



HAL
open science

Descriptive temporal sensory evaluation with consumers

Michel Visalli

► **To cite this version:**

Michel Visalli. Descriptive temporal sensory evaluation with consumers. Food engineering. Université de Bourgogne, 2023. English. NNT: . tel-04533721v1

HAL Id: tel-04533721

<https://hal.science/tel-04533721v1>

Submitted on 5 Apr 2024 (v1), last revised 22 Apr 2024 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THESE DE DOCTORAT

Pour obtenir le grade de Docteur de l'Université de Bourgogne
en Sciences de l'Alimentation

Préparée à l'Institut National de Recherche pour l'Agriculture, l'Alimentation
et l'Environnement, dans l'Unité Mixte de Recherche Centre des Sciences du
Goût et de l'Alimentation

Ecole doctorale n°554 : Environnements – Santé

Descriptive temporal sensory evaluation with consumers

Par Michel VISALLI

Thèse présentée et soutenue à Dijon le 12/12/2023

Composition du jury :

M. Hervé CARDOT	Professeur, Université de Bourgogne	Examineur
Mme Anne SAINT-EVE	Professeure, AgroParisTech	Rapportrice
M. Pascal SCHLICH	Directeur de recherche, INRAE	Directeur de thèse
Mme Claire SULMONT-ROSSE	Directrice de recherche, INRAE	Examinatrice
M. Ronan SYMONEAUX	Chargé de recherche HDR, ESA	Rapporteur
Mme Evelyne VIGNEAU	Professeure, ONIRIS	Examinatrice

"Aussi audacieux soit-il d'explorer l'inconnu, il l'est plus encore de remettre en question le connu."

Walter Kasper

Abstract

The transformation of food in the mouth and the release of aromatic and sapid molecules leads to temporal changes in the sensory perception of tasters. This dynamic of sensory perception has been studied for many years using temporal sensory analysis methods. Initially, these methods were developed with trained panels, but over the last years, they shifted from quantitative to qualitative measurements, resulting in rapid methods which can be applied with consumers. While this evolution is desirable to make measurements more representative of perception of consumers, it was likely driven more by practicalities than by scientific considerations. Indeed, rapid methods compromise between costs, ecological validity and relevance of data collection.

A preliminary systematic scoping review of the literature has evidenced that despite 30 years of research, there is still no clear recommendation on the application of temporal methods with consumers, and it is unclear to what extent implementation details may impact the results. It was therefore legitimate to ask the question "how appropriate are rapid sensory evaluation methods for measuring the temporal perception of consumers?".

First, the granularity of temporal sensory data was questioned. The results obtained with sensory evaluation methods recording data concurrently with perception (during tasting) were compared with those obtained with a new method collecting data retrospectively to the tasting and summarizing perception in three periods. Different variants of this method were implemented to evaluate several product categories, varying the measured concepts (dominance or applicability), the mode of data acquisition (list of attributes or free comment), and the environment (laboratory or home). No severe loss of temporal information was observed with retrospective measurements, the latter making it possible to discriminate between products as well as concurrent methods while simplifying collection and analysis of temporal data.

Then, the performances (validity, reliability, resolution and discrimination ability) of methods were compared to better understand the nature of differences observed between the methods. From data collected on controlled temporal stimuli (delivered in the mouth using a gustometer), it has been shown that concurrent methods can be used with consumers to record relevant sequences of sensations. With complex food products, a large heterogeneity

in temporal descriptions was observed among consumers. Neither concurrent nor retrospective methods were able to reliably catch subtle temporal differences between products, discrimination being mostly related to presence/absence of sensory attributes regardless of temporal aspects.

These results have demonstrated that it is possible to measure the temporal perception of consumers using rapid temporal sensory evaluation methods, but that one should not expect more from the methods than what they can measure given the complexity of the phenomenon studied. These new insights may help both manufacturers (who wish to gain information about their products) and researchers (who want to understand temporal perception) to choose the most appropriate rapid method of temporal data collection depending on their question, to implement it effectively with consumers, and finally to analyse temporal data in such a way that robust conclusions can be drawn.

This research work has also highlighted the difficulty of generalizing the conclusions on the basis of a few studies. In the future, perhaps meta-analyses carried out on large databases of sensory studies will make it possible to refine or challenge the conclusions of this work. Thus, we have tried to initiate the process by making all the collected datasets public, and by proposing a template for sharing data accompanied by guidelines for reporting research results on the topic.

Résumé

La transformation des aliments en bouche et la libération de molécules aromatiques et sapides entraînent des changements temporels dans la perception sensorielle des dégustateurs. Cette dynamique est étudiée depuis longtemps à l'aide de méthodes d'analyse sensorielle temporelle, qui ont été initialement développées avec des panels entraînés. Ces dernières années, le passage de mesures quantitatives à qualitatives a rendu ces méthodes utilisables par des consommateurs. Bien que cette évolution soit souhaitable pour rendre les mesures plus écologiques, elle a probablement été davantage motivée par des intérêts pratiques que par des considérations scientifiques.

Une revue systématique de la littérature a mis en évidence que malgré 30 ans de recherche, il n'y a toujours pas de recommandation claire sur l'application des méthodes d'analyse sensorielle temporelle avec des consommateurs, et on ne sait pas dans quelle mesure les détails de leur implémentation peuvent avoir un impact sur les résultats. Il était donc légitime de se poser la question « dans quelle mesure les méthodes rapides d'analyse sensorielle sont-elles appropriées pour mesurer la perception temporelle de consommateurs ? ».

Les résultats obtenus avec des méthodes collectant les données pendant la dégustation ont été comparés à ceux obtenus avec une nouvelle méthode recueillant les données rétrospectivement à la dégustation résumant la perception en trois périodes. Différentes variantes de cette méthode ont été implémentées en faisant varier le concept mesuré (dominance ou applicabilité), le mode d'acquisition des données (liste d'attributs ou commentaires libres), et l'environnement de collecte (laboratoire ou domicile). Aucune perte majeure d'information temporelle n'a été observée avec les mesures rétrospectives, et ces dernières ont permis de discriminer les produits aussi bien que les méthodes simultanées tout en simplifiant la collecte et l'analyse des données, ce qui questionne sur la granularité des mesures temporelles.

Les performances (validité, fiabilité, résolution temporelle et pouvoir discriminant) des différentes méthodes ont donc été comparées afin de mieux comprendre la nature des différences de mesures observées. À partir de données d'évaluations de stimuli temporels contrôlés (délivrés avec un gustomètre), il a été montré que les méthodes simultanées peuvent être utilisées avec des consommateurs pour enregistrer des séquences de perception

pertinentes. Avec des aliments complexes, ces méthodes n'ont pas été en mesure de détecter systématiquement des différences temporelles subtiles entre produits, la discrimination étant principalement liée à la présence/absence d'attributs sensoriels (indépendamment de leur temporalité).

Ces résultats suggèrent qu'il est possible de mesurer la perception temporelle des consommateurs à l'aide de méthodes rapides, mais qu'il ne faut pas attendre plus de ces dernières que ce qu'elles peuvent mesurer compte tenu de la complexité du phénomène étudié. Ces nouvelles connaissances pourront aider à la fois les industriels (pour obtenir des informations sur leurs produits) et les chercheurs (pour comprendre les mécanismes de la perception temporelle) à choisir la méthode la plus appropriée, à l'implémenter efficacement, et enfin à analyser les données de telle manière à pouvoir en tirer des conclusions valides.

Ce travail a également mis en évidence la difficulté de généraliser des conclusions sur la base de quelques études uniquement. A l'avenir, des méta-analyses réalisées sur des bases de données pourraient permettre d'affiner les conclusions obtenues. Ainsi, nous avons essayé d'initier le processus en rendant publics tous les jeux de données collectés et en proposant un modèle de données accompagné de lignes directrices permettant de mieux communiquer les résultats de la recherche impliquant des méthodes d'analyse sensorielle.

Acknowledgments

I am extremely grateful to Pascal Schlich, not only for his supervision, but also for having involved me in research all these years, until I finally decide to undertake this research work. A huge thank you for the trust and freedom he has always given me even when we did not agree (that is to say often). I will miss our (sometimes heated) debates, the next ones may be calmer, on a boat or by a river!

I would like to express my deepest gratitude to Anne Saint-Eve, Claire Sulmont-Rossé, Evelyne Vigneau, Hervé Cardot and Ronan Symoneaux for agreeing to form my defense committee. Many thanks to Anne and Ronan for having accepted to report on this work, and sorry for the length of the manuscript! Many thanks to Evelyne for her advices. Many thanks to Claire for her training on systematic reviews, it changed my perspectives! Many thanks to Hervé for his teachings, a long time ago (what a journey since we met at CESAER!).

I also could not have undertaken this journey without my co-authors. Special thanks to Caroline, Benjamin and Guillaume, to whom I repeated that I will never defend a thesis, it would seem that they changed my mind! Special thanks to Mara for following me (and for restraining me when necessary) in the huge work of the scoping review. Special thanks also to Magalie, it was nice working (and drinking wine) together! Thanks to Takahiro, now it is your turn for the PhD! I am also very grateful to Betty, Catherine, Léna, Noëlle, Marine, Marion and Sylvie for their essential contributions in the management of the experimental part of the studies.

I warmly thank my closest colleagues, in particular Anne-Laure, Arnaud, Carole, Cédric, Christine, Christine, Christophe, Eric and Mitch. I don't forget the other people from CSGA or INRAE who do shadow work (Benjamin, Christine, Michaël) and all other people I met during this work (and before), too numerous to mention.

Lastly, I would be remiss in not mentioning my friends and family, especially my partner, and step-children. I can never thank Fouzia enough for supporting and accompanying me for so many years, I wouldn't be here without her. Many thanks to Anissa and Elias, I am proud of the adults you have become. I would also like to thank my parents for everything, with a special thought for my deceased grandmother who always believed in me. Thanks to my friends for not pretending to be interested by what I did, it helps keep my feet on the ground! Thanks to my cats Kalouch and Luna for their company during the (very) many hours I spent in my office.

Table of contents

Abstract	5
Résumé	7
Acknowledgments	9
Table of contents	11
List of figures and tables	15
Figures.....	15
Tables.....	19
Glossary	23
List of abbreviations related to temporal methods	31
Foreword	33
Short presentation.....	33
Publications in peer-reviewed journals	33
Research protocol	33
Data papers	33
Research articles	34
Ethics statement	35
Declaration of interests	35
Chapter I. Introduction	37
I.1. Sensory perception of foods.....	39
I.1.1. What is sensory perception?	39
I.1.2. Why human sensory perception of food products is so important?.....	40
I.2. Measuring food sensory perception	41
I.2.1. The periods of development of sensory-consumer science.....	41
I.2.2. Main categories of sensory evaluation methods.....	43
I.3. Temporal sensory evaluation methods.....	45
I.3.1. Why the temporal dimension of perception is important?.....	45
I.3.2. Multi-attribute temporal descriptive methods in sensory analysis applied in food science: protocol for a scoping review*	47
I.3.3. Multi-attribute temporal descriptive methods in sensory analysis applied in food science: a systematic scoping review*	75
I.4. The challenge of measuring temporal perception of consumers	147

