), John Wiley & Sons, Inc., New Jersey.

Albert, A., Salvador, A., Schlich, P., & Fiszman, S. (2012). Comparison between temporal dominance of sensations (TDS) and key-attribute sensory profiling for evaluating solid food with contrasting textural layers: Fish sticks. *Food Quality and Preference*, *24*(1), 111–118.

Ares, G., Castura, J. C., Antúnez, L., Vidal, L., Giménez, A., Coste, B., Picallo, A., Beresford, M. K., Chheang, S. L., & Jaeger, S. R. (2016). Comparison of two TCATA variants for dynamic sensory characterization of food products. *Food Quality and Preference*, *54*, 160–172.

Ares, G., Jaeger, S. R., Antúnez, L., Vidal, L., Giménez, A., Coste, B., Picallo, A., & Castura, J. C. (2015). Comparison of TCATA and TDS for dynamic sensory characterization of food products. *Food Research International*, *78*, 148–158.

Berget, I., Castura, J. C., Ares, G., Næs, T., & Varela, P. (2020). Exploring the common and unique variability in TDS and TCATA data – A comparison using canonical correlation and orthogonalization. *Food Quality and Preference*, *79*, 103790.

Cardot, H., Lecuelle, G., Schlich, P., & Visalli, M. (2019). Estimating finite mixtures of semi‐Markov chains: an application to the segmentation of temporal sensory data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *68*(5), 1281–1303.

Cerri, J.; Thøgersen, J., Testa, F. (2019)? Social desirability and sustainable food research: a systematic literature review. *Food Quality and Preference, 71*, 2019, 136-140.

Castura, J. C., Antúnez, L., Giménez, A., & Ares, G. (2016). Temporal Check-All-That-Apply (TCATA): A novel dynamic method for characterizing products. *Food Quality and Preference*, *47*, 79–90.

Di Monaco, R., Su, C., Masi, P., & Cavella, S. (2014). Temporal Dominance of Sensations: A review. *Trends in Food Science & Technology*, *38*(2), 104–112.

Dinnella, C., Masi, C., Naes, T., & Monteleone, E. (2013). A new approach in TDS data analysis: A case study on sweetened coffee. *Food Quality and Preference*, *30*(1), 33–46.

Dinnella, C., Masi, C., Zoboli, G., & Monteleone, E. (2012). Sensory functionality of extra-virgin olive oil in vegetable foods assessed by Temporal Dominance of Sensations and Descriptive Analysis. *Food Quality and Preference*, *26*(2), 141–150.

Dunn, P. K., Smyth, G. K. (1996). Randomized Quantile Residuals. *Journal of Computational and Graphical Statistics,* 5(3), 236-244.

Esmerino, E. A., Castura, J. C., Ferraz, J. P., Tavares Filho, E. R., Silva, R., Cruz, A. G., Freitas, M. Q., & Bolini, H. M. A. (2017). Dynamic profiling of different ready-to-drink fermented dairy products: A comparative study using Temporal Check-All-That-Apply (TCATA), Temporal Dominance of Sensations (TDS) and Progressive Profile (PP). *Food Research International*, *101*, 249–258.

Frost, S. C., Blackman, J. W., Ebeler, S. E., & Heymann, H. (2018). Analysis of temporal dominance of sensation data using correspondence analysis on Merlot wine with differing maceration and cap management regimes. *Food Quality and Preference*, *64*, 245–252.

Galmarini, M. V., Visalli, M., & Schlich, P. (2017). Advances in representation and analysis of mono and multi-intake Temporal Dominance of Sensations data. *Food Quality and Preference*, *56*, 247–255.

Goodman, L. A. (1965). On Simultaneous Confidence Intervals for Multinomial Proportions. *Technometrics*, *7*(2), 247–254.

Grainger, K. (2009). *Wine Quality : Tasting and Selection*. Food industry briefing seeries. Chichester, UK, and MAes, IA: Wiley-Blackwell.

Harrington, R. J. (2008). *Food and wine pairing : a sensory experience*. John Wiley & Sons Inc, United States.

Hutchings, S. C., Foster, K. D., Grigor, J. M. V., Bronlund, J. E., & Morgenstern, M. P. (2014). Temporal dominance of sensations: A comparison between younger and older subjects for the perception of food texture. *Food Quality and Preference*, *31*, 106–115.

ISO, 2016. ISO. Sensory analysis — Methodology — General guidance for establishing a sensory profile, ISO standard 13299:2016(E). International Organization for Standardization, Geneva, Switzerland (2016).

Jack, F. R., Piggott, J. R., & Paterson, A. (1994). Analysis of Textural Changes in Hard Cheese during Mastication by Progressive Profiling. *Journal of Food Science*, *59*(3), 539–543.

Jaeger, S.R., Hort, J., Porcherot, C., Ares, G., Pecore, S., & MacFie, H. J. H. (2017). Future directions in sensory and consumer science: Four perspectives and audience voting. *Food Quality and Preference*, *56*, 301–309.

Jaeger, Sara R., Alcaire, F., Hunter, D. C., Jin, D., Castura, J. C., & Ares, G. (2018). Number of terms to use in temporal check-all-that-apply studies (TCATA and TCATA Fading) for sensory product characterization by consumers. *Food Quality and Preference*, *64*, 154–159.

Jaeger, Sara R., Beresford, M. K., Hunter, D. C., Alcaire, F., Castura, J. C., & Ares, G. (2017). Does a familiarization step influence results from a TCATA task? *Food Quality and Preference*, *55*, 91–97.

Jaeger, Sara R, & Porcherot, C. (2017). Consumption context in consumer research: methodological perspectives. *Current Opinion in Food Science*, *15*, 30–37.

Kawasaki, H., Yoshimura, W., Wakita, A., & Kasamatsu, C. (2019). Visualization of temporal differences between dominant perceptions in temporal dominance of sensations (TDS) and temporal check-all-that-apply (TCATA) perceptions using dominance-highlighted TCATA (dTCATA) curves. *Food Quality and Preference*, *77*, 166–171.

Kemp, S. E., Hort, J., & Hollowood, T. (2017). Time-Dependent Measures of Perception in Sensory Evaluation. In J. Hort, S. E. Kemp, & T. Hollowood (Eds.), *Time-Dependent Measures of Perception in Sensory Evaluation (pp. 1-23)*. John Wiley & Sons Ltd, Chichester, UK.

Kuesten, C., Bi, J., Feng, Y. (2013). Exploring taffy product consumption experiences using a multi-attribute time–intensity (MATI) method, *Food Quality and Preference, 30(2)*, 260-273,

Lecuelle, G., Visalli, M., Cardot, H., & Schlich, P. (2018). Modeling Temporal Dominance of Sensations with semi-Markov chains. *Food Quality and Preference*, *67*.

Lee, W. E. I., & Pangborn, M. (1986). Time–intensity: The temporal aspects of sensory perception. *Food Technology*,

566 *40*, 78–82.

567 Lenfant, F., Loret, C., Pineau, N., Hartmann, C., & Martin, N. (2009). Perception of oral food breakdown. The concept
568 of sensory trajectory. *Appetite*, *52*(3), 659–667.

569 Lepage, M., Neville, T., Rytz, A., Schlich, P., Martin, N., & Pineau, N. (2014). Panel performance for Temporal
570 Dominance of Sensations. *Food Quality and Preference*, *38*, 24–29.

571 Likert, R. A. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *40*, 5–53.

572 Meiselman, H. L. (2013). The future in sensory/consumer research: ……..evolving to a better science. *Food Quality*
573 *and Preference*, *27*(2), 208–214.

574 Mesurolle, J., Saint-Eve, A., Déléris, I., & Souchon, I. (2013). Impact of fruit piece structure in yogurts on the dynamics
575 of aroma release and sensory perception. *Molecules*, *18*(5), 6035–6056.

576 Methven, L., Rahelu, K., Economou, N., Kinneavy, L., Ladbrooke-Davis, L., Kennedy, O. B., Mottram, D. S., & Gosney,
577 M. A. (2010). The effect of consumption volume on profile and liking of oral nutritional supplements of varied
578 sweetness: Sequential profiling and boredom tests. *Food Quality and Preference*, *21*(8), 948–955.

579 Meyners, M, & Castura, J. (2018). The analysis of temporal check-all-that-apply (TCATA) data. *Food Quality and*
580 *Preference*, *67*, 67–76.

581 Meyners, Michael. (2020). Temporal methods: Are we comparing apples and oranges? *Food Quality and Preference*,
582 79, 103615.

583 Meyners, Michael, & Castura, J. C. (2019). Did assessors select attributes by chance alone in your TDS study, and how
584 relevant is it to know? *Food Research International*, *119*, 571–583.

585 Nguyen, Q. C., Næs, T., & Varela, P. (2018). When the choice of the temporal method does make a difference:
586 TCATA, TDS and TDS by modality for characterizing semi-solid foods. *Food Quality and Preference*, *66*, 95–106.

587 Osterland, E. A. (2012). *Power entertaining : secrets to building lasting relationships, hosting unforgettable events,*
588 *and closing big deals from America's 1st master sommelier*. John Wiley & Son, US.

589 Pecore, S., Rathjen-Nowak, C., Tamminen, T. (2011). Temporal Order of Sensations. *9th Pangborn Sensory Science*
590 *Symposium, 4–8 September 2011, Toronto, Canada*.

591 Peltier, C., Visalli, M., Thomas, A. (2019). Using temporal dominance of emotions at home. Impact of coffee
592 advertisements on consumers' behavior and methodological perspectives. *Food Quality and Preferenc*e, 71,
593 311-319.

594 Pineau, N., Schlich, P., Cordelle, S., Mathonnière, C., Issanchou, S., Imbert, A., Rogeaux, M., Etiévant, P., & Köster, E.
595 (2009). Temporal Dominance of Sensations: Construction of the TDS curves and comparison with time–
596 intensity. *Food Quality and Preference*, *20*(6), 450–455.

597 Pineau, Nicolas, de Bouillé, A. G., Lepage, M., Lenfant, F., Schlich, P., Martin, N., & Rytz, A. (2012). Temporal
598 Dominance of Sensations: What is a good attribute list? *Food Quality and Preference*, *26*(2), 159–165.

599 R Core team. (2017). *R: A language and environment for statistical computing. R Foundation for Statistical*
600 *Computing, Vienna, Austria.* url: http://www.R-project.org.