I.4.1. Temporal methods, from analytical to rapid temporal measurements.....	147
I.4.2. Aims and objectives of this work	152
I.4.3. Plan of the manuscript.....	154
Chapter II. Methods	155
II.1. A dataset on concurrent and immediate retrospective measures of sensory perception and preferences of dark chocolates*	157
II.2. A dataset on concurrent and retrospective sensory perception and preferences for lemon-flavoured carbonated alcoholic drinks*	167
II.3. A dataset on the sensory and affective perception of Bordeaux and Rioja red wines collected from French and Spanish consumers at home and international wine students in the lab*	177
II.4. A dataset of sensory perception of chocolates, guacamoles, iced teas and crisps collected with consumers using six temporal methods*	191
II.5. A dataset on sensory evaluation of gustometer controlled stimuli with TDS, TCATA and AEF*	207
Chapter III. Results.....	217
III.1. Concurrent vs. retrospective temporal sensory data collection	219
III.1.1. Context and aims	219
III.1.2. Concurrent vs. retrospective temporal data collection: Attack-Evolution-Finish as a simplification of Temporal Dominance of Sensations?*	221
III.1.3. Concurrent vs. immediate retrospective temporal data collection. A case study on lemon-flavoured carbonated alcoholic drinks*	245
III.1.4. Relevance of free comment to describe wine temporal sensory perception: An application with panels varying in culture and expertise*	269
III.1.5. Main findings of sections III.1	293
III.2. Performances of rapid temporal sensory evaluation methods	295
III.2.1. Context and aims	295
III.2.2. Assessment of the reliability and validity of temporal sensory evaluation methods used with consumers on controlled stimuli delivered by a gustometer*	297
III.2.3. Temporal precision and resolution of TDS and TCATA used with consumers* ..	331
III.2.4. Product discrimination ability of temporal sensory evaluation methods used with consumers*	385
III.2.5. Main findings of section III.2.....	417
Chapter IV. General discussion	419
IV.1. Limitations	421

IV.2. Implications of key findings.....	422
IV.2.1. About the methods: what is asked vs. what is done	422
IV.2.2. About the data: what is collected vs. what is analysed.....	425
IV.3. Recommendations.....	426
IV.3.1. For the method choice.....	426
IV.3.2. For the implementation of the methods.....	428
IV.3.3. For the analysis of temporal data	430
IV.3.4. For reporting research findings and sharing data related to temporal measurements	432
Chapter V. Conclusion	433
V.1. Temporal methods can be used with consumers, but too much should not be expected	435
V.2. Recommendations for future research	436
V.3. Use of temporal sensory evaluation methods with consumers: a position paper* ...	439
References.....	451

*: article

List of figures and tables

Figures

Figure 1. Wheel of descriptive temporal sensory evaluation methods.....	31
Figure 2. Sensory information processing, from physical stimulus to perceptual response... 39	39
----- In "Multi-attribute temporal descriptive methods in sensory analysis applied in food science: protocol for a scoping review":	
Figure 3. PRISMA flow chart to be completed in the final scoping review.....	58
Figure 4. Summary of the process of evaluation of the articles.	72
----- In "Multi-attribute temporal descriptive methods in sensory analysis applied in food science: a systematic scoping review":	
Figure 5. Main steps of the reviewing process.	78
Figure 6. PRISMA diagram for article identification, selection and inclusion.....	80
Figure 7. Main characteristics of descriptive sensory evaluation methods.	104
Figure 8. Number of articles published by year and type of method used.	106
Figure 9. Main objectives of the use of temporal methods.....	107
Figure 10. Other measurements found in articles which had multi-attribute temporal descriptive measurements.....	108
Figure 11. Categories of products evaluated using multi-attribute temporal methods	110
Figure 12. Distribution of number of products per study.....	110
Figure 13. Types of panel involved in multi-attribute temporal descriptive measurements over the years.....	111
Figure 14. Details of training by type of panel.....	112
Figure 15. Distribution of the number of subjects and evaluations by type of panel	114
Figure 16. Use of attributes with temporal evaluation methods	116
Figure 17. Characterization of the tastings	119
Figure 18. Examples of data collected at continuous times, fixed pre-determined moments of consumption and recapitulative time periods.....	125
Figure 19. Quality appraisal of the included articles.....	135
Figure 20. Field Weighted Citation Impact of articles by area of knowledge.....	137
Figure 21. Evolution of the use of multi-attribute temporal sensory evaluation methods by trained and consumer panels over years.....	148
Figure 22. Standardized TDS curves and difference curves of CH3, CH4 and CH4_rep.....	150
Figure 23. Results of the ESN survey.....	151
----- In "A dataset on concurrent and immediate retrospective measures of sensory perception and preferences of dark chocolates":	
Figure 24. Procedure of data collection	165
----- In "A dataset on concurrent and retrospective sensory perception and preferences for lemon-flavoured carbonated alcoholic drinks":	
Figure 25. Description of the tasting of the drink over seven sips	174
Figure 26. Evaluation of sips 1, 4, 7 for panel RET using retrospective CATA.....	174

Figure 27. Evaluation of sips 1, 4, 7 for panel SIM using simultaneous CATA	175
----- In "A dataset on the sensory and affective perception of Bordeaux and Rioja red wines collected from French and Spanish consumers at home and international wine students in the lab":	
Figure 28. Standardized labels of the samples.....	185
Figure 29. Participant flowchart.....	186
----- In "A dataset of sensory perception of chocolates, guacamoles, iced teas and crisps collected with consumers using six temporal methods":	
Figure 30. PCA of the averaged scores of the items of the complexity questionnaire by product category.	198
Figure 31. Experimental procedure chart	205
----- In "A dataset on sensory evaluation of gustometer controlled stimuli with TDS, TCATA and AEF":	
Figure 32. Gustometer model Burghart GU002.....	212
Figure 33. Screens displayed to the consumers.....	214
----- In "Concurrent vs. retrospective temporal data collection: Attack-Evolution-Finish as a simplification of Temporal Dominance of Sensations?":	
Figure 34. TDS and AEF measurement screens.....	224
Figure 35. Histograms of consumer behaviour observed during the TDS sequences	230
Figure 36. TDS curves and AEF barplots of the proportion of dominances for C54, C65, C68, C70, and C73.....	232
Figure 37. Multidimensional maps, axes 1 and 2.....	235
----- In "Concurrent vs. immediate retrospective temporal data collection. A case study on lemon-flavoured carbonated alcoholic drinks":	
Figure 38. Description of the tasting of the drink over seven sips	251
Figure 39. Evaluation of sips 1, 4, 7 for panel RET using retrospective CATA.....	251
Figure 40. Evaluation of sips 1, 4, 7 for panel SIM using simultaneous CATA	251
Figure 41. Use of descriptors	254
Figure 42. mrCA of 36 evaluations representing the product by sip sensory trajectories over periods on dimensions 1 and 2.	255
Figure 43. mrCA of 12 evaluations representing the product sensory trajectories over periods on dimensions 1 and 2.	257
Figure 44. mrCA of the four products, by period and averaged over all periods on dimensions 1 and 2, for panels RET and SIM.....	258
Figure 45. mrCA of the three periods, by product and averaged over all products on dimensions 1 and 2	260
----- In "Relevance of free comment to describe wine temporal sensory perception: An application with panels varying in culture and expertise":	
Figure 46. Summary of the experimental procedure.....	274
Figure 47. Screen for FC-AEF measurement.	275
Figure 48. mrCA biplot and table summarizing the results of the hypergeometric test between products	283

Figure 49. mrCA biplot and table summarizing the results of the hypergeometric test within product.....	286
Figure 50. mrCA biplot and table summarizing the results of the hypergeometric test within product for the WC panel.....	287
Figure 51. mrCA biplot and table summarizing the results of the hypergeometric test between products for the WC panel.....	287
----- In "Assessment of the reliability and validity of temporal sensory evaluation methods used with consumers on controlled stimuli delivered by a gustometer":	
Figure 52. Overview of the experimental procedure.....	304
Figure 53. Evaluations of stimuli with same concentrations and no simultaneous delivery of compounds.....	315
Figure 54. Evaluations of stimuli with constant concentrations and simultaneous delivery of compounds.....	316
Figure 55. Evaluations of stimuli with varying concentrations and simultaneous delivery of compounds.....	317
Figure 56. Distribution of the individual repeatability scores averaged over times or periods of the consumers, by product and method	321
Figure 57. Word cloud of the words used to describe the compound	329
----- In "Temporal resolution of sensory evaluation methods used with consumers":	
Figure 58. Representation of within-product differences between attributes for Crisp C1 (TDS)	348
Figure 59. Representation of within-product differences between attributes for Crisp C1 (TCATA).....	351
Figure 60. Representation of differences between C1, C1_rep and C4, by attribute (TDS, with time-standardization).....	354
Figure 61. Representation of differences between C1, C1_rep and C4, by attribute (TCATA, with time-standardization).....	356
Figure 62. Screenshots of the survey administrated to members of the European Sensory Network.....	367
Figure 63. Results of the ESN survey.....	367
Figure 64. Representation of differences between C1, C1_rep and C4, by attribute (TDS, without time-standardization).	368
Figure 65. Representation of differences between C1, C1_rep and C4, by attribute (TCATA, without time-standardization).	369
Figure 66. Representation of within-product differences between attributes for Chocolate CH4.	370
Figure 67. Representation of differences between CH1, CH4 and CH4_rep, by attribute (TDS, without time-standardization).	371
Figure 68. Representation of differences between CH1, CH4 and CH4_rep, by attribute (TDS, with time-standardization).	372
Figure 69. Representation of differences between CH1, CH4 and CH4_rep, by attribute (TCATA, without time-standardization).	373

Figure 70. Representation of differences between CH1, CH4 and CH4_rep, by attribute (TCATA, with time-standardization).....	374
Figure 71. Representation of within-product differences between attributes for Guacamole G1.	375
Figure 72. Representation of differences between G1, G1_rep and G4, by attribute (TDS, without time-standardization).	376
Figure 73. Representation of differences between G1, G1_rep and G4, by attribute (TDS, with time-standardization).	377
Figure 74. Representation of differences between G1, G1_rep and G4, by attribute (TCATA, without time-standardization).	378
Figure 75. Representation of differences between G1, G1_rep and G4, by attribute (TCATA, with time-standardization).	379
Figure 76. Representation of within-product differences between attributes for Ice tea IT3	380
Figure 77. Representation of differences between G1, G3 and G3_rep, by attribute (TDS, without time-standardization).	381
Figure 78. Representation of differences between G1, G3 and G3_rep, by attribute (TDS, with time-standardization).	382
Figure 79. Representation of differences between G1, G3 and G3_rep, by attribute (TCATA, without time-standardization).	383
Figure 80. Representation of differences between G1, G3 and G3_rep, by attribute (TCATA, with time-standardization).	384
----- In "Product discrimination ability of temporal sensory evaluation methods used with consumers":	
Figure 81. Controlled temporal stimuli delivered using a gustometer	388
Figure 82. Biplots of multivariate analyses with 95% confidence ellipses.....	394
Figure 83. Pairwise differences in citation rates plotted over time.....	395
Figure 84. Summary of pairwise differences between controlled stimuli, by attribute and method.	396
Figure 85. Guacamoles - Biplots of multivariate analyses with 95% confidence ellipses.....	405
Figure 86. Guacamoles - Pairwise differences in citation rates plotted over time.....	406
Figure 87. Guacamoles - Summary of pairwise differences, by attribute and method.....	407
Figure 88. Ice teas - Biplots of multivariate analyses with 95% confidence ellipses	408
Figure 89. Ice teas - Pairwise differences in citation rates plotted over time	409
Figure 90. Ice teas - Summary of pairwise differences, by attribute and method	410
Figure 91. Crisps - Biplots of multivariate analyses with 95% confidence ellipses.....	411
Figure 92. Crisps - Pairwise differences in citation rates plotted over time.....	412
Figure 93. Crisps - Summary of pairwise differences, by attribute and method	413
Figure 94. Chocolates - Biplots of multivariate analyses with 95% confidence ellipses.....	414
Figure 95. Chocolates - Pairwise differences in citation rates plotted over time.....	415
Figure 96. Summary of pairwise differences, by attribute and method.....	416
Figure 97. Flowchart for choosing the method of measuring consumer perception.....	427

Figure 98. Evolution of the use of multi-attribute temporal sensory evaluation methods by consumers and trained panels over the years.	440
Figure 99. Flowchart for choosing the method of measuring consumer perception.	444
Figure 100. Flowchart for data analyses.	448

Tables

----- In "Multi-attribute temporal descriptive methods in sensory analysis applied in food science: protocol for a scoping review":

Table 1. Main sensory descriptive temporal methods published in peer-reviewed articles ..	49
Table 2. Existing reviews related to all temporal methods in sensory science.	52
Table 3. Meta-data exported from WOS and Scopus.	58
Table 4. Generic standards applicable for reporting studies involving temporal sensory methods.....	60
Table 5. Extraction form based on topic-related questions.....	64
Table 6. Indicators of quality of articles.	70
Table 7. Reasons for not checking quality criteria.	79