601 Rizo, A., Vidák, K., Fiszman, S., & Tarrega, A. (2020). Influence of fading duration on TCATA evaluation. *Food Quality*
602 *and Preference*, *79*, 103619.

603 Saporta, G., & Hatabian, G. (1986). Régions de confiance en analyse factorielle. *E. D., et Al. (Eds.), Data Analysis and*
604 *Informatics, Elsevier Science Publisher B.V. (1986)*.

605 Schlich, P. (2017). Temporal Dominance of Sensations (TDS): a new deal for temporal sensory analysis. *Current*
606 *Opinion in Food Science*, *15*, 38–42.

607 Schwarz, N. (2012). Retrospective and Concurrent Self-Reports: The Rationale for Real-Time Data Capture. In Mehl &

608 Connor (Eds.), *Handbook of Research Methods for Studying Daily Life.* Guilford.

609 Spence, C., & Wang, Q. J. (2018). What does the term 'complexity' mean in the world of wine? *International Journal*
610 *of Gastronomy and Food Science*, *14*, 45–54.

611 Stone, A. A., Shiffman, S. S., & DeVries, M. W. (1999). Ecological momentary assessment. In D. Kahneman, E. Diener,
612 & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (p. 26–39). Russell Sage Foundation.

613 Thomas, A., Visalli, M., Cordelle, S., & Schlich, P. (2015). Temporal Drivers of Liking. *Food Quality and Preference*, *40*,
614 365–375.

615 Torres-Moreno, M., Hort, J., Tarrega, A. (2016). The challenges of evaluating temporal sensations in beverages:
616 Performance of TDS, TCATA and TOS in multiple-sip experiments. *Presented at 7th European Conference on*
617 *Sensory and Consumer Research (Eurosense), Dijon, France, September 2016*.

618 Urbano, C., Mahieu, B., Thomas, A., Schlich, P., Visalli, M. (2018). Recording facial mimics and temporal dominance of
619 sensations and emotions. *8th European Conference on Sensory and Consumer Research, Verona, Italy,*
620 *September 2018.*

621 van Bommel, R., Stieger, M., Schlich, P., & Jager, G. (2019). Dutch consumers do not hesitate: Capturing implicit 'no
622 dominance' durations using Hold-down Temporal Dominance methodologies for Sensations (TDS) and
623 Emotions (TDE). *Food Quality and Preference*, *71*, 332–342.

624 Vandeputte, A., Romans, J., Lenfant, F., Pineau, N. (2011). Innovative methods to assess the evolution of the sensory
625 characteristics during the tasting of a full product portion (several bites). *9th Pangborn Sensory Science*
626 *Symposium, Toronto, Canada, 4-8th September 2011*.

627 Varela, P., Antúnez, L., Carlehög, M., Alcaire, F., Castura, J. C., Berget, I., Giménez, A., Næs, T., & Ares, G. (2018).
628 What is dominance? An exploration of the concept in TDS tests with trained assessors and consumers. *Food*
629 *Quality and Preference*, *64*, 72–81.

630 Visalli, M., Lange, C., Mallet, L., Cordelle, S., & Schlich, P. (2016). Should I use touchscreen tablets rather than
631 computers and mice in TDS trials? *Food Quality and Preference*, *52*.

632 Warton, D.I., Lyons, M., Stoklosa, J. and Ives, A.R. (2016). Three points to consider when choosing a LM or GLM test
633 for count data. *Methods in Ecology and Evolution*, 7, 882-890.

START

Dry	Sweet	Cocoa
Melting	Sour	Fruity
Fat	Bitter	Floral
Sticky	Astringent	Woody

STOP

What sensations did you perceive during the tasting, in chronological order?

(Click on the drop-down lists to answer)

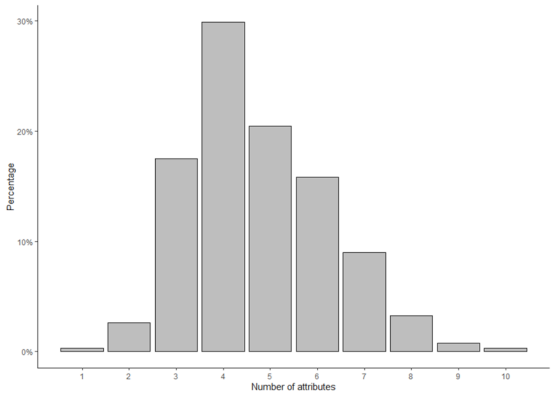
At first, I perceived this chocolate

then after a few moments I perceived it

and at the end of the tasting I perceived it

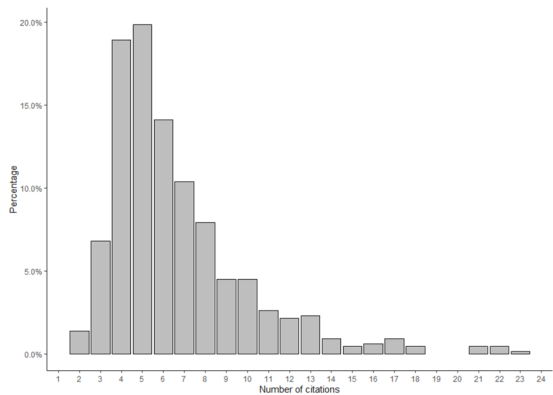
NEXT

A



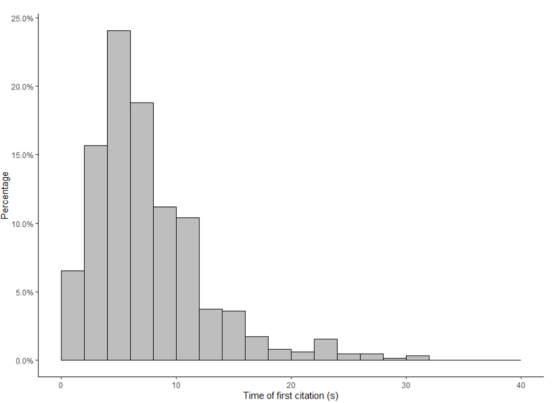
Mode: 4, Median: 4, Mean: 4.7, Sd: 1.5, CV: 32%

B



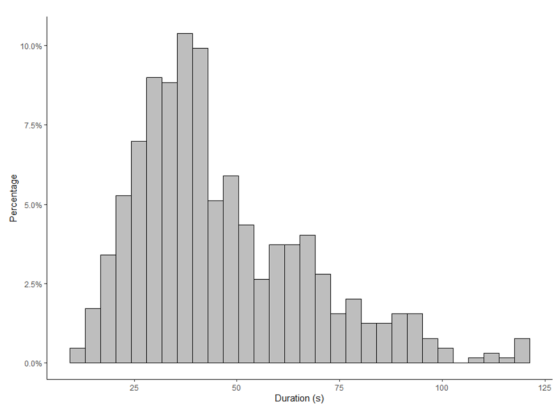
Mode: 5, Median: 5, Mean: 6.7, Sd: 3.4, CV: 51%

C



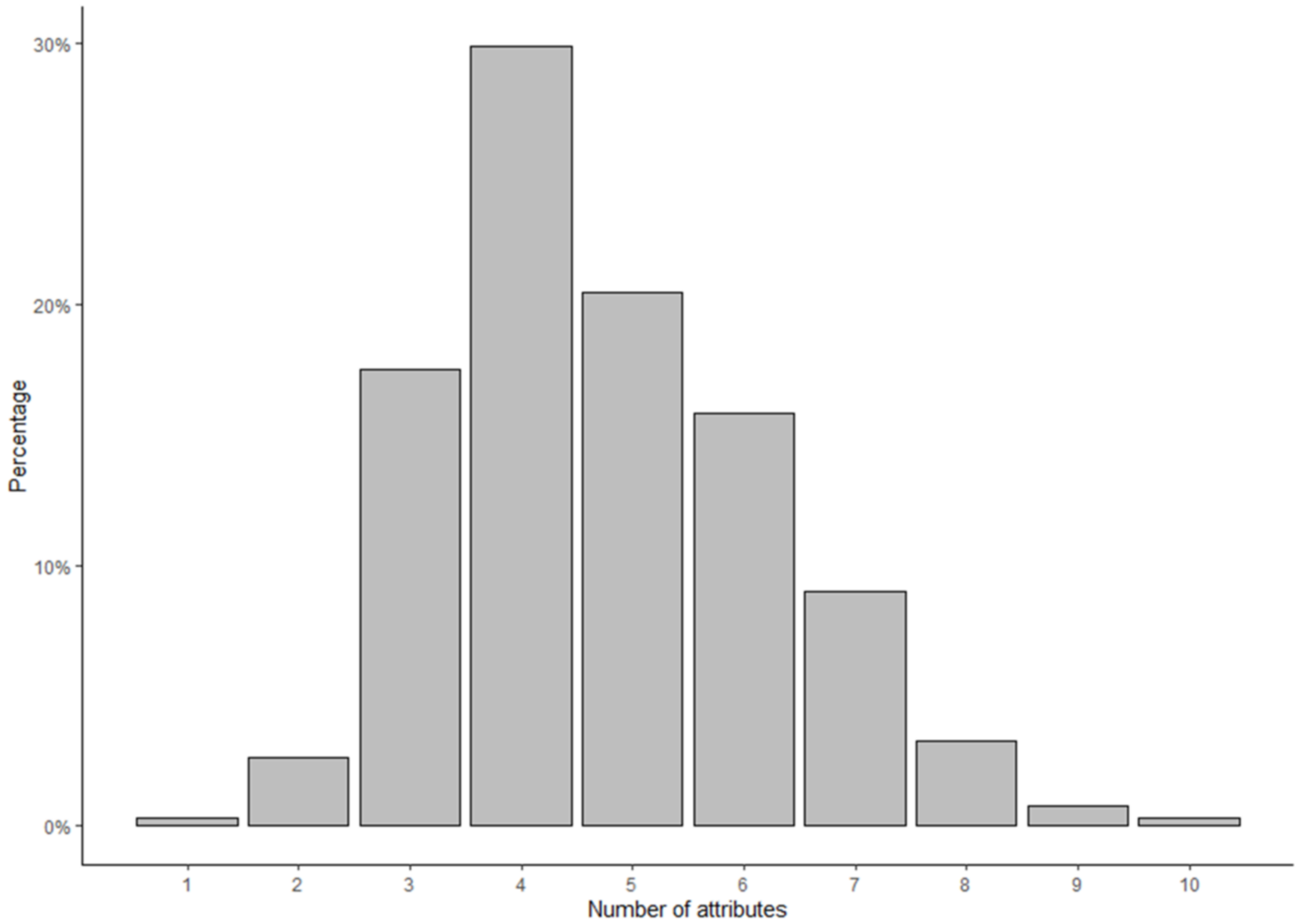
Mode:5, Median: 6.3, Mean: 7.7, Sd: 6.4, CV: 83%