----- In "Multi-attribute temporal descriptive methods in sensory analysis applied in food science: a systematic scoping review":

Table 8. Articles included in the present scoping review	82
Table 9. Descriptive temporal sensory evaluation methods	102
Table 10. Attributes most frequently used in main product categories.....	118
Table 11. Details of implementation of temporal methods with products evaluated over successive intakes	122
Table 12. Main variables derived from data collected with temporal methods	127
Table 13. Main statistical analyses reported in the literature.	130
Table 14. Main conclusions reported in articles involving method comparisons	133
Table 15. Checklist for reporting research involving sensory evaluation methods.....	139
Table 16. Template for sharing sensory data.....	142
Table 17. SWOT analysis of rapid temporal sensory evaluation methods	152
Table 18. Summary of datasets collected.	156

----- In "A dataset on concurrent and immediate retrospective measures of sensory perception and preferences of dark chocolates":

Table 19. Specifications table.....	158
-------------------------------------	-----

----- In "A dataset on concurrent and retrospective sensory perception and preferences for lemon-flavoured carbonated alcoholic drinks":

Table 20. Specifications table.....	168
Table 21. Information about products	171

----- In "A dataset on the sensory and affective perception of Bordeaux and Rioja red wines collected from French and Spanish consumers at home and international wine students in the lab":

Table 22. Specifications table.....	178
Table 23: Objective information about wines.....	183

Table 24. Subjective information about wines.	184
Table 25. Participant characteristics.	187
----- In "A dataset of sensory perception of chocolates, guacamoles, iced teas and crisps collected with consumers using six temporal methods":	
Table 26. Summary of the main characteristics of the six temporal methods.	196
Table 27. Product categories, codes and composition	197
Table 28. Averaged scores of the items of the complexity questionnaire	198
Table 29. Attributes used with the TDS, TCATA, AEF-D and AEF-A methods	199
Table 30. Number of consumers having evaluated each product category, by panel.	200
Table 31. Individual characteristics of the consumers, by panel	201
----- In "A dataset on sensory evaluation of gustometer controlled stimuli with TDS, TCATA and AEF":	
Table 32. Specifications table.....	208
Table 33. Composition of the consumer panels.	212
Table 34. Information about the sapid and aromatic compounds used in the stimuli	213
----- In "Concurrent vs. retrospective temporal data collection: Attack-Evolution-Finish as a simplification of Temporal Dominance of Sensations?":	
Table 35. ANOVA of durations or citations by descriptor, method and period.....	234
Table 36. Mean answers to the questionnaire and their 95% confidence intervals	237
----- In "Concurrent vs. immediate retrospective temporal data collection. A case study on lemon-flavoured carbonated alcoholic drinks":	
Table 37. Percentages of citations by attribute, product and period for panel RET	259
Table 38. Percentages of citations by attribute, product and period for panel SIM	259
Table 39. Percentages of citations by attribute, product and period for panel RET	261
Table 40. Percentages of citations by attribute, product and period for panel SIM	261
---- In "Relevance of free comment to describe wine temporal sensory perception: An application with panels varying in culture and expertise":	
Table 41. Characteristics of participants, by panel	279
Table 42. Average number of words used per participant to describe wine samples	280
Table 43. Groups of words obtained after Greenacre classification	281
----- In "Assessment of the reliability and validity of temporal sensory evaluation methods used with consumers on controlled stimuli delivered by a gustometer":	
Table 44. Description of the sapid and aromatic compounds used in the stimuli	303
Table 45. Variables derived from the controlled stimuli	308
Table 46. Percentage of citations.....	311
Table 47. Construct validity of the measurements	318
Table 48. Content validity of the measurements.....	318
Table 49. Criterion validity of the measurements	318
Table 50. Words associated with "Fruity" and "Vegetal"	329
Table 51. Comparisons of mean durations of perception between the stimuli.	329
Table 52. Content validity of AEF-RATA intensity measurements.	330
Table 53. Criterion validity of AEF-RATA intensity measurements.....	330

Table 54. Reliability of the measurements.	330
----- In "Temporal resolution of sensory evaluation methods used with consumers":	
Table 55. Products evaluated by the consumers	337
Table 56. Mean and standard deviation of total duration of perception, number of distinct attributes selected and number of citations per consumer, by method and product category.....	345
Table 57. Relative standard deviation of temporal measurements	346
Table 58. Agreement and repeatability of temporal measurement.....	347
Table 59. Time necessary between two temporal events so that they are significantly discriminated.....	358
Table 60. Characterization of datasets.....	390
Table 61. Times of begin and end of delivery of the controlled temporal stimuli	390
----- In "Product discrimination ability of temporal sensory evaluation methods used with consumers":	
Table 62. Count of discriminated pairs of samples	398
Table 63. Count of discriminated pairs of samples, by variable, all categories combined....	400

Glossary

----- A -----

Absolute (perception) threshold: smallest amount of stimulation needed for a **subject** to detect that stimulus 50% of the time.

Acceptability: degree of appreciation of a **product**, evaluated overall or for particular **sensory attributes**.

Analytical (or objective) sensory evaluation method: any method in which the effects of personal opinions, attitudes, beliefs, feelings and emotions are minimized.

Analysis of variance (ANOVA): statistical model and associated estimation procedure used to analyse the differences among means.

Affective (or subjective) sensory evaluation method: any method based on personal opinions, attitudes, beliefs, feelings and emotions.

Applicability (of a sensory attribute): **qualitative concept** associated with an **attribute** relevant to describe the **perception** of a **product**.

(Sensory) attribute (or descriptor): perceptible characteristic of a **product**.

----- B -----

Between products difference: difference between identical or distinct **attributes evaluated** in distinct **products**.

Bootstrapping: statistical procedure consisting in resampling a **dataset** to create many simulated samples and allowing to construct confidence intervals and perform hypothesis testing.

----- C -----

Canonical Variate Analysis (or CVA): statistical method used for analysing group structure in multivariate data.

Categorical Functional Data Analysis (or CFDA): statistical method extending the multiple **Correspondence Analysis** towards functional data.

Concurrent (or simultaneous) sensory evaluation: evaluation carried-out during the tasting (as opposed to **retrospective sensory evaluation**).

Confirmatory data analyses: process used to evaluate evidence by producing parameters estimates and their confidence intervals and testing hypotheses about parameters.

Consumer panel: sensory panel having little or no experience with **sensory evaluation** but who consumes or uses the **product** of interest.

Construct validity: degree to which a measurement accurately measures the theoretical construct (or concept) it is intended to measure.

Content validity: degree to which a measurement evaluates all aspects of the concept that it is designed to measure.

Continuous time signal: signal defined over an uninterrupted range of time.

Count data: data reflecting the number of occurrences of an event (in a fixed period of time).

Correspondence Analysis (CA): visualization technique used for finding out and displaying the relationship between two categorial variables.

Criterion validity: degree to which a measurement agrees with a gold standard or an external criterion of the phenomenon being measured.

----- D -----

Data paper: document published in a peer reviewed journal describing a **dataset** and offering free access to it.

Dataset: collection of information obtained through observations, measurements, study, or analysis.

Descriptive sensory evaluation method: any method used to describe or quantify the sensory **attributes** of **products** by a **sensory panel**.

Difference threshold: value of the smallest perceptible difference in the physical intensity of a stimulus.

Discrete time signal: sequence of values that correspond to particular instants in time.

Discriminative ability: ability to capture quantitative and/or qualitative differences between two or more **products**.

Discrimination (statistical): qualitative and/or quantitative differentiation between two or more **products**.

Discrimination test: any method involving direct comparison between **products** to determine if differences are perceptible.

Dominance (of a sensory attribute): **qualitative concept** associated with an **attribute** that catch the attention (at a given time).

Dynamic sensory evaluation: **sensory evaluation** carried out at different time point during the tasting, combination of **concurrent** and **temporal sensory evaluation** (as opposed to **static sensory evaluation**).

----- E -----

Ecological validity: degree to which a measurement is generalizable and predictable of behaviours in real-life settings.

Expert panel: group of **subjects** having professional sensory experience with the **product**, but not necessarily introduced to the methods of **sensory evaluation**.

Exploratory data analysis: graphical and quantitative ways of representing data to help derive hypotheses to be tested.

Face validity: degree to which a measurement appears to measure what it claims to measure based on a subjective judgment.

----- F -----

FAIR data: data meeting the principles of findability, accessibility, interoperability, and reusability.

Familiarization: short **training** (<10 min) where **subjects** become acquainted with the list of **sensory attributes** and applied them to a practice sample.

----- G -----

Granularity (of data): level of detail in a data structure.

Gustometer: device used to deliver a controlled amount of liquid stimulus varying along time into the **subject'** mouth.

Gustatory working memory: dynamic encoding, maintenance, and retrieval of taste on short time-scales.

----- H -----

Hedonic: relating to like or dislike.

----- I -----

Intensity (of perception): magnitude of the perceived **sensation**.

Intensity rating: number associated with the **intensity** of a **perception**.

----- M -----

Markov model: chain of **stochastic** events, in which the probability of each event transition depends only on the state of the previous event.

Measurement error: difference between an observed value and the true value.

Meta-analysis: formal quantitative study used to systematically assess the results of previous research by combining the results of multiple studies.

Metadata: data that describe data.

Multiple-Response Correspondence Analysis (mrCA): extension of **Correspondence Analysis** which considers evaluations (instead of citations) as experimental units.

----- N -----

Narrative (or traditional) literature review: review providing an overview of a research topic with no clear methodological approach.

Noise (in data): unwanted or irrelevant variation in the measurement that does not reflect the true state of the phenomenon being measured.

----- O -----

Open science: movement to make scientific research, data and their dissemination more transparent and democratic.

----- P -----

Palatability: expected positive **hedonic** evaluation of food's sensory characteristics.

Perception: awareness of the effects of a sensory stimulation.

Performance (of a sensory evaluation method): metric(s) used to quantify the efficiency and/or effectiveness of a method.

Preference: selection by a **subject** of one **product** over the others in a given set based on **hedonic** criteria.

Principal Component Analysis (PCA): technique that transforms high-dimensions data into lower-dimensions while retaining as much information as possible.

Product: object (edible or not) assessed using **sensory evaluation**.

Psychophysics: study of relationships between measurable stimuli and the corresponding sensory responses.

----- Q -----

Qualitative concept: concept for which measurements result in nominal or ordinal variables.

Qualitative measurement: description of the **sensory attributes** of a **product** without **intensity ratings**.

Quality (or critical) appraisal: process of systematic assessment of an article of a **systematic review** to check if it meets the necessary criteria to be considered relevant and trustworthy.

Quantitative concept: concept for which measurements result in interval or ratio variables.

Quantitative measurement: description of the **sensory attributes** of a **product** with **intensity ratings**.

----- R -----

Random error: difference between the observed and true values due to chance and averaging close to zero over a sufficient large number of observations.

Rapid sensory evaluation method: method that can be implemented and performed quickly since minimal or no **training** is required prior to the assessment.

Reference-based sensory evaluation method: method using reference **products** against which other **products** in the set are compared.

Reliability: consistency of a measurement.

Repeatability: extent to which a **subject** does not vary in its repeated evaluations of the same **product** in identical testing conditions.

Replicate: repeated measurement carried out with the same combination of parameters.

Reproducibility: variation in measurements made under controlled changing conditions.

Retrospective sensory evaluation: evaluation carried-out after the tasting (as opposed to **concurrent sensory evaluation**).

----- S -----

Sample: specimen of a **product** presented for **sensory evaluation**.

Scaling: application of numbers to quantify sensory experience.

Scoping literature review: review mapping the size, characteristics, or scope of existing literature in a field of interest.

Semi-trained panel: group of **subjects** slightly trained or having previous experience with the method but not specifically trained to evaluate the **product** of interest.

Sensation: psychophysiological reaction resulting from a sensory stimulus.

Sensitivity: ability of **subjects** to perceive, identify and differentiate stimuli (qualitatively or quantitatively).

Sensometrics: discipline involving the application of mathematical, statistical, and computational methods to the field of sensory-consumer science.

Sensory evaluation (or sensory analysis): discipline involved with the assessment of the **sensory attributes** of a **product** by the senses.

Sensory modality: term referring to the different channels through which sensory information is perceived (texture, taste, etc.)

(Sensory) panel: group of **subjects** participating in **sensory evaluation**.

Sensory property: intrinsic characteristic of a **product** that can be perceived by the human senses.

Sensory-specific satiety: decline in pleasantness associated with a specific eaten food while other non-eaten foods remain pleasant.

Static sensory evaluation: evaluation carried out at a single time point before, during or after the tasting (as opposed to **dynamic sensory evaluation**).

Stochastic (or random) variable: variable for which observations are subject to random variation.

Stochastic process: observations of a **stochastic variable** through time and/or space.

Subject (or panellist, assessor): any person taking part in a **sensory evaluation**.

Systematic error (or bias): error which systematically leads to an estimate of a parameter with a mathematical expectation departing from the true value.

Systematic literature review: review following a structured methodology and based on a strict protocol defining specific inclusion and exclusion criteria to identify, appraise, and synthesize the relevant literature.

----- T -----

Temporal sensory evaluation: **sensory evaluation** carried out with consideration for the temporal aspects of **perception** (as opposed to **static sensory evaluation**).

Temporal resolution (of a method): term referring either to the temporal precision in data collection (continuous time, discrete time, periods, global) or to the minimum size of detectable differences by the method **within and between products** based on temporal aspects (durations, times of citations).

Time series: realisation of a **stochastic process** indexed by the time.

Trained panel: group of **subjects** intensively trained on the method and specifically trained to evaluate the type of **products** of interest.

Training: series of sessions in which **subjects** are oriented to the tasks to be completed in assessing particular types of **products**, which may include relevant product characteristics, standard rating scales, techniques of evaluation and terminology.

----- U -----

Ultra-processed food: food processed using many ingredients including food additives improving **palatability**, processed raw materials or ingredients rarely used in home cooking.

----- V -----

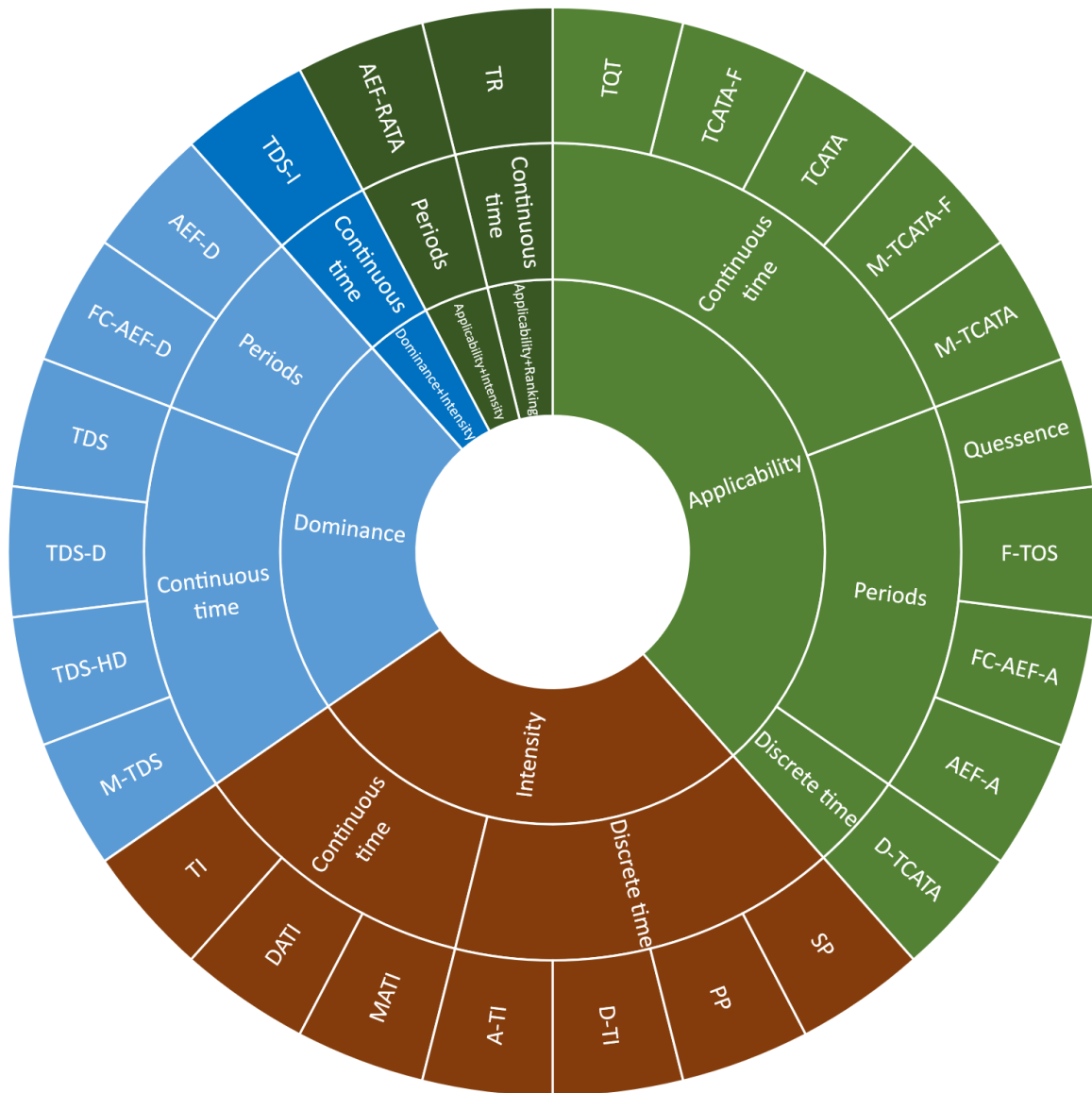
Verbal-based sensory evaluation methods: methods based on evaluation of individual **attributes** referenced by words.

----- W -----

Within product difference: temporal difference in times of citation or durations between distinct **attributes** in the same **product**.

List of abbreviations related to temporal methods

Figure 1. Wheel of descriptive temporal sensory evaluation methods



A-TI: Alternated Time-Intensity; **AEF-A:** Attack-Evolution-Finish Applicability; **AEF(-D):** Attack-Evolution-Finish Dominance; **AEF-RATA:** Attack-Evolution-Finish with RATA; **D-TCATA:** Discrete time TCATA; **D-TI:** Discrete time TI; **DATI:** Dual-Attribute Time-Intensity; **F-TOS:** Free TOS; **FC-AEF(-A):** Attack-Evolution-Finish Applicability with free comment; **FC-AEF-D:** Attack-Evolution-Finish Applicability with free comment; **M-TCATA:** TCATA by modality; **M-TCATA-F:** TCATA fading by modality; **M-TDS:** TDS by modality; **MATI:** Multi-Attribute Time-Intensity; **PP:** Progressive Profile; **Quessence;** **SP:** Sequential Profile; **TCATA:** Temporal Check-All-That-Apply; **TCATA-F:** TCATA fading; **TDS:** Temporal Dominant of Sensations; **TDS-D:** Dual TDS; **TDS-HD:** TDS hold down; **TDS-I:** TDS with intensity; **TI:** Time-Intensity; **TQT:** Time-Quality Tracking; **TR:** Temporal ranking.

Foreword

Short presentation

I have joined the ChemoSens platform of the Centre for Taste and Feeding Behaviour (CSGA, Dijon, France) in 2008 as a computer scientist. For more than ten years, I was in charge of the development of TimeSens, a software for sensory data collection and analysis. This gave me the opportunity to discover food science through the prism of methodological developments and sensometrics.

I have contributed as a co-author to numerous research programs aimed at exploring consumer perception of food, most involving temporal sensory measurements. However, this thesis by publication only refers to articles submitted or published as a first or last author since 2020, when I started to think about a PhD. This research work was carried out under the scientific direction of Pascal Schlich, INRAE research director.

Publications in peer-reviewed journals

Research protocol

Visalli, M., Galmarini M. (2022). *Multi-attribute temporal descriptive methods in sensory analysis applied in food science: protocol for a scoping review*. PlosONE, 17(7). <10.1371/journal.pone.0270969> <hal-03744281>

Data papers

Béno, N., Nicolle, L., Visalli, M. (2023). *A dataset of consumer perceptions of gustometer-controlled stimuli measured with three temporal sensory evaluation methods*. Data in Brief, 48, 109271. <10.1016/j.dib.2023.109271> <hal-04145806>

Visalli, M., Cordelle, S., Mahieu, B., Pedron, C., Hoffarth, B., Praudel, M., Coutière, M., & Schlich, P. (2022). *A dataset of sensory perception of chocolates, guacamoles, ice teas and crisps collected with consumers using six temporal methods*. Data in Brief, 45, 108708. <10.1016/J.DIB.2022.108708> <hal-03863058>

Visalli, M., Dubois, M., Schlich, P., Ric, F., Cardebat, J.-M., & Georgantzis, N. (2023). *A dataset on Bordeaux and Rioja red wines sensory and affective perception collected from French and Spanish consumers at home and international wine students in lab*. *Data in Brief*, 46, 108873. <10.1016/j.dib.2022.108873> <hal-03985366>

Visalli, M., Wakihira, T., Schlich, P. (2022). *A dataset on concurrent and retrospective methods for measuring sensory perception and preferences of lemon-flavoured carbonated alcoholic drinks*. *Data in Brief*, 43, 108346. <10.1016/j.dib.2022.108346> <hal-03729303>

Visalli, M., Mahieu, B., Schlich, P. (2023). *A dataset on concurrent and immediate retrospective measures of sensory perception and preferences of dark chocolates*. *Data in Brief*, 49, 109314. <10.1016/j.dib.2023.109314> <hal-04158500>

Research articles

Published research articles

Visalli, M., Schlich, P., Thomas, A., Mahieu, B. (2020). *Concurrent vs. retrospective temporal data collection: Attack-Evolution-Finish as a simplification of Temporal Dominance of Sensations?* *Food Quality and Preferences*, 85, 103956. <10.1016/j.foodqual.2020.103956> <hal-02869715v1>

Visalli, M., Wakihira T., Schlich P. (2022). *Concurrent vs. retrospective temporal data collection. A case study on lemon-flavoured carbonated alcoholic drinks*. *Food Quality and Preference*, 101, 104629. <10.1016/j.foodqual.2022.104629> <hal-03711757>

Visalli, M., Dubois, M., Schlich, P., Ric, F., Cardebat, J.M., Georgantzis, N. (2023). *Relevance of free comment to describe wine temporal sensory perception: An application with panels varying in culture and expertise*. *Food Quality and Preferences*, 105, 104785. <10.1016/j.foodqual.2022.104785> <hal-03985522>

Visalli, M., Béno, N., Nicolle, N., Schlich, P. (2023). *Assessment of the reliability and validity of temporal sensory evaluation methods used with consumers on controlled stimuli delivered by a gustometer*. *Food Quality and Preferences*, 110, 104942. <10.1016/j.foodqual.2023.104942> <hal-04176922v1>

Visalli, M., Galmarini, M. V., Schlich, P. (2023). *Use of temporal sensory evaluation methods with consumers: a position paper*. *Current Opinion in Food Science*, 54, 101102. <10.1016/j.cofs.2023.101102> <hal-04313440v1>

Visalli, M., Galmarini, M. V. *Multi-attribute temporal descriptive methods in sensory analysis applied in food science: a systematic scoping review*. *Comprehensive Reviews in Food Science and Food Safety*, 23, 1–45. <10.1111/1541-4337.13294>

Visalli, M., Cordelle, S., Béno, N., Schlich, P. (2024). *Product discrimination ability of temporal sensory evaluation methods used with consumers*. *Food Quality and Preferences*, 115, 105123. <10.1016/j.foodqual.2024.105123>

Submitted research articles

Visalli, M., Mahieu, B., Peltier, C., Cordelle, S., Schlich, P. *Temporal precision and resolution of sensory evaluation methods used with consumers*. Submitted to *Food Quality and Preference* (under revision).