D



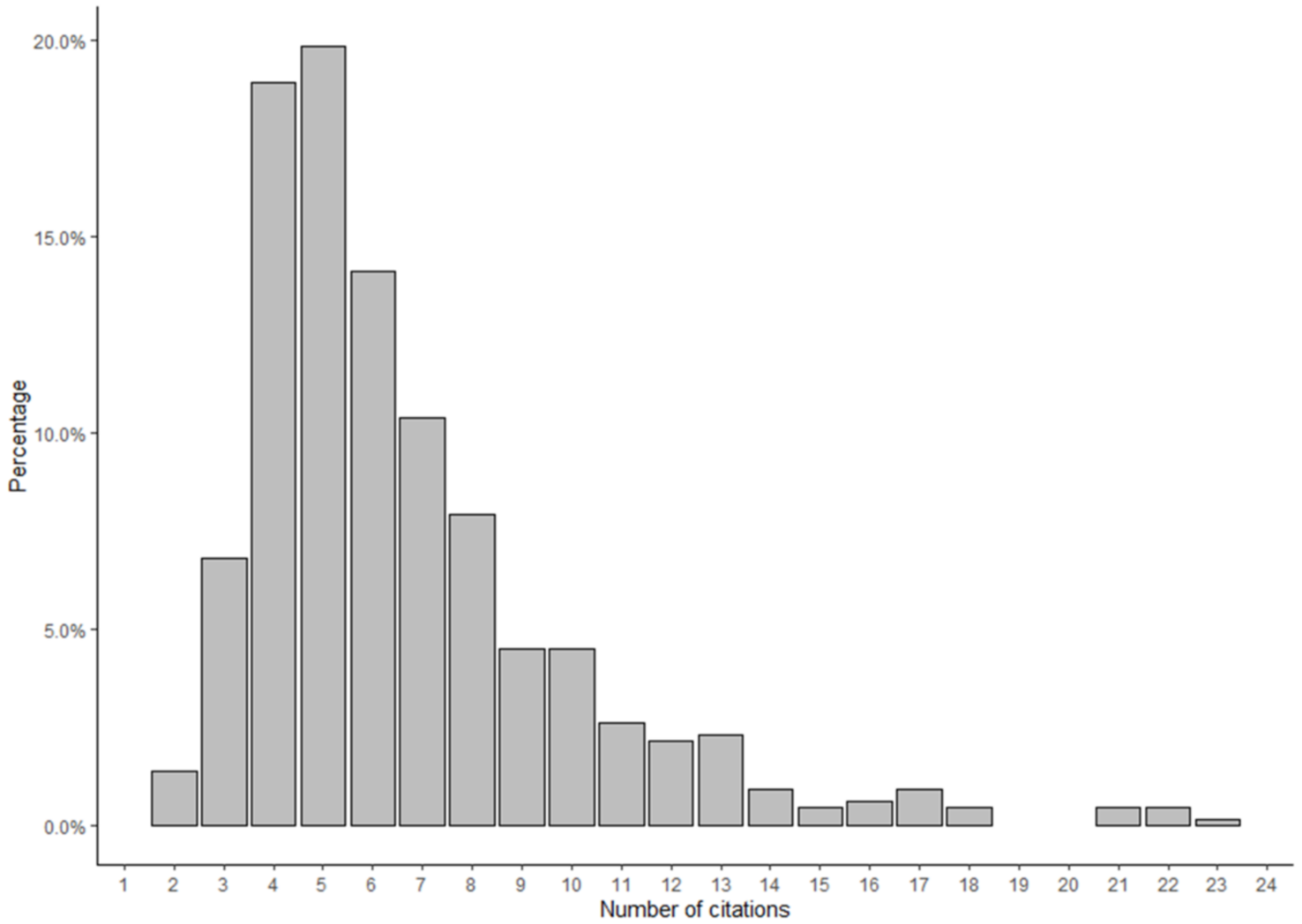
Mode: 40, Median: 40.7, Mean: 46.2, Sd: 21.1, CV: 46%

A



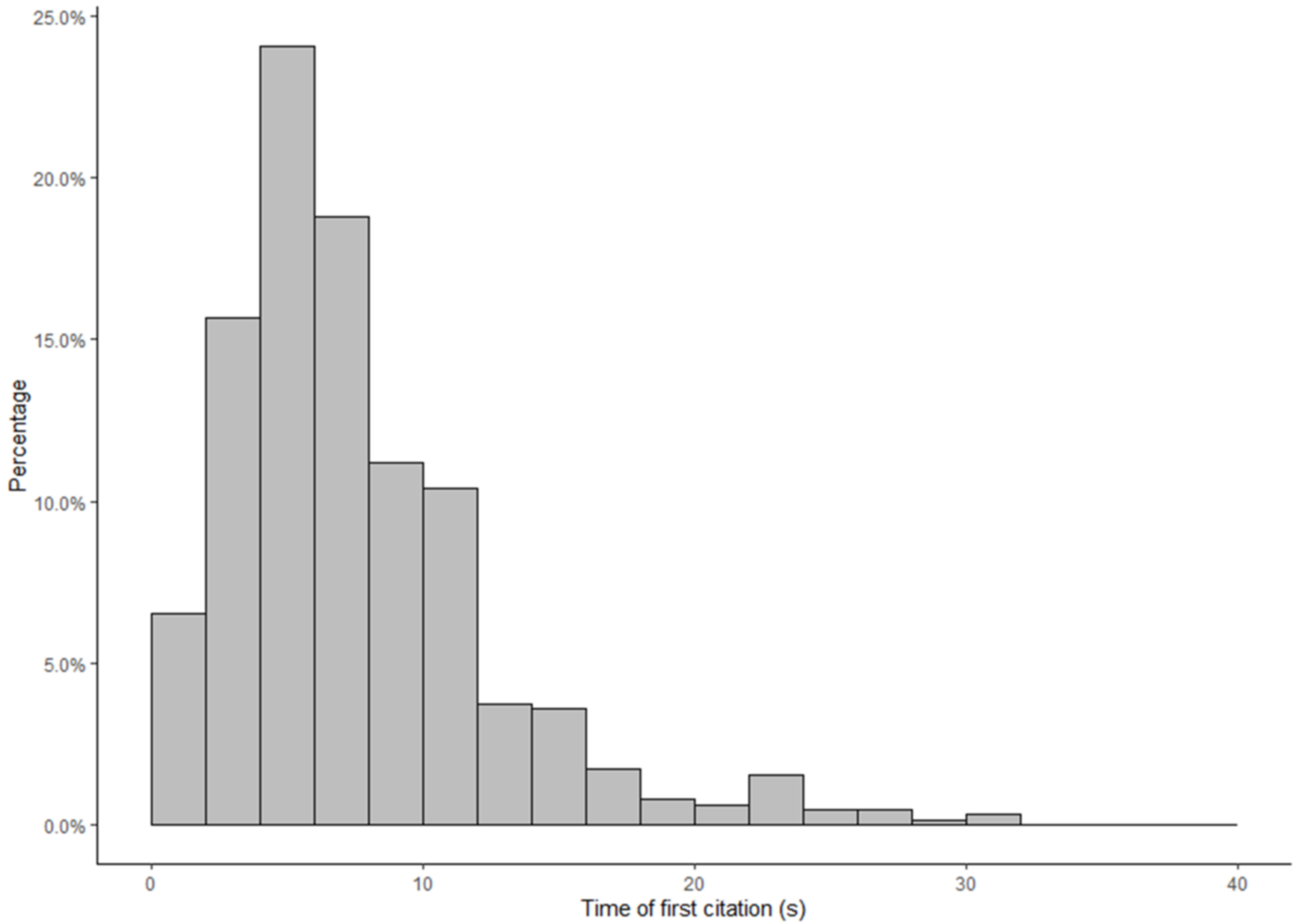
Mode: 4, Median: 4, Mean: 4.7, Sd: 1.5, CV: 32%

B



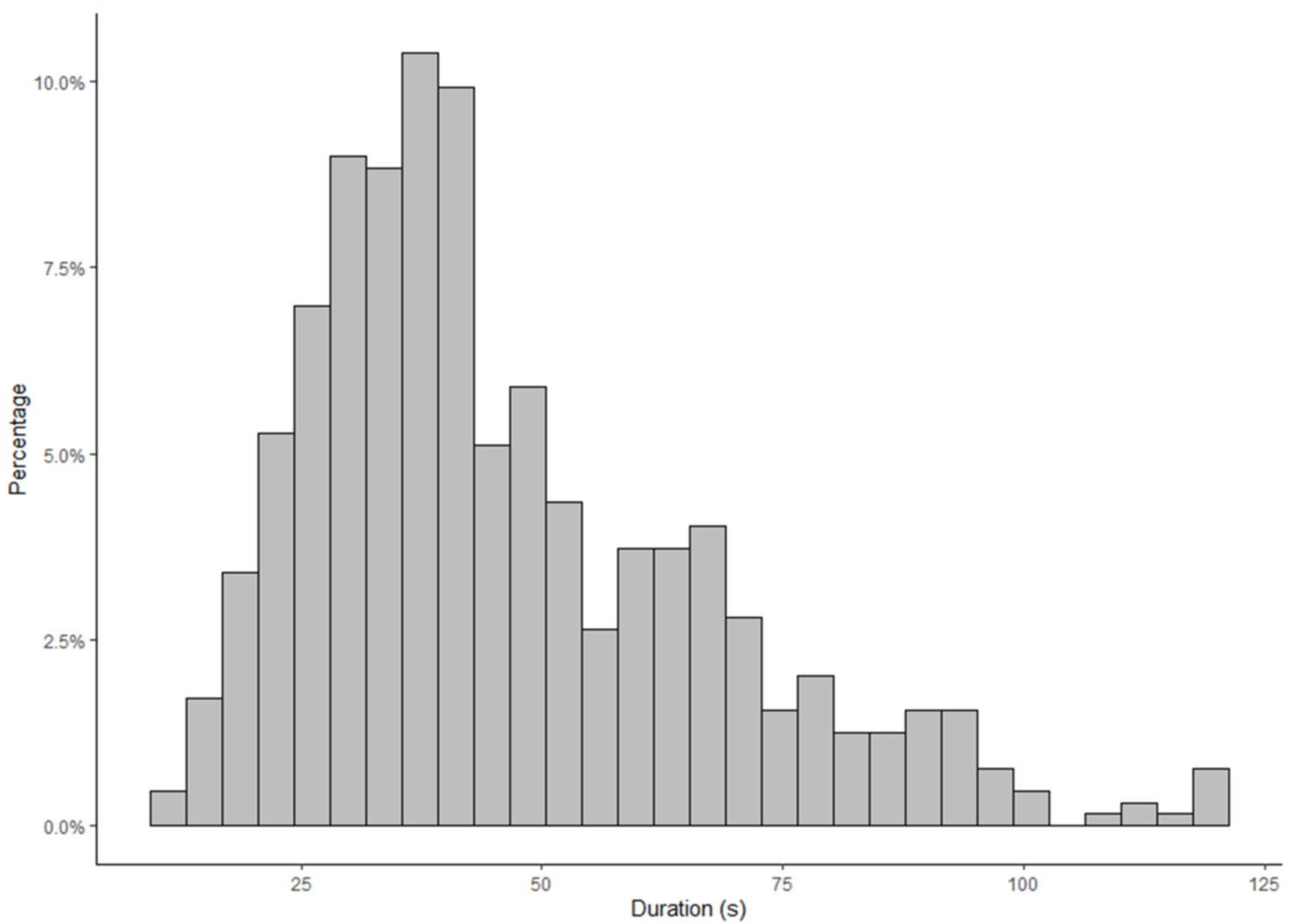
Mode: 5, Median: 5, Mean: 6.7, Sd: 3.4, CV: 51%