Note: In order to make the manuscript more uniform, all references cited in the articles presented have been grouped into a single section, and all Tables and Figures have been renumbered.

Ethics statement

This research was carried out in conformity with the Declaration of Helsinki. In each presented study, the participants were informed of the conditions for participating and validated a consent form.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in the presented articles.

Chapter I.

Introduction

I.1. Sensory perception of foods

I.1.1. What is sensory perception?

Sensory perception is defined as "the organization, identification, and interpretation of sensory information in order to represent and understand the presented information or environment" (Schacter et al., 2011). Sensation (physical process) and perception (psychological process) are two distinct stages in the processing of human sensing. Sensory perception is the interpretation of a sensation by an individual: it is based on the activation of sensory receptors, but it occurs in the brain.

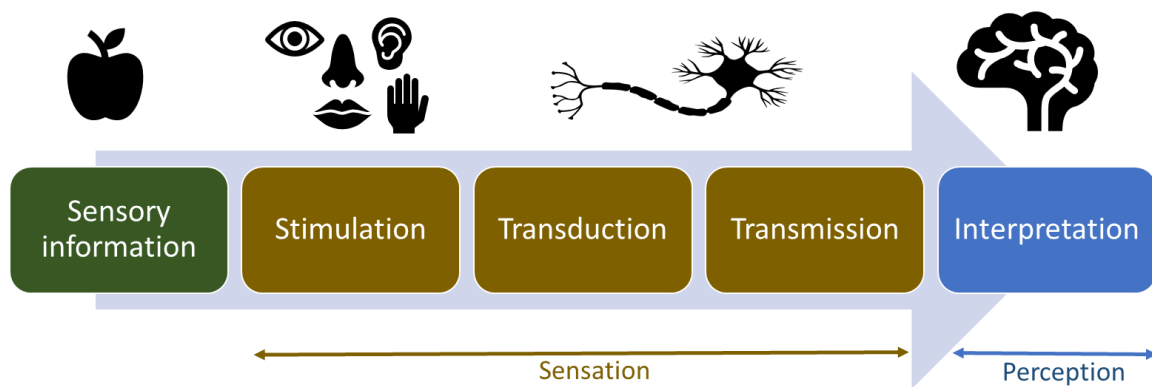


Figure 2. Sensory information processing, from physical stimulus to perceptual response.

Figure 2 shows the four stages of sensory information processing:

- Stimulation is the activation by stimuli of sensory receptors (chemoreceptors: taste, smell, etc.; mechanoreceptors: pressure, vibration, etc.; proprioceptors: spatial position, thermoreceptors: warming, cooling, etc., nociceptors: heat, crushing, etc.)
- Transduction is the translation of a sensory signal by the sensory receptor to an electrical signal in the nervous system.
- Transmission is the passage of encoded sensory information (type, location, duration, and relative intensity of the stimulus) to the central nervous system via the axons of the sensory receptors.
- Interpretation is the processing of incoming signals by the brain.

There are five basic human senses: touch, sight, hearing, smell and taste. Each sensory organ requires a minimal amount of stimulation to detect a stimulus. This absolute threshold can vary according to external and internal factors. The difference threshold is the minimum

required difference between two stimuli of different intensities to be noticed. Perception is multimodal: it depends on information provided by multiple sensory modalities at the same time combined into a single perceptual experience. It means that information from one sense can influence how information from another is perceived.

I.1.2. Why human sensory perception of food products is so important?

Sensory perception gives information about the environment and drives most behavioural decisions of animals related to survival, reproduction and behavioural adaptation (Bradbury & Vehrencamp, 1998). All animals possess gustatory sensing capabilities that are crucial for foraging as they enable to select the most nutritionally advantageous food sources (Demi et al., 2021). Indeed, the gustatory system acts as a screening mechanism that evaluates the quality of foods based on their chemical compositions and drives to consumption of energy-dense food and avoidance of potentially harmful food (Demi et al., 2021). Notably, omnivorous species had to develop complex feeding strategy due to the range of foods they were exposed. This led them to associate the food sensory properties with nutritional rewards and punishments that shaped food preferences (Breslin, 2013).

Sweet, salt and umami ("good tastes") generally stimulate consumptive responses for foods that supply compounds essential for growth and metabolism, while bitter and sour ("bad tastes") generate aversive responses to foods containing toxic substances (Breslin, 2013). These associations have remained anchored in taste genes of animals and humans (Wooding et al., 2021), and taste sensory inputs still consciously and unconsciously influence our behaviour toward foods (Breslin, 2013). For example, studies have demonstrated that sweet taste is innately liked by new-borns, while bitter taste is rejected (Steiner et al., 2001). This expected positive hedonic evaluation of food's sensory characteristics is called palatability (McCrickerd & Forde, 2016). Its influence on human appetite and food intake has been widely studied, and almost all studies have shown a positive correlation between food intake and palatability (Sørensen et al., 2003). Beyond palatability, perceptual feedback from smell, appearance, taste and texture have been shown to be involved in sensory-specific satiety, while the role of macronutrients and energy content of foods was less clear (Sørensen et al., 2003). Specifically, the smell and appearance of food stimulate a specific appetite and influence consumption decisions, taste informs the arrival of nutrients and the onset of satiety, texture moderates the eating rate and calorie ingestion (Forde, 2016). Even foods

similarly liked but presenting subtle differences in sensory qualities can influence everyday consumption differently (Forde, 2016).

Many foods in the Western diet, especially ultra-processed food, in addition to being convenient, shelf-stable, and affordable, are engineered by the food industry to be highly palatable. The repeated consumption of these “super-stimuli” is one of the main causes of obesity. Indeed, some studies have shown that the pathways activated in human brain in response to palatable foods are the same as those involved in drug addiction (Johnson & Wardle, 2014). The sensory qualities of food have been reported in many countries as the most important factor in food choices (Caltabiano & Shellshear, 1998; Glanz et al., 1998) and purchases (Allès et al., 2017; Glanz et al., 1998; Honkanen & Frewer, 2009; Januszewska et al., 2011; Milošević et al., 2012). This makes the study of the sensory response to foods central to understanding eating behaviours (Forde, 2016) and developing more sustainable diets (Aschemann-Witzel et al., 2019; Hoek et al., 2017; Knaapila, 2022).

I.2. Measuring food sensory perception

I.2.1. The periods of development of sensory-consumer science

Humans have been interested in describing food perception since ancient times (Shapin, 2016), but the development of sensory-consumer science was intrinsically linked to the industrialization of the food system (Lahne, 2016). Meiselman et al. (2022) have identified three time periods of development of sensory-consumer science: the early period (foundation), the rapid expansion period (formalisation), and the current period (consolidation).

Between 1940 and 1970s ("Period 1"), in a context of competition between manufacturers, producing food appealing to most consumers has become of primary importance (Shapiro & Hubert, 2005). History remembers as the foundation of sensory evaluation the development by the US army of acceptance-testing research to solve the problem of enlisted soldiers refusing to eat nutritious rations (Meiselman & Schutz, 2003). This led to the development of the nine-point hedonic scale, which is still in use. From there, the food industry began to support this new developing science, and various sensory evaluation methods were developed.

Sensory evaluation, attempting to become an objective science, has evolved into a discipline within food science, based on methodologies from experimental physiology, psychology, and psychophysics (Meiselman et al., 2022). A definition of sensory evaluation was proposed in 1975 by the institute of food technologists as "a scientific discipline used to evoke, measure, analyse, and interpret reactions to those characteristics of food and materials as they are perceived by the senses of sight, smell, taste, touch, and hearing". Stevens (1969) has demonstrated that the human subject could assign numbers reflecting the perceived intensity of stimuli. Thus, psychophysicists started to use direct methods of scaling to learn about sensory perception, considering the human as a sensory measuring instrument (Moskowitz, 2003). O'Mahony (1979) claimed that "a sensory panel is a flavour-measuring instrument. It consists of judges who use specified procedures for assessing flavours of foods. The data from a sensory panel have the status of data from any laboratory instrument." The psychophysical approach required all aspects of sensory evaluation to be under controlled conditions in order to produce results attributable to an experimental intervention (Lahne, 2016). Standardization of test conditions included panellist training, blind tasting in individual booths, maintain of constant parameters (lighting, temperature, humidity, sample size), etc. (Vilela, 2021).

Sensory evaluation methods have been proven performant tools for accurately determining the sensory characteristics of foods. However, the psychophysical approach assumes that the subject is a passive receiver of both the stimulus and the psychological image that comprises perception (Lahne, 2016). This implies that the sensory properties of a food product remain constant regardless of the context except for random variation, which does not reflect the empirical reality of human sensory perception (Lahne, 2016). Moreover, the perception of trained panels – calibrated to be objective measurement tools – can be very different from that of consumers. In this sense, it might be meaningless outside of the industrial context, when it comes to understanding everyday life experiences (Lahne, 2016). Thus, some researchers started to consider that research findings would be more relevant if the products were described by consumers (Meiselman et al., 2022).

Authors like Köster (2003) believed that testing food in the laboratory removed the natural consumption setting, hindered the ability of subjects to provide accurate data and threatened the ecological validity of the tests. Thus, during the 1980s and 1990s ("Period 2"), the need for methods beyond traditional descriptive methods motivated the development of new

descriptive sensory evaluation methods that can be used with consumers (Meiselman et al., 2022). Interest in research conducted out of the laboratory or in natural settings has increased, as have connections with consumer research on topics such as context, environment, attitudes or health (Meiselman et al., 2022). During this period, it has been recognized that the context in which food were consumed could have a significant impact on the perception of these foods (Stelick & Dando, 2018).

During Period 2, sensory evaluation also became more integrated with quality control and product development (Meiselman et al., 2022). The expansion of the discipline continued in the agri-food industry but also extended to many other sectors (e.g. pet food, textile, cosmetics, tobacco, automobile, etc.)

From the 2000s to today ("Period 3"), the period is marked by a growing interest in understanding foods from outside the industrial paradigm (Lahne, 2016). Sensory-consumer science is now characterized by a combination of commercial and academic research (Meiselman et al., 2022). Sensory evaluation is today part of the tools used to formulate better products from a nutritional perspective, but also to design interventions aimed at promoting public health (Forde, 2016) or favouring more sustainable diets (Meiselman et al., 2022).

1.2.2. Main categories of sensory evaluation methods

Even the most advanced analytical techniques of quantification of physicochemical properties of foods cannot replicate the complexity of human sensory perception (Forde, 2016). Thus, sensory evaluation relies on subjects' self-reported measurements (Torrico, 2021) expected to be representative of the sensory properties of products (Lahne, 2018). Sensory properties differ from material properties measured by instruments in the sense that they are not necessarily related to physical quantities (Chen, 2020). A sensory property can be defined as "the human physiological-psychological perception of a number of physical and other properties of food (product) and their interactions" (Rahman & McCarthy, 2009). A sensory property can be related to a current, a recalled or an expected perception (Cornelis et al., 2017).

Over the years, numerous sensory evaluation methods were developed. Sensory scientists from the psychophysical school commonly categorize them as "analytical" or "affective"

(Stone, 2018). International standards (ISO 5492:2008) use the wordings "objective" by opposition to "subjective".

Affective/subjective methods include hedonic tests (evaluation of the degree of liking of products) and preference tests (determination of preferred products). For affective tests, subjects should be representative of the population for which the product is intended and are therefore most often regular users of the category of product. Between 75 and 100 subjects ("consumer panel") are generally required (Stone, 2018).

Analytical/objective methods include discrimination tests (perception of differences between products) and descriptive analysis (characterization of products based on sensory intensities). For analytical methods, subjects need to be trained in order to be able to discriminate between samples of a product category (Stone, 2018). For descriptive analysis, between 10 and 12 subjects ("trained panel") are recommended, and for difference tests between 20 and 30 (Stone, 2018). During the training process, the subjects learn to use their senses and familiarize themselves with the products and methods until being repeatable (i.e. being able to make consistent and repeatable sensory assessments of various products - ISO 5492:2008). Examples of sensory evaluation methods for discrimination tests include triangle tests (Helm & Troille, 1946) or duo-trio (Peryam & Swartz, 1950). Examples of evaluation methods for descriptive analysis include Flavour profile™ (Cairncross & Sjostrom, 1950) or Quantitative Descriptive Analysis™ (QDA, Stone et al., 1974).

Although still in use, this classification has been challenged by the evolution of the sensory-consumer science field and the emergence of rapid methods (Delarue et al., 2015) in Period 2. Rapid methods enable sensory product characterization by untrained panels (including consumers), thus blurring the perimeter of analytical tests (Meiselman et al., 2022; Varela & Ares, 2012). Valentin et al. (2012) categorised rapid methods into three groups according to the psychological processes required from the subjects who evaluate products, namely verbal-based, similarity-based and reference-based methods. Verbal-based methods rely on the ability of subjects to describe their perception with words (Valentin et al., 2012), using pre-determined lists of sensory attributes (e.g. Check-All-That-Apply, CATA, Adams et al., 2007) or terms provided by the subjects themselves (e.g. free comment, FC, ten Kleij & Musters, 2003). Most of these methods rely on analysis of count data rather than intensity ratings. Similarity-based methods consist in holistic evaluations aimed at identifying similarities and differences

between samples (e.g. Sorting, Schiffman et al., 1981). Reference-based methods use reference samples against which other samples in the set are compared (e.g. Polarised Sensory Positioning, Teillet et al., 2010).

I.3. Temporal sensory evaluation methods

I.3.1. Why the temporal dimension of perception is important?

The transformation of foods in the mouth results in complex mixtures of information involving the senses of smell, taste and touch. This information is perceived independently and dynamically by the different senses which integrate it into a continuous perceptual output (Forde, 2016). Measuring dynamically this continuous perceptual output has been a challenge since Period 1. In order to measure perception as close as possible to the moment when the sensations are perceived, many temporal sensory evaluation methods were developed, in particular in Period 2, after the rise of computerized systems which facilitated the recording of time during data collection. Temporal sensory evaluation methods are sub-categories of analytical or affective methods aimed at measuring perception by considering the temporal dimension of the perceptual output, most often concurrently (during) to the tasting. The topic has become so important that entire books were devoted to temporal sensory evaluation methods (Castura, 2018; Kemp et al., 2017).

To date, only narrative reviews have been carried out on temporal evaluation methods (Cliff & Heymann, 1993; Devezeaux de Lavergne et al., 2017; Di Monaco et al., 2014; Dijksterhuis & Piggott, 2000; Fizman & Tarrega, 2018; Foster et al., 2011; Keefer et al., 2023; Schlich, 2017). Narrative reviews suffer from the absence of objective and systematic criterion for the selection of articles, which can bias the interpretation and conclusions of the authors (Pae, 2015). Conversely, systematic scoping reviews are structured using rigorous methodology, including an assessment of risk of bias, which makes the results transparent, reproducible and objective. Scoping reviews map the body of literature on a thematic area and present an overview of a broad selection of articles (Pham et al., 2014).

Section I.3.2 presents the protocol used to perform a systematic scoping review on multi-attribute temporal sensory evaluation methods used in food science. Section I.3.3 presents the results of the review and gives a detailed picture of the use of descriptive temporal sensory evaluation methods in food science.

I.3.2. Multi-attribute temporal descriptive methods in sensory analysis applied in food science: protocol for a scoping review

This article was *published in PlosONE*:
<https://doi.org/10.1371/journal.pone.0270969>

M. V. Galmarini: Conceptualization, Methodology, Writing - Original Draft, Visualization; M. Visalli: Conceptualization, Methodology, Writing - Original Draft, Visualization.

Sensory perception is a temporal phenomenon highly present in food evaluation. Over the last decades, several sensory analysis methods have been developed to determine how our processing of the stimuli changes during tasting. These methods differ in several parameters: how attributes are characterized (intensity, dominance or applicability), the number of attributes evaluated, the moment of sample characterization (simultaneously with the tasting in continuous or discrete time, or retrospectively), the required panel (trained subjects or consumers), etc. At the moment, there is no systematic review encompassing the full scope of this topic. This article presents the protocol for conducting a scoping review on multi-attribute temporal descriptive methods in sensory analysis in food science.

The protocol was developed according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) extension for Scoping Reviews checklist. The research question was "how have multi-attribute temporal descriptive methods been implemented, used and compared in sensory analysis?". The eligibility criteria were defined using the PICOS (Population, Intervention, Comparator, Outcome, Study design) framework. This protocol details how the articles of the final review will be retrieved, selected and analysed. The search will be based on the querying of two academic research databases (Scopus and Web of Science). The main topics reported in research involving sensory analyses methods will be identified and summarized in a data extraction form. This form (detailed in the protocol) will be used to report pertinent information regarding the objectives of the review. It could also be reused as a guideline for carrying out and reporting results of future research in a more standardized way. A quality appraisal process was derived from literature. It will be applied on the included articles of the review, and could also be re-used to ensure that future publications meet higher quality levels. Finally, for the sake of transparency, the limitations of the protocol are discussed.

1. Introduction

1.1. Background

As a result of mastication, bolus formation, contact with saliva and body temperature, food and beverage perception changes during oral processing and for this reason it is considered a temporal phenomenon. Sensory analysis aims at understanding the sensory perception of products by measuring subject responses. Thus, several temporal methods have been developed in the past 50 years trying to capture, study, describe and quantify these changes in perception. These methods and their main characteristics are listed in Table 1. One main reference is cited for each, in case the reader needs further detail and information.

The time lapse studied goes mostly from the moment in which the product gets in contact with the mouth until some seconds after swallowing. That is to say, the time it takes for a person to evaluate one intake of the product (e.g., one sip, one bite). However, food and beverage consumption have another temporality: bite after bite (or sip after sip) perception can also change due to cumulative sensory phenomenon. This type of temporality (multi-intake or full portion evaluation) was only recently studied and, even though many of the methods could be applied to this, only a few have been used.

Time Intensity (TI, Lee & Pangborn, 1986) was the first temporal method developed. It measures (by means of a scale) the intensity of one given attribute over a continuous period of time. It was conceived as a sort of temporal version of the Quantitative Descriptive Analysis (Stone et al., 1974) but allowing only to measure one attribute at a time. TI has long been the temporal method of reference, but it presents several limitations: measuring only one descriptor at a time, which results in halo-dumping effect (Clark & Lawless, 1994). Moreover, the "signature" effect (Dijksterhuis & Eilers, 1997) (evaluators have a characteristic shape of the curve) requires a higher training to reduce variability and obtain curves that respond to product characteristics and not to individual differences, resulting also in panellist fatigue (Dijksterhuis & Piggott, 2000). All other temporal methods have been developed trying to compensate for these limitations.

Table 1. Main sensory descriptive temporal methods published in peer-reviewed articles, in chronological order.

Name of the method	Abbreviation	Number of attributes	Attribute generation	Variable measured	Temporal resolution	Main reference
Time-Intensity	TI	1	Determined before the test (by the panel leader, etc.)	Intensity	Continuous	Lee & Pangborn (1986)
Intensity Variation Descriptive Methodology	IVDM	>1	Determined before the test (by the panel leader, by consensus, etc.)	Intensity	Discontinuous	Gordin (1987)
Discontinuous Time-Intensity	DTI	>1	Determined before the test (by the panel leader, by consensus, etc.)	Intensity	Discontinuous	Clark & Lawless (1994)
Progressive Profile	PP	>1	Determined before the test (by the panel leader, by consensus, etc.)	Intensity	Discontinuous	Jack et al. (1994)
Dual-Attribute Time-Intensity	DATI	2	Determined before the test (by the panel leader, etc.)	Intensity	Continuous	Duizer et al. (1997)
Temporal Dominance of Sensations	TDS-I TDS	>1	Determined before the test (by the panel leader, by consensus, etc.)	Dominance and intensity (TDS-I) or dominance only (TDS)	Continuous	Pineau et al. (2009)
Sequential Profile	SP	>1	Determined before the test (by the panel leader, by consensus, etc.)	Intensity	Discontinuous	Methven et al. (2010)
Multi-attribute Time-Intensity	MATI	>1	Determined before the test (by the panel leader, by consensus, etc.)	Intensity	Continuous	Kuesten et al. (2013)
Temporal Check All That Applies	TCATA	>1	Determined before the test (by the panel leader, by consensus, etc.)	Applicability	Continuous	Castura et al. (2016)
Attack Evolution Finish	AEF	>1	Determined during the test (free comment)	Dominance	3 periods (retrospective)	Visalli et al. (2020)
Free-Comment Attack Evolution Finish	FC-AEF	>1	Determined during the test (free comment)	Applicability	3 periods (retrospective)	Mahieu et al. (2020)

As in every quantitative method, the use of scales calls for trained assessors. But, unlike other Descriptive Analysis techniques, TI requires a higher concentration since the evaluator needs to be focused on the perception and changes in intensity for the given attribute over a period of time. This higher concentration and the continuous temporal manner of the measurement requires some extra training in comparison. In addition to the extra training sessions to manage this, measuring only one attribute increases the number of sessions needed if a multi-attribute temporal description of the product is required.

Aiming at reducing the number of sessions to attempt a description with more than one attribute, Dual Time Intensity (DATI, Duizer et al., 1997) and Multi Attribute Time Intensity (MATI, Kuesten et al., 2013) were developed. But they were not widely implemented, probably due to the difficulty of the task: quantifying different attributes over continuous scales at the same time. Discontinuous time alternatives have been proposed to simplify the process and enable the recording of intensities within a single bite or sip (“single-intake”) at uniform intervals steps or at specific moments using Intensity Variation Descriptive Methodology (IVDM, Gordin, 1987), Discontinuous Time-Intensity (DTI, Clark & Lawless, 1994), or Progressive Profile (PP, Jack et al., 1994), or over repeated or consecutive consumptions (“multiple-intakes”) using Sequential Profile (SP, Methven et al., 2010).

Methods developed after this tried to simplify the task by recording only qualitative data. Temporal Dominance of Sensations (TDS, Pineau et al., 2009) introduced the concept of *dominance* (different from intensity) asking the subject to choose (from a given list) the sequence of dominant sensations. As a matter of fact, when TDS was first presented, panellists were also asked to rate the intensity of the chosen attributes. It was soon deemed too difficult and the intensity scoring was disused. Temporal Check-All-That-Apply (TCATA, Castura et al., 2016) was developed as an alternative to TDS registering the presence/absence (“applicability”) of all attributes along time. TCATA added the time dimension to the static Check-All-That-Apply (CATA) method (Adams et al., 2007). After having been used with trained panels, TDS and TCATA were gradually more and more used with consumer panels allowing also to better understand preferences in addition to product description.

The newest methods, changed from the simultaneous tasting-evaluating paradigm and proposed a retrospective measure, describing the product right after tasting but taking into account the perceived temporality. Attack-Evolution-Finish (AEF, Visalli et al., 2020) methods

summarize perception as a sequence of three attributes corresponding to three subjective periods: “Attack”, “Evolution” and “Finish”. FC-AEF (Mahieu et al., 2020) mixed static free comment method (ten Kleij & Musters, 2003), AEF and applicability, allowing the subjects to characterize their temporal perception using their own words instead of predefined list of descriptors.

1.2. Rationale for conducting the review

As can be seen in Table 1, many temporal methods have been developed over the years. Despite their differences, they all aim at measuring the same phenomenon. In food science, Ares (2015) highlighted that the most important challenge for new methodologies for sensory characterization is the identification of their limitations. Although it was referring to non-temporal DA, it also applies to temporal descriptive methods. It has not been clearly established yet in which situations methodologies provide equivalent information and when their application is or is not recommended. Meyners (2020a) recently pointed out: *“Many (sensory and statistical) methods are developed and deployed, but they are rarely compared exhaustively and objectively with alternative existing methods. Why would I adopt any new method when I have something that currently (seemingly) addresses the same task in a similar way? What benefit does it bring, and is it important enough for me to bother? How can I make findings actionable to inform product design?”*.