C



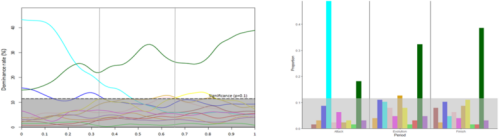
Mode:5, Median: 6.3, Mean: 7.7, Sd: 6.4, CV: 83%

D



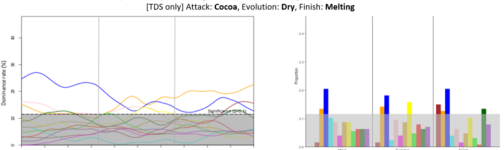
Mode: 40, Median: 40.7, Mean: 46.2, Sd: 21.1, CV: 46%

C54



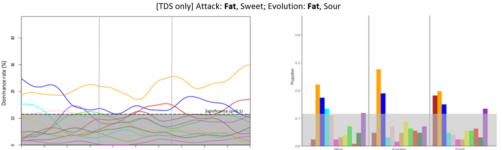
[TDS & AEF] Attack: **Dry**, **Sweet**; Evolution: Fruity, **Sweet**; Finish: **Sweet**
[TDS only] Attack: **Cocoa**, Evolution: **Dry**, Finish: **Melting**

C65



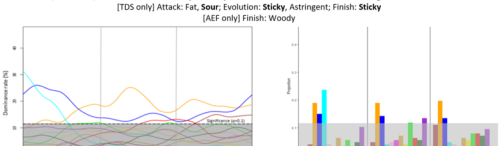
[TDS & AEF] Attack: **Bitter**[J], **Cocoa**; Evolution: **Bitter**[T], **Cocoa**, **Melting**; Finish: **Bitter**[T], **Cocoa**, **Astringent**[4], Sweet
[TDS only] Attack: **Fat**, Sweet; Evolution: **Fat**, Sour

C68



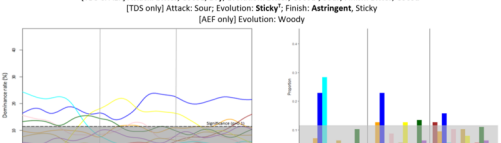
[TDS & AEF] Attack: **Bitter**, **Cocoa**, **Dry**[T], **Woody**[T]; Evolution: **Bitter**, **Cocoa**; Finish: **Astringent**, **Bitter**, **Cocoa**
[TDS only] Attack: Fat, **Sour**; Evolution: **Sticky**, Astringent; Finish: **Sticky**
[AEF only] Finish: Woody

C70



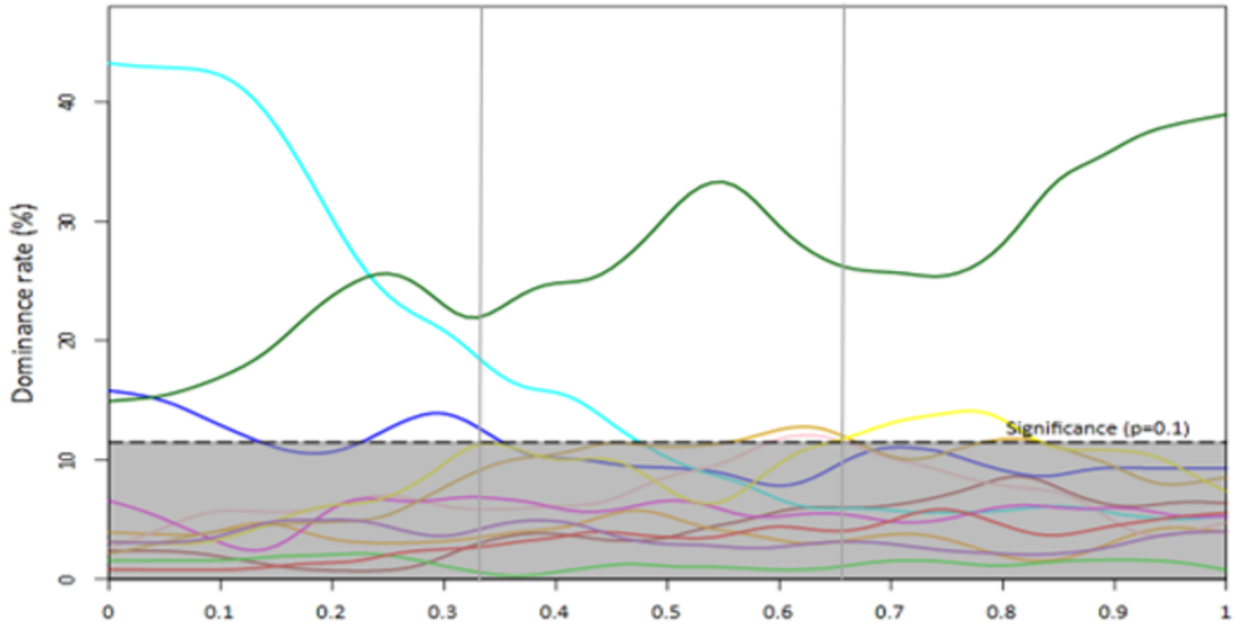
[TDS & AEF] Attack: **Bitter**, **Cocoa**, **Dry**; Evolution: **Bitter**, **Cocoa**[T], Sour; Finish: **Bitter**, **Cocoa**[T]
[TDS only] Attack: Sour; Evolution: **Sticky**[T]; Finish: **Astringent**, Sticky
[AEF only] Evolution: Woody

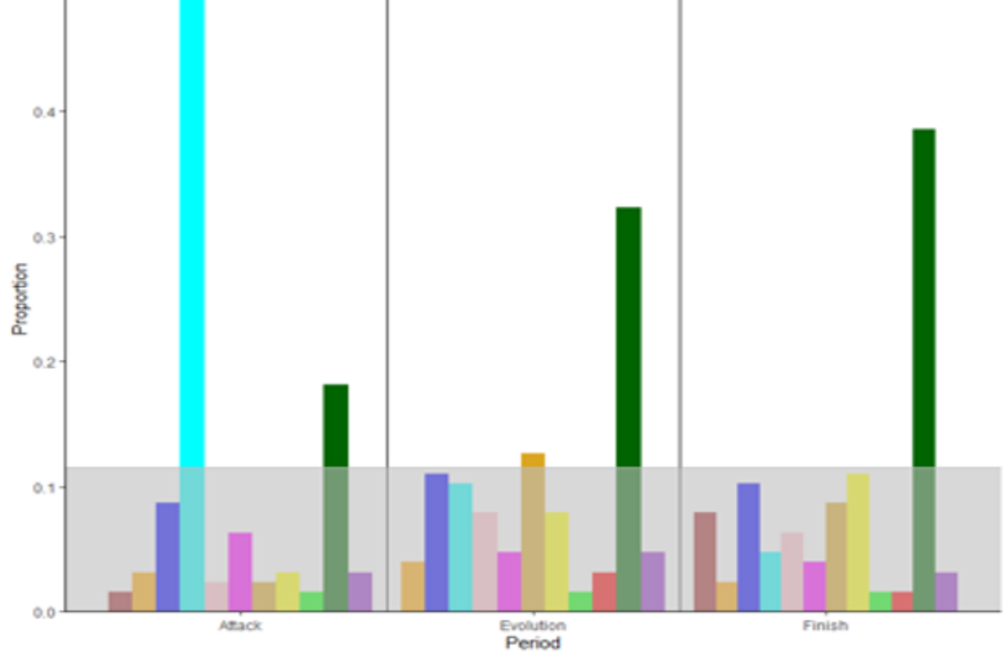
C73



[TDS & AEF] Attack: **Cocoa**, **Dry**; Evolution: Bitter, **Cocoa**, **Melting**[T], **Sweet**[T], Finish: **Astringent**[T], **Cocoa**
[TDS only] Attack: **Fat**, **Melting**, **Sweet**; Evolution: Fat, Finish: **Bitter**, Melting

C
5
4

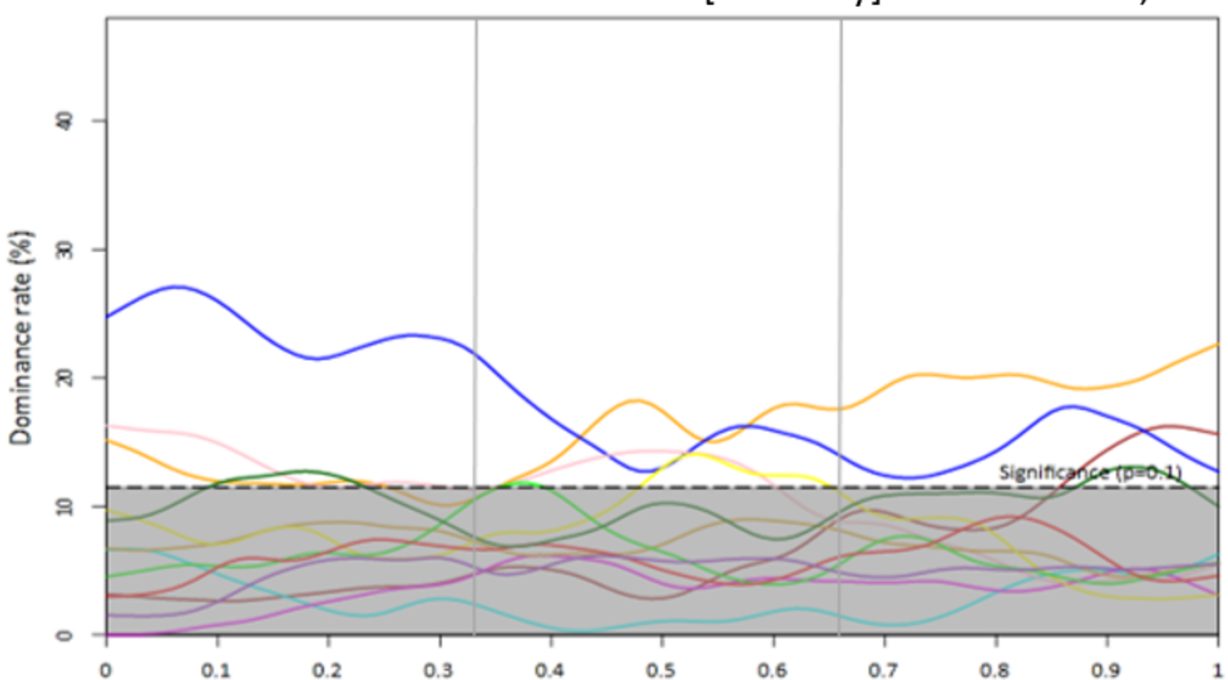


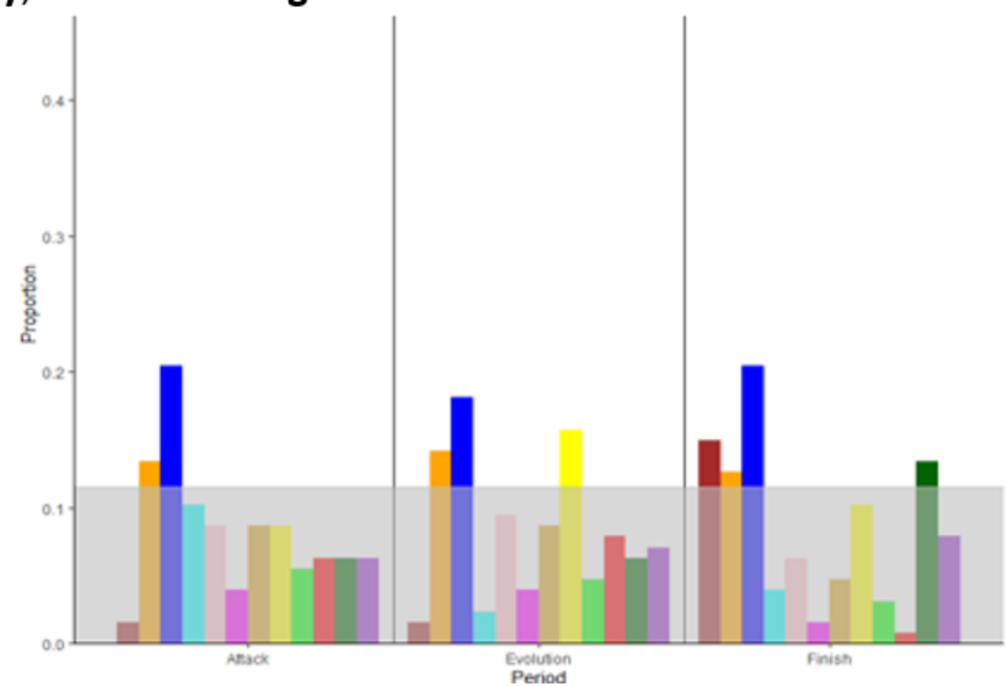


[TDS & AEF] Attack: **Dry**, **Sweet**; Evolution: Fruity, **Sweet**; Finish: **Sweet**

[TDS only] Attack: **Cocoa**, Evolution: **Dry**, Finish: **Melting**

C
6
5

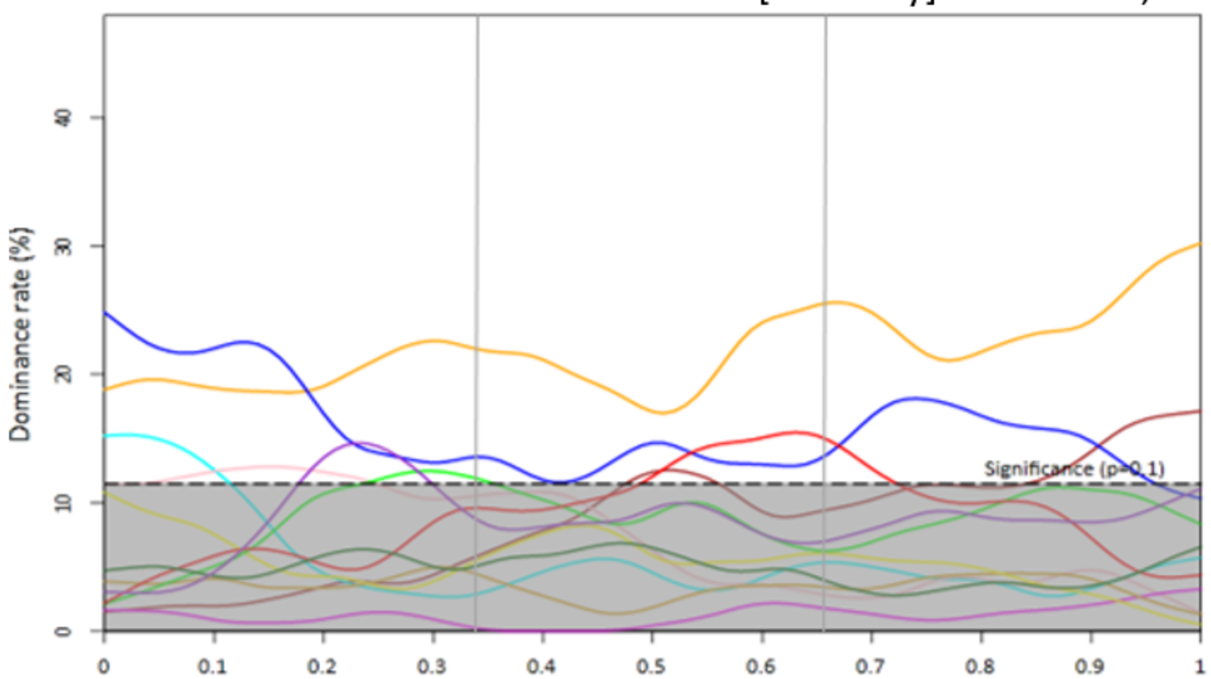


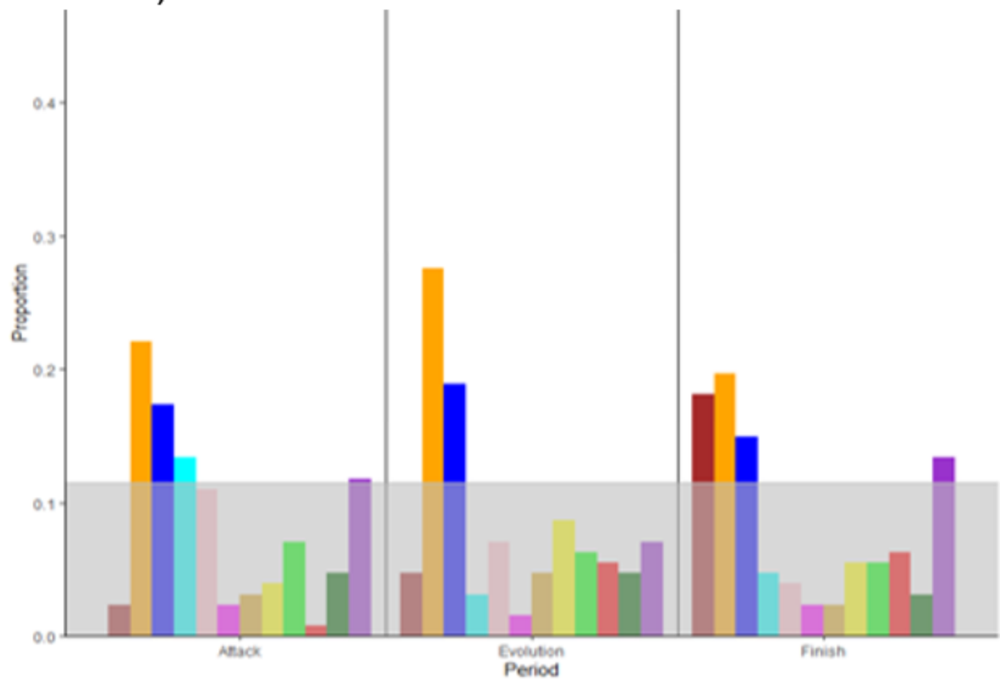


[TDS & AEF] Attack: **Bitter**^{T}, **Cocoa**; Evolution: **Bitter**^{T}, **Cocoa**, **Melting**; Finish: **Bitter**^{T}, **Cocoa**, **Astringent**^{A}, Sweet