However, to date, no such exhaustive comparison of the multi-attribute temporal sensory methods exists. Indeed, the searching methodology described in this article allowed finding several reviews on the subject, but none of them addresses all the existing multi-attribute temporal descriptive methods. Moreover, as it can be observed in Table 2, there is no systematic review on the topic.

Cliff & Heymann (1993) and Dijksterhuis & Piggott (2000) are obviously no longer up to date. Foster et al. (2011) and Devezeaux de Lavergne et al. (2017) mainly focus on applications in food oral processing research. Di Monaco et al. (2014), Schlich (2017) and Fiszman & Tarrega (2018) only reviewed studies related to TDS. By way of comparison, more than 350 articles (without TI) will be considered for inclusion in the review using the protocol described in this article.

Table 2. Existing reviews related to all temporal methods in sensory science.

Reference	Type of review	Scope	Number of references*
Cliff & Heymann (1993)	Narrative	TI	66
Dijksterhuis & Piggott (2000)	Narrative	TI, PP	56
Foster et al. (2011)	Narrative	TI, TDS in studies related to food oral processing	136
Di Monaco et al. (2014)	Narrative	TDS	43
Devezeaux de Lavergne et al. (2017)	Narrative	TI, TDS, TCATA in studies related food oral processing	112
Schlich (2017)	Narrative	TDS and variants	49
Fizman & Tarrega (2018)	Narrative	TDS, in studies related to texture	48

*Total number of references in each review, as reported in Scopus

Thus, the relevance of conducting a systematic review of the academic research on temporal sensory methods seems established to set guidelines based on scientific evidence. To overcome the limitations of the previous reviews, it is necessary to include all the temporal methods, with the exception of TI which is singular in the sense that it is the only one which characterizes a single attribute. As the heterogeneous nature of the studies was not amenable to a more precise systematic review, a scoping review will be carried out.

1.3. Objectives

The objectives of this review, in accordance with Daudt et al. (2013) are:

- (1) to map the scientific literature to make an exhaustive and objective inventory of the methods available for multi-attribute temporal descriptive sensory analysis of food products,
- (2) to clarify working definitions and inform practices in the field,
- (3) to summarize findings and recommendations based on (1) and (2) and to identify research gaps in the existing literature,
- (4) to disseminate research findings.

A specific focus will be accorded to methodology and articles comparing methods.

2. Materials and methods

The protocol was drafted in order to ensure that the scoping review will be conformed to the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist (Tricco et al., 2018).

This section was organized based on the five stages of a scoping review (Arksey & O'Malley, 2007):

- (1) identification of the research question,
- (2) identification of relevant databases and literature,
- (3) selection of articles,
- (4) data extraction,
- (5) summarization, interpretation and dissemination of the results.

Each stage is more detailed below in line with the objectives of the current scoping review.

2.1. Identification of the research question

2.1.1. Main research question and sub-questions

The main research question is: "how have multi-attribute temporal descriptive methods been implemented, used and compared in sensory analysis?"

The main research sub-questions concern:

Method and protocol

- What temporal sensory methods (and their variants) were used?
- What characterizes the method(s)?
- Which product categories were evaluated using the methods?
- What were the reported limitations, advantages and disadvantages of each method?
- How were the methods compared in terms of advantages and disadvantages?
- Are there any remaining unanswered questions related to the method and protocol implemented?

Data analysis

- How was the data collected with these methods analysed?
- How was performance (accuracy, discrimination, repeatability, reproducibility) measured?
- How were the methods compared in terms of their performances?
- Are there any remaining unanswered questions related to data analysis?

Contribution to existing scientific knowledge

- In which scientific fields (food science, psychology, etc.) were these methods applied?
- What type of temporal information was obtained with these methods?
- Did this information provide additional knowledge compared to other sensory (static measures, liking, etc.) or instrumental measures?
- Was the contribution used by other researchers? (quality, number of citations)

Demographics of the research

- What are the geographical and historical coverages of the methods?
- What are the main journals and authors implied?
- Is the scientific area FAIR? (standardization of metadata, open access, open data, ethical, etc.)

2.1.2. Inclusion and exclusion criteria

The PICO(S) (Population, Intervention, Comparator, Outcome, Study design) eligibility criteria (“Popping the (PICO) question in research and evidence-based practice,” 2002) for inclusion will be as follows:

Population. Any human panel (trained or semi-trained panellists, consumers) will be eligible for inclusion, without any limitation on its composition.

Intervention. Any study aiming to evaluate the sensory properties of food or drink using a multi-attribute temporal descriptive method will be relevant for inclusion. Any methodological article related to temporal data collection or statistical analysis of multi-attribute temporal descriptive methods will also be eligible for inclusion.

TI studies will be included at the identification and screening phases to present the magnitude of the use and study of the method in comparison to the multi-attribute temporal ones. However, they will not be considered thereafter for inclusion.

Studies exclusively related to temporality of preferences or emotions without consideration to product sensory descriptions will not be considered for inclusion.

Comparators. Studies with or without comparator will be eligible for inclusion.

Outcomes. Not applicable.

Study design. All types of study design will be eligible for inclusion.

To ensure that the articles will be available for the future readers of the review, only peer-reviewed articles having a DOI will be eligible for inclusion. In order to avoid article duplication reviews, opinion papers, congress proceedings, doctoral thesis and book chapters will be excluded. Only articles written in English will be considered.

Literature search will include published works until January 16, 2022.

2.2. Identification of relevant databases and literature

2.2.1. Information sources

The following databases were searched: Scopus, Web of Science Core Collection (WOS), and Google scholar (Martín-Martín et al., 2018). Only Scopus and WOS were retained (see 2.2.2 for details on this choice). Furthermore, we will search the reference lists and citing articles of included studies and related systematic reviews.

2.2.2 Search strategy

The Peer Review of Electronic Search Strategies (PRESS, McGowan et al., 2016) checklist was used to help constructing and validating the search strategy.

First, keywords were listed according to the PICOS criteria. No keyword related to population, comparators, outcomes and study design were added. Keywords related to intervention were identified, including the names of the published temporal methods (Table 1) established based on the knowledge of the authors: “progressive profile/profiling”, “sequential profile/profiling”, “temporal dominance of sensations”, “TDS”, “temporal check all that apply”, “TCATA”, “T-CATA”, “attack evolution finish”. Although TI was outside the scope of this research, the keywords “time intensity” and “time-intensity” were added to retrieve references related to DATI and MATI, but also to potentially detect articles citing TI and to get

an idea of the bibliographic volume related to this method. Keywords referring to multiple intakes were also added: “multi”, “multiple”, “bite”, “sip”, “intake”. The wildcard character (*) was used (when applicable) to ensure that variations of each keyword were found. The proximity operator was also preferred (when possible) to the “AND” operator to limit the number of results of the query.

Second, Scopus and WOS were queried on title, abstract and keywords, then results were quickly screened. The test query allowed to notice that the name of the method was not always present in the title, abstract or keywords. Thus, new generic keywords related to temporal perception of sensations (“temporal”, “dynamic”, “perception”, “description” and “sensory analysis”) were added and their combinations added to the search. The test query also allowed to identify major “subject areas” in Scopus (Agricultural and Biological Sciences; Chemistry; Psychology) and “research areas” in WOS (Food science; Chemical analytics; Psychology). They were later used to limit the number of results of the query.

Third, previous reviews related to the research question were identified (see Table 2). It allowed to check if the queries enabled to retrieve all cited references. To this end, no restriction on the type of article was added in the search. Several iterations were needed to refine the queries in WOS and Scopus. It was concluded that, with this search criteria, Google Scholar did not bring new references, and it was therefore not added as a database.

2.2.3 Research equations

The following research equations were retained for each database:

Scopus

```
TITLE-ABS-KEY((("progressive" PRE/1 "profil*") OR ("sequential" PRE/1 "profil*") OR ("dynamic" PRE/1 "profil*") OR ("time" near/1 "intensity") OR ("time-intensity") OR ("temporal" AND "dominance" AND "sensation*") OR ("TDS") OR ("temporal" AND "check" AND "all" AND "that" AND "apply") OR ("TCATA") OR ("T-CATA") OR ("attack evolution finish") OR ("intensity variation descriptive methodology") OR ("temporal" PRE/1 "profil*") OR ("temporal" PRE/1 "perception") OR ("temporal" PRE/1 "description") OR ("temporal" AND "sensory analysis") OR ("temporal method*") OR ("dynamic" PRE/1 "perception") OR ("dynamic" PRE/1 "description") OR ("multi*" PRE/1 "sip*") OR ("multi*" PRE/1 "bite*") OR
```

("multi" PRE/1 "intake*") AND (LIMIT-TO (SUBJAREA, "CHEM") OR LIMIT-TO (SUBJAREA, "AGRI") OR LIMIT-TO (SUBJAREA, "PSYC"))*

WOS

TS= (("progressive" NEAR/1 "profil") OR ("sequential" NEAR/1 "profil*") OR ("dynamic" NEAR/1 "profil*") OR ("time" NEAR/1 "intensity") OR ("time-intensity") OR ("TI") OR ("temporal" AND "dominance" AND "sensation*") OR ("TDS") OR ("temporal" AND "check" AND "all" AND "that" AND "apply") OR ("TCATA") OR ("T-CATA") OR ("attack evolution finish") OR ("intensity variation descriptive methodology") OR ("temporal" NEAR/1 "profil*") OR ("temporal" NEAR/1 "perception") OR ("temporal" NEAR/1 "description") OR ("temporal" AND "sensory analysis") OR ("temporal method*") OR ("dynamic" NEAR/1 "perception") OR ("dynamic" NEAR/1 "description") OR ("multi*" NEAR/1 "sip*") OR ("multi*" NEAR/1 "bite*") OR ("multi*" NEAR/1 "intake*"))*

2.3. Selection of articles

Following the search, all identified records will be collated into Mendeley citation manager and duplicates removed. The process of study selection will be presented using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR, Page et al., 2021) as in Figure 3.

The two authors will independently screen (i) the titles and (ii) abstracts retrieved from the database search for potentially eligible studies. The full texts of these studies will be obtained and further screened for eligibility based on the inclusion and exclusion criteria. Corresponding authors will eventually be contacted to retrieve articles not available. Potential disagreements regarding eligibility will be resolved through discussion and consensus. Reasons for exclusion of full-text assessed articles will be reported in the PRISMA diagram.

The articles cited in the reviews and included articles will be manually screened to look for potential relevant missing articles. The articles citing the included articles will be retrieved using Scopus and screened in the same way. The additional references not retrieved by the research equations will be reported in the PRISMA diagram.

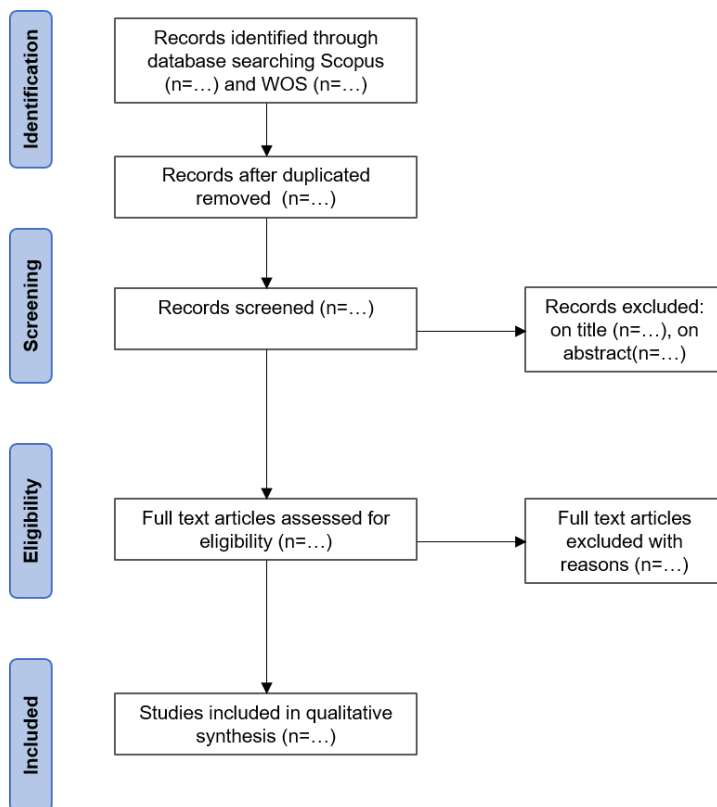


Figure 3. PRISMA flow chart to be completed in the final scoping review.