[TDS only] Attack: **Fat**, Sweet; Evolution: **Fat**, Sour

C
6
8



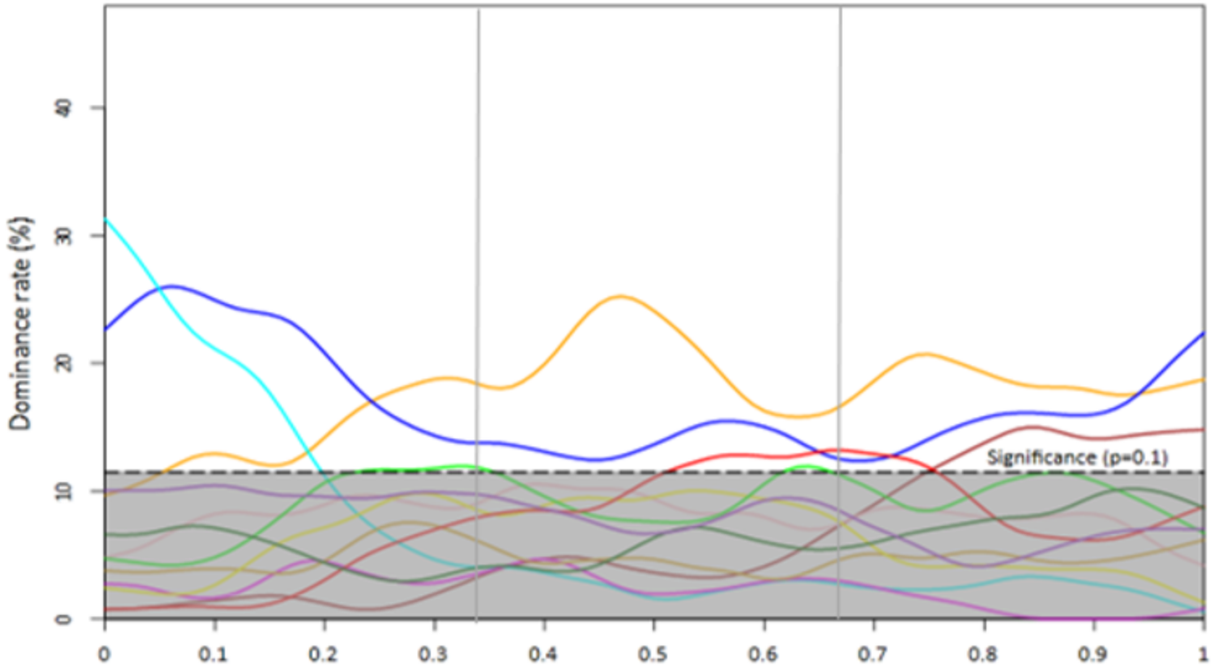


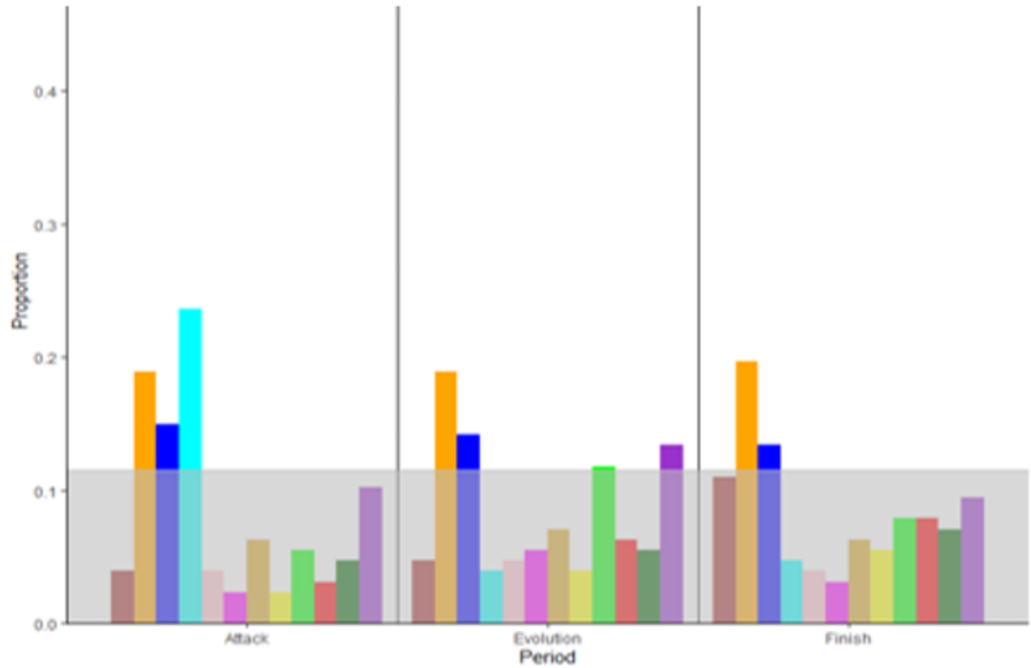
[TDS & AEF] Attack: **Bitter**, **Cocoa**, **Dry**^{T}, **Woody**^{T}; Evolution: **Bitter**, **Cocoa**; Finish: **Astringent**, **Bitter**, **Cocoa**

[TDS only] Attack: Fat, **Sour**; Evolution: **Sticky**, Astringent; Finish: **Sticky**

[AEF only] Finish: Woody

C
7
0



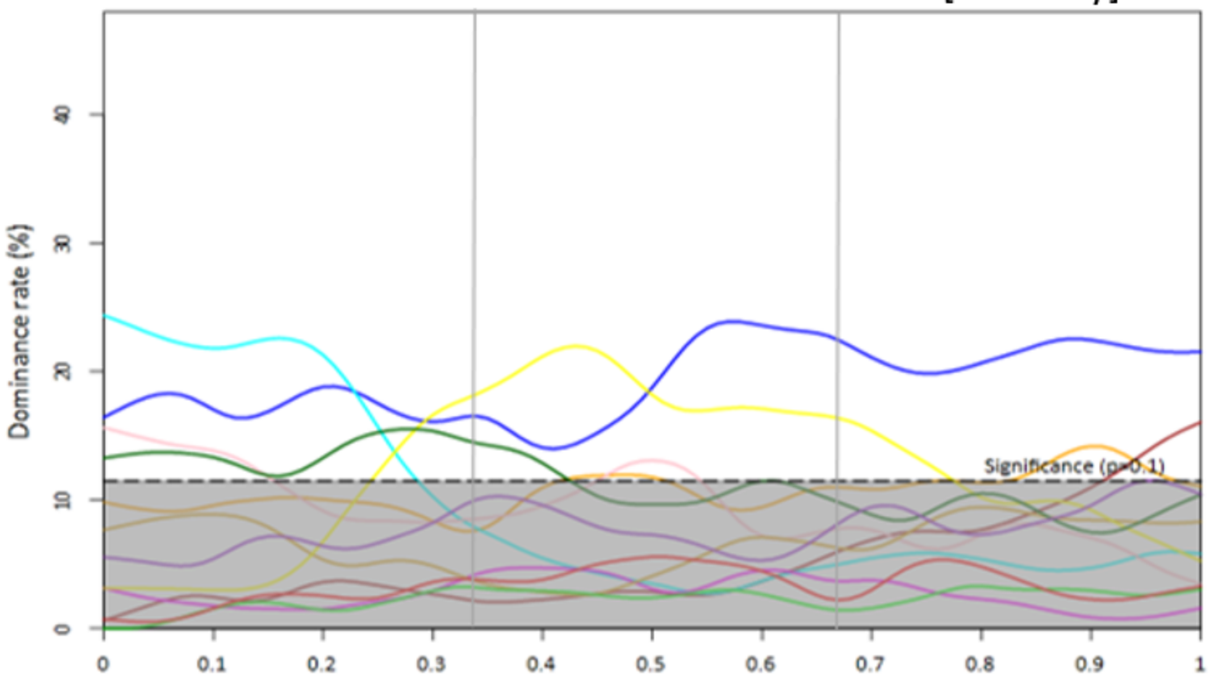


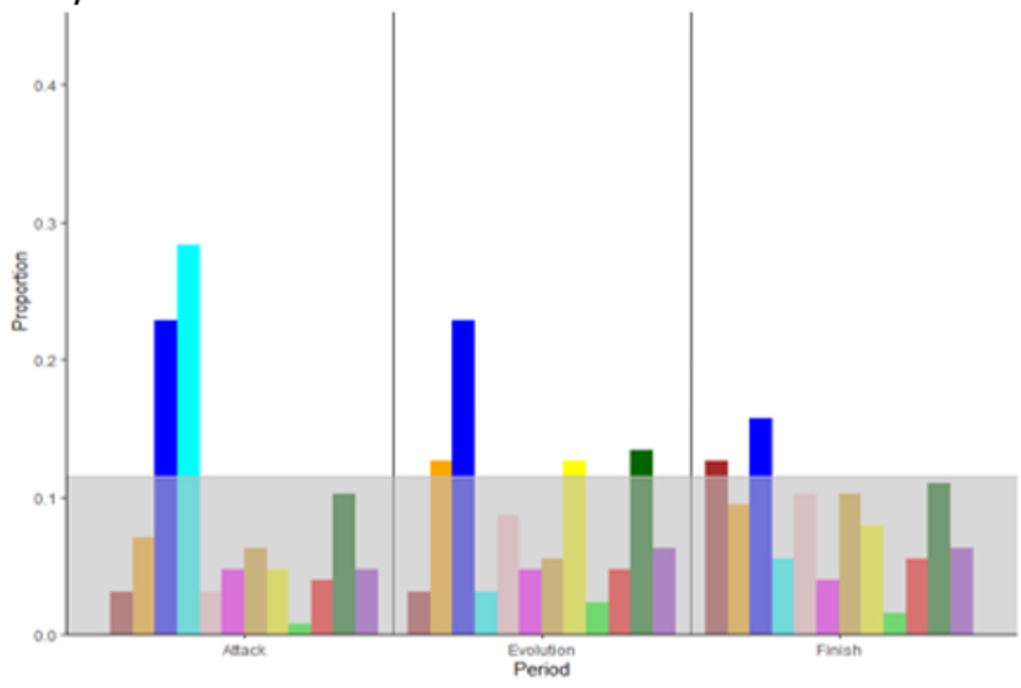
[TDS & AEF] Attack: **Bitter**, **Cocoa**, **Dry**; Evolution: **Bitter**, **Cocoa**^{T}, Sour; Finish: **Bitter**, **Cocoa**^{T}

[TDS only] Attack: Sour; Evolution: **Sticky**^{T}; Finish: **Astringent**, Sticky

[AEF only] Evolution: Woody

C
7
3

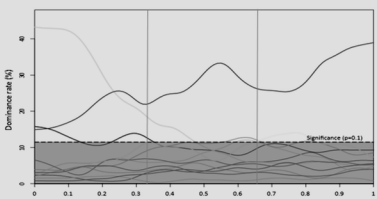


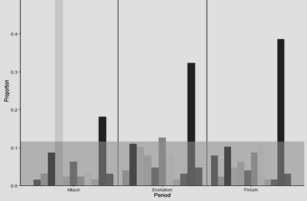


[TDS & AEF] Attack: **Cocoa**, **Dry**; Evolution: Bitter, **Cocoa**, **Melting**^{T}, **Sweet**^{T}; Finish: **Astringent**^{T}, **Cocoa**