2.4. Extraction of relevant information and critical appraisal

2.4.1. Extraction of article metadata

The meta-data will be reported as exported from the databases.

Table 3. Meta-data exported from WOS and Scopus.

Meta-data	Description
DOI	Digital object identifier of the article
Title	Title of the article
Authors	Authors of the article
Abstract	Abstract of the article
Year	Year of publication of the article
Source title	Title of the peer-reviewed journal in which the article is published
Keywords	Keywords reported by the authors
Language	Language of the article
Subject/research area	Subject area of the journal
Number of references	Number of references cited by the article
Number of citations	Number of references citing this article
Open access to the manuscript	Type of open access

2.4.2. Identification of generic standards for reporting researching involving temporal sensory methods

A strategy was defined to identify relevant data in the most exhaustive way. As no comparable review exists, it was not possible to rely on previous similar research. Many standards have been developed to report quantitative or qualitative research using specific designs, but no guideline corresponded to the designs used for the studies that will be included in the review. Indeed, most of them used quasi-experimental research designs. “Quasi-experimental research is similar to experimental research in that there is manipulation of an independent variable. It differs from experimental research because either there is no control group, no random selection, no random assignment, and/or no active manipulation.” (Abraham & MacDonald, 2011). For product-oriented questions, the independent variable (the product) was most often studied using within-subject (counterbalanced or randomized) experimental designs, with no control group. For subject or method-oriented questions, factorial designs still with no control group were mostly used.

Thus, generic standards adapted for this review were identified from “Journal Article Reporting Standards for Quantitative Research in Psychology: The APA Publications and Communications Board Task Force Report” (JARS, Appelbaum et al., 2018). Applicable topics were identified from JARS (from part “Information Recommended for Inclusion in Manuscripts That Report New Data Collections Regardless of Research Design”). Some topics were renamed or grouped to be more consistent with the nomenclatures used in the articles in our area of interest. The expected content of each topic was completed based on literature when necessary. The result is summarized in Table 4.

Table 4. Generic standards applicable for reporting studies involving temporal sensory methods, adapted from JARS.

Topic	Expected content
<i>Meta-data</i>	
Title	Title should provide a concise description of the nature and topic of the study, including the name of the data collection method (O'Brien et al., 2014). General recommendations can be found in Hays (2010).
Abstract	Abstract should report objectives (state of the problem under investigation, main hypotheses); description of participants (pertinent characteristics); study method (research design, sample size, materials/methods used, outcome measured, data analysis procedures); main findings (including statistical significance levels); conclusions (beyond just results, reporting implications or applications). Adapted from Appelbaum et al. (2018).
Keywords	Keywords should be related to controlled vocabularies. Adapted from Ishida et al. (2020).
<i>Introduction</i>	
Problem	Introduction should describe the importance of the problem, including an overview of what is known about the problem, gaps in current knowledge and practical implications that make the study necessary. Adapted from O'Brien et al. (2014). Introduction should include a succinct review of relevant scholarship, including relation to previous work. Included references should be relevant to the problem studied, and cited in accordance with content and context. Self-citation should be justified. Adapted from Penders (2018).
Objective(s)	Introduction should report a statement framed as one or more research questions, purposes, goals, or objectives that should set readers' expectations for the methods, findings and discussion sections of the manuscript. Adapted from O'Brien et al. (2014).
<i>Materials and methods</i>	
Participants	A participants section should be reported, indicating: inclusion and exclusion criteria (if any); major demographic characteristics as well as important research-specific characteristics; recruitment procedure; settings, locations and dates for data collection; consent and retribution made to participants; institutional review board agreements; ethical standards met and safety monitoring (if any); intended and achieved sample size if different from intended; determination of sample size (power analysis or methods used to determine the number). Adapted from Appelbaum et al. (2018).
Data collection	Authors should describe in detail their data collection design and method(s) and justify them in relation to the research question(s). The authors should describe all instruments, guides, and protocols, including their development and cite relevant literature, theories or conceptual frameworks as appropriate. Adapted from O'Brien et al. (2014).

Data analysis	<p>Materials and methods should include a data analysis section describing the analytic process so that readers can follow the logic of inquiry from the research question(s) to the analysis and findings. The authors should cite the guiding literature and describe their processes in sufficient detail so readers can judge the extent to which the processes align with the guiding approach. If modification to or deviations from the guiding approach occurred, the authors should explain and justify these modifications. Adapted from O’Brien et al. (2014).</p> <p>Data analysis section should report: the variables measured and their nature (intensity, citation rate, etc.); the planned data diagnostics (criteria for post-data collection exclusions of participants); the criteria for deciding when to infer missing data and methods used for imputation of missing data; the definition and processing of statistical outliers; the analyses of data distributions; the data transformation to be used; the statistics methods used (including details of the models and references to the appropriate literature if required); the data analysis software. Adapted from Appelbaum et al. (2018). Before engaging in statistical inferences, level of expected statistical probability (e.g. $p < 0.05$, $p < 0.01$) should be established on the basis of reasonable knowledge of the phenomena under investigation and the caution necessary for interpreting comparisons (Norris et al., 2015).</p>
Results	
Synthesis	<p>Results should report information detailing the statistical and data-analytic methods, including: missing data; characterization of the data (n, means, standard deviations, etc.); inferential statistics (including exact p-values, minimally sufficient set of statistics needed to construct the tests: df, mean square effects, mean square error, etc.); reporting of any problems with statistical assumptions and/or data distributions that could affect the validity of findings. Adapted from Appelbaum et al. (2018).</p>
Findings	<p>Results should report evidence to substantiate the more general and abstract concepts or inferences presented as findings. Authors should report counter-examples and concrete details related to their findings. Judicious use of tables and figures can help communicate such findings. Adapted from O’Brien et al. (2014).</p>
Quality of measurements	<p>“Just because one obtains a graphical display or a series of tables with associated statistical significance does not mean it has any meaning or external validity” (Stone et al., 2012).</p> <p>For new methods, results should provide validity evidence (either directly in the study itself, e.g., via pilot testing, or indirectly based on previous research) supporting the use of the given measurement instruments for the intended construct interpretations (Norris et al., 2015). More on validity can be found in Heale & Twycross (2015).</p> <p>Results should also report estimates related to the reliability of measures (reliability estimates from other studies should only be used for comparison purposes). More on reliability can be found in Heale & Twycross (2015).</p>

<i>Discussion</i>	
Support of original hypotheses	Discussion should begin with a short summary of the main findings as a remainder for the readers helping them assess whether the subsequent interpretation and implications formulated are supported by the findings. Adapted from O'Brien et al. (2014). It should provide a statement of support or non-support for all hypotheses. Adapted from Appelbaum et al., (2018).
Connection to prior works	Discussion should elaborate on similarities and differences between reported results and work of others. Adapted from Appelbaum et al. (2018).
Interpretation & limitations	Discussion should provide an interpretation of the results and elaboration on findings in relation to the study purpose. Specific elements, decisions or events of the study that could influence interpretation should be identified. Adapted from O'Brien et al. (2014). Authors should also take into account: sources of potential bias; imprecision of measurement protocols; overall number of tests; adequacy of sample size (Appelbaum et al., 2018).
Contribution to the field	Discussion should consider contribution to the field (O'Brien et al., 2014), generalizability of the findings (taking into account target population and other contextual issues) and implications for future research (Appelbaum et al., 2018).
<i>Transparency</i>	
Declaration of interest	This section should identify (if applicable) any real or potential conflicts of interest that might have influenced or could appear to have influenced the research. Authors should also explain how these conflicts were managed in the conduct of the study, and describe the potential impact on study findings and/or conclusions (O'Brien et al., 2014).
Funding	This section should describe any sources of funding and other support for the study and the role of funders in data collection, data analysis and reporting, if applicable (O'Brien et al., 2014).
Contributions	This section should report the roles played by each author of the article, using Contributor Roles Taxonomy (Brand et al., 2015).

2.4.3. Definition of specific standards for reporting research involving temporal sensory methods

Fifty articles (called hereafter "test articles") identified thanks to the database query (see 2.2.3) were randomly selected, then read. Following this reading, new topics specific to sensory analyses studies were identified. To facilitate the work of extraction following the reading of each article, a data extraction form was derived from Table 3 and Table 4 as a list of topic-related questions. When possible, the use of closed questions was preferred to maximize the agreement between the reviewers and facilitate quantitative analysis of the data in the final review. When possible, a list of pre-determined answers was suggested (the list could be extended during the final evaluation process). When the content of the information was not related to one of the identified research questions but the presence/absence of the information still relevant to be reported for other purpose, the predetermined answers were "yes", "no" or "not applicable". The result is summarized in Table 5. The two authors will independently fill in the data extraction form materialized by an Excel sheet (that will be included as a supplementary material in the final review). As the nature of the answers is mainly objective, the differences in reporting will be resolved by checking the article until agreement between the two reviewers.

Table 5. Extraction form based on topic-related questions.

Item	Question
	<i>Predetermined answers</i>
<i>Meta-data</i>	
DOI	What is the digital object identifier of the article? <i>As reported in Scopus/WOS.</i>
Title	What was title of the article? <i>As reported in Scopus/WOS.</i>
Authors	Who are the authors of the article? <i>As reported in Scopus/WOS.</i>
Abstract	What is the content of the abstract of the article? <i>As reported in Scopus/WOS.</i>
Year	What is the year of publication of the article? <i>As reported in Scopus/WOS.</i>
Source title	What is the title of the peer-reviewed journal in which the article is published? <i>As reported in Scopus/WOS.</i>
Keywords	What are the keywords reported by the authors? <i>As reported in Scopus/WOS.</i>
Language	What is the language of the article? <i>As reported in Scopus/WOS.</i>
Subject/research area	What is the subject area of the journal? <i>As reported in Scopus/WOS.</i>
Number of references	How many references are cited by the article? <i>As reported in Scopus/WOS.</i>
Number of citations	How many references cite the article at the moment of the review? <i>As reported in Scopus/WOS.</i>
Open access to the manuscript	What is the type of open access? <i>As reported in Scopus/WOS.</i>
Open access to the data	Does the reader have access to the data, in a public repository or data paper? (extends "Open access to the manuscript"). <i>Yes/no.</i>
Open access to the source code	Does the reader have access to the source code used for data analysis? (extends "Open access to the manuscript"). <i>Yes/no.</i>
<i>Introduction/Problem</i>	
Review of scholarship*	Did the introduction include an overview of what is known of the problem based on a review of the scholarship (related to temporal sensory analysis)? <i>Yes/no.</i>

Relevance	Did the introduction identify gaps in current knowledge and/or practical implications that make the study necessary? <i>Yes/no.</i>
<i>Introduction/Objectives</i>	
Objective(s)*	What were the objectives of the research? <i>As reported by the authors/not reported.</i>
Area of knowledge	What was the area of knowledge produced by the research? <i>Methodological/product oriented/etc. (deduced from objective(s))</i>
<i>Materials and methods/participants</i>	
Selection criteria	Were the selection criteria of the participants reported? <i>Yes/no/not applicable.</i>
Recruitment modalities	Were the recruitment modalities reported? <i>Yes/no/not applicable.</i>
Determination of sample size	What was the criterion for determining sample size? <i>Literature/power calculation/not applicable/not reported.</i>
Number*	What was the final number of participants that actually participated in the evaluation? <i>As reported by the authors/not applicable/not reported.</i>
Demographics	Were the characteristics of the participants (at least one characteristic: age, gender, frequency of consumption, etc.) reported? <i>Yes/no/not applicable.</i>
Country	In which country did the experiment take place? <i>As reported by the authors (or deduced from context of the experiment, but not by authors affiliation)/not reported/not applicable.</i>
Location	Where did the data collection take place? <i>Lab/home/not applicable/etc.</i>
Ethics review board	Did the authors report an approval