[TDS only] Attack: **Fat**, **Melting**, **Sweet**; Evolution: Fat, Finish: **Bitter**, Melting

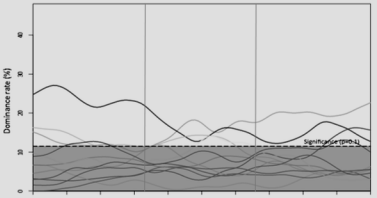
C54

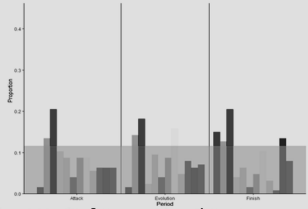




[TDS & AEF] Attack: **Dry, Sweet**; Evolution: Fruity, **Sweet**; Finish: **Sweet**
[TDS only] Attack: **Cocoa**, Evolution: **Dry**, Finish: **Melting**

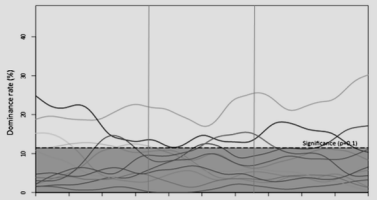
C65

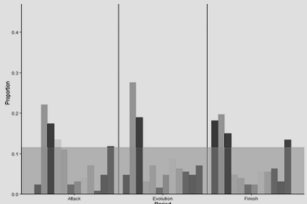




[TDS & AEF] Attack: **Bitter**^{T}, **Cocoa**; Evolution: **Bitter**^{T}, **Cocoa**, **Melting**; Finish: **Bitter**^{T}, **Cocoa**, **Astringent**^{A}, Sweet
[TDS only] Attack: **Fat**, Sweet; Evolution: **Fat**, Sour

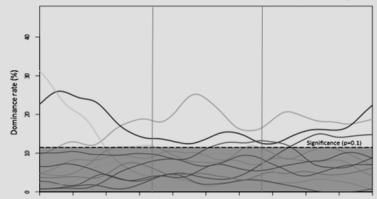
C68

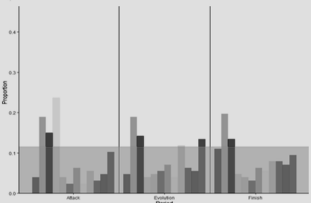




[TDS & AEF] Attack: **Bitter**, **Cocoa**, **Dry**^{T}, **Woody**^{T}; Evolution: **Bitter**, **Cocoa**; Finish: **Astringent**, **Bitter**, **Cocoa**
[TDS only] Attack: Fat, **Sour**; Evolution: **Sticky**, Astringent; Finish: **Sticky**
[AEF only] Finish: Woody

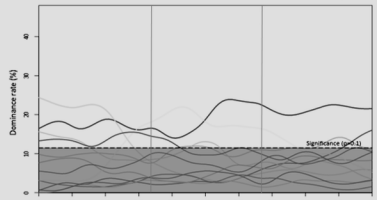
C70

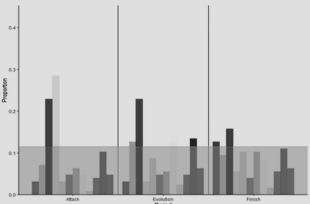




[TDS & AEF] Attack: **Bitter**, **Cocoa**, **Dry**; Evolution: **Bitter**, **Cocoa**^{T}, Sour; Finish: **Bitter**, **Cocoa**^{T}
[TDS only] Attack: Sour; Evolution: **Sticky**^{T}; Finish: **Astringent**, Sticky
[AEF only] Evolution: Woody

C73



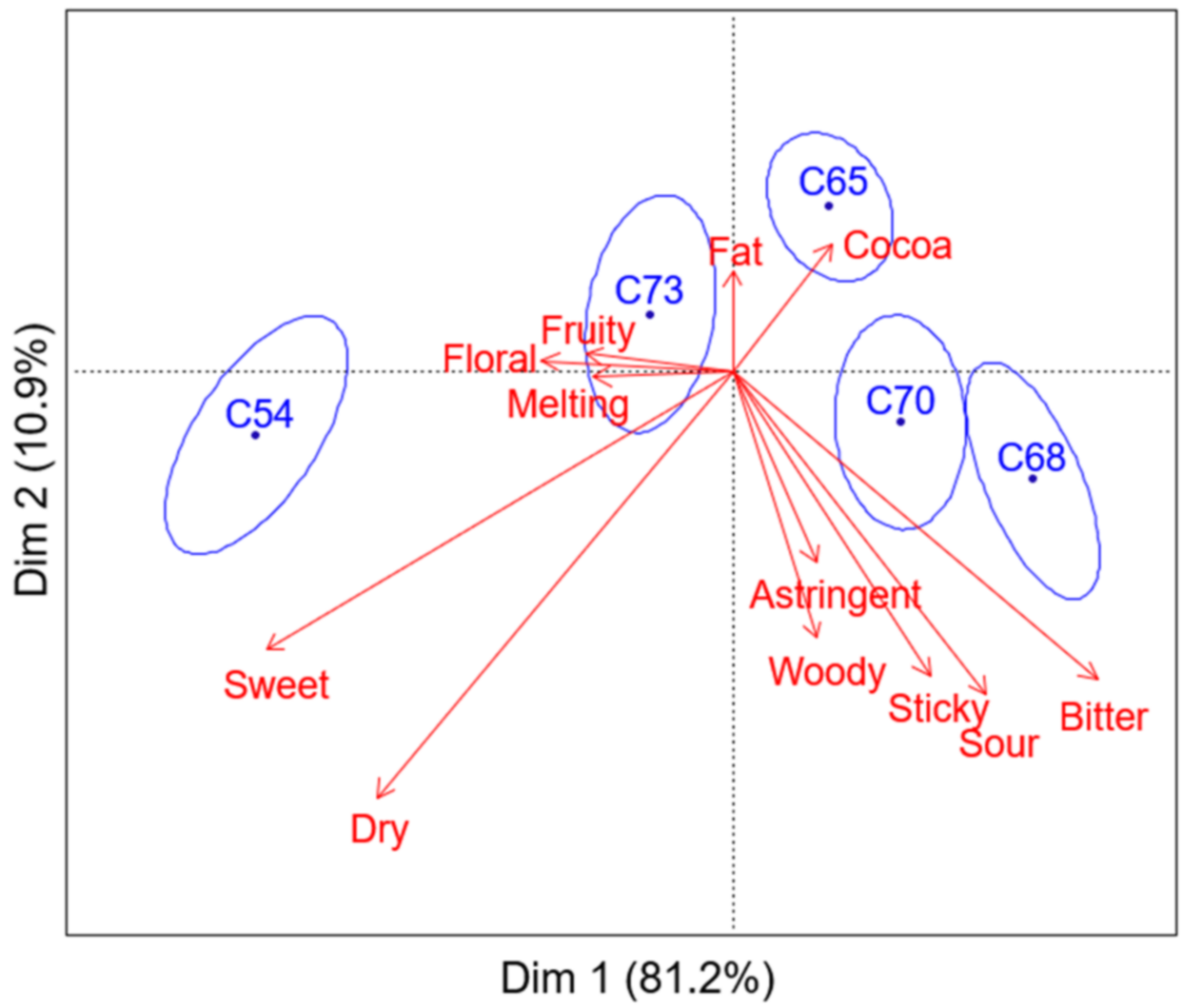


[TDS & AEF] Attack: **Cocoa**, **Dry**; Evolution: Bitter, **Cocoa**, **Melting**^{T}, **Sweet**^{T}; Finish: **Astringent**^{T}, **Cocoa**
[TDS only] Attack: **Fat**, **Melting**, **Sweet**; Evolution: Fat, Finish: **Bitter**, Melting

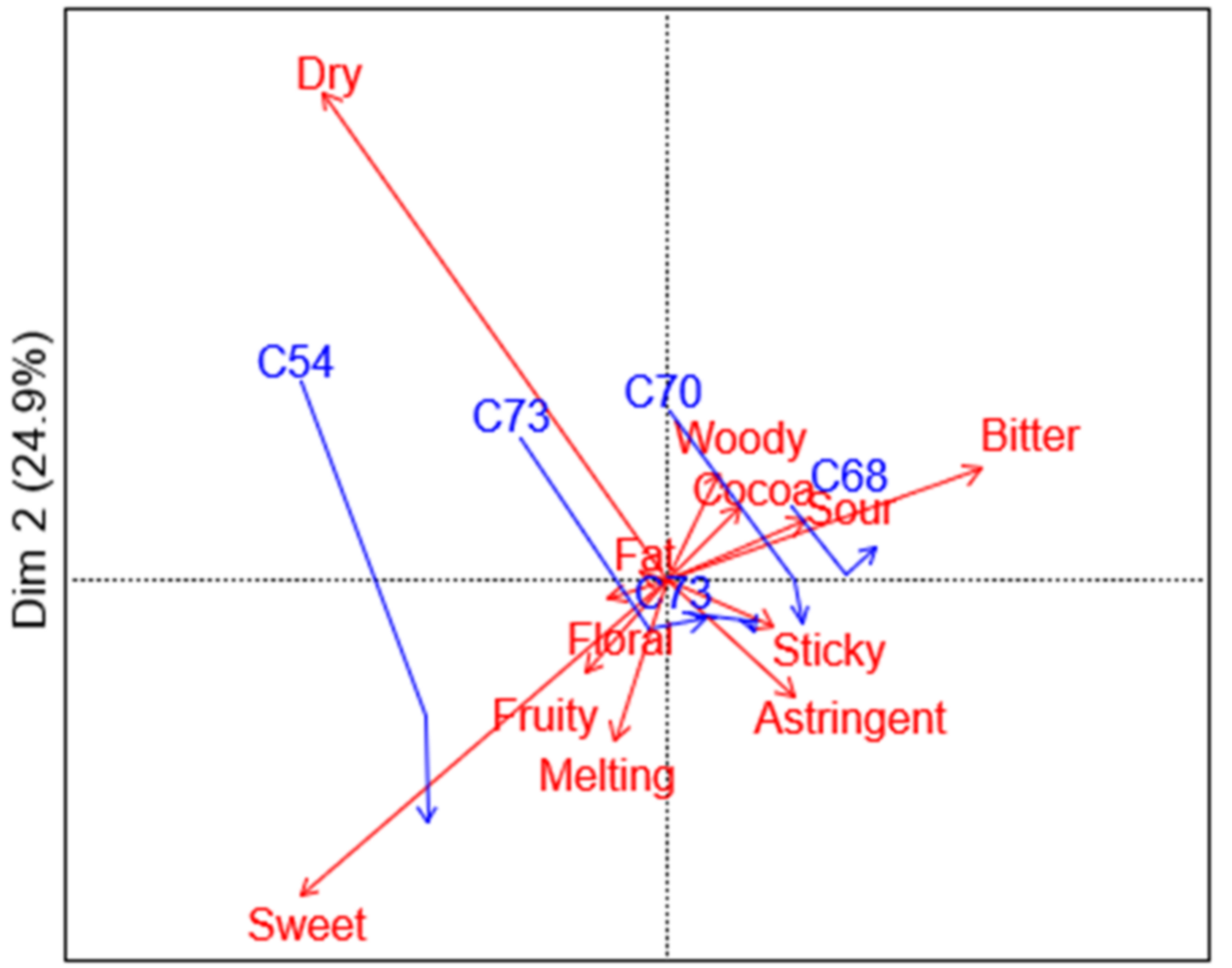
Astring. Bitter Cocoa Dry Fat Floral Fruity Melting Sour Sticky Sweet Woody

Astring.
Bitter
Cocoa
Dry
Fat
Floral
Fruity
Melting
Sour
Sticky
Sweet
Woody

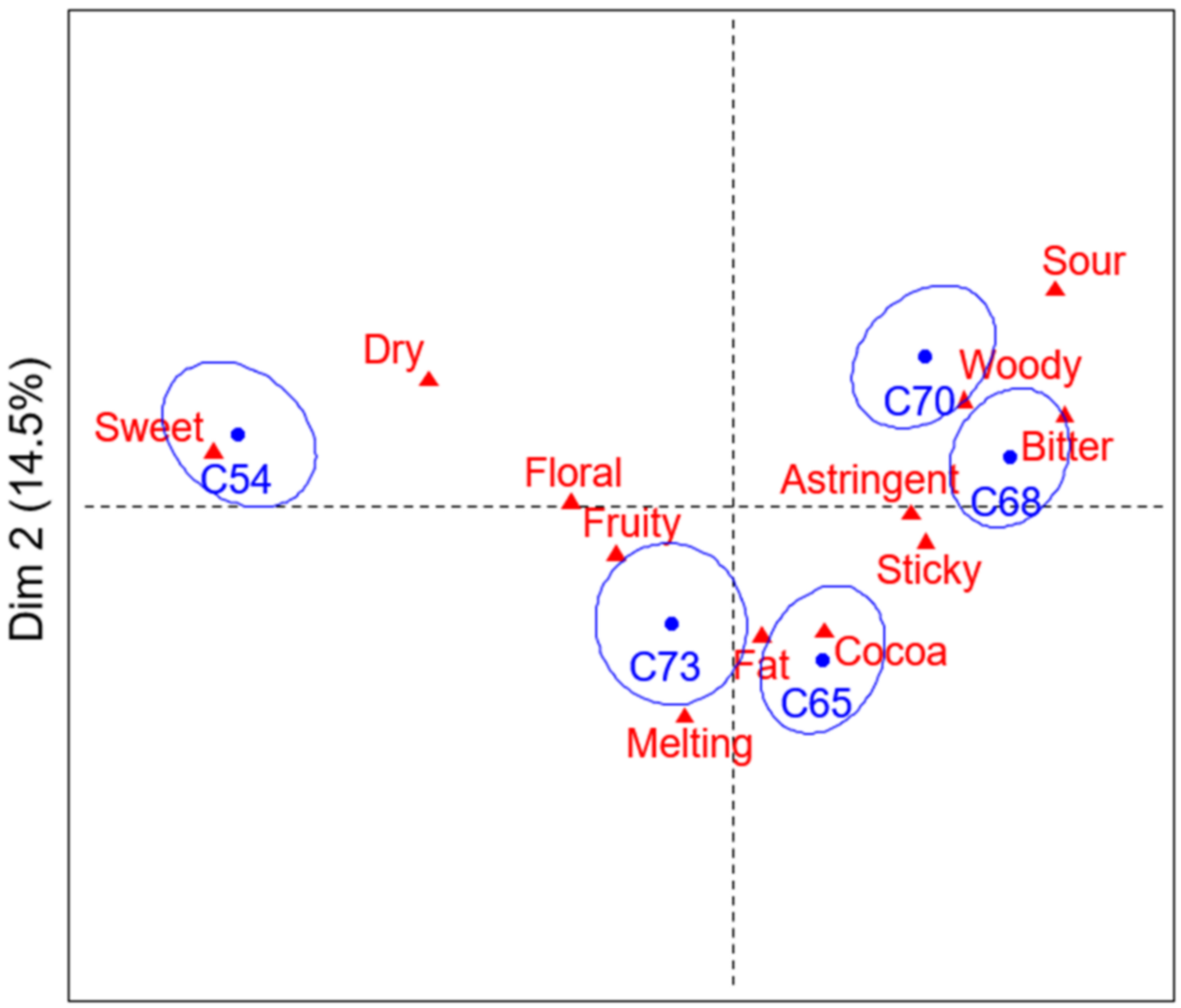
A


Dim 2 (10.9%)
C65
Cocoa
Fat
C73
Fruity
Floral
Melting
C54
C70
C68
Astringent
Woody
Sticky
Sour
Bitter
Sweet
Dry
Dim 1 (81.2%)

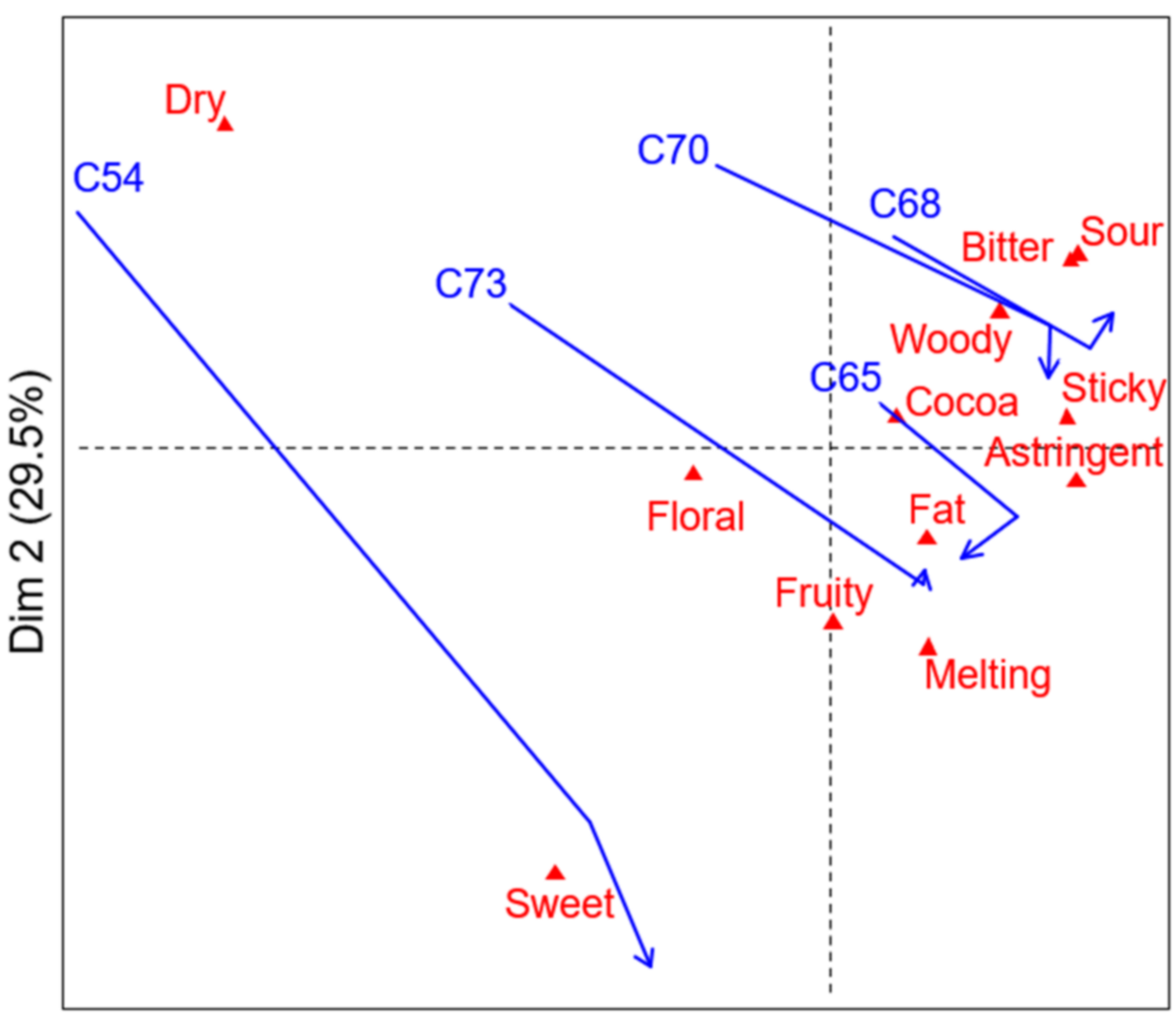
B


Dry
C54
C73
C70
Woody
Bitter
Cocoa
C68
Sour
Fat
C73
Floral
Sticky
Fruity
Astringent
Melting
Sweet
Dim 2 (24.9%)
Dim 1 (56.7%)

C


Sour
Dry
Woody
C70
Sweet
C54
Bitter
C68
Floral
Astringent
Fruity
Sticky
C73
Fat
Cocoa
C65
Melting
Dim 2 (14.5%)
Dim 1 (79.9%)

D


Dry
C54
C70
C68
Bitter
Sour
C73
Woody
C65
Cocoa
Sticky
Astringent
Floral
Fat
Fruity
Melting
Sweet
Dim 2 (29.5%)
Dim 1 (45.3%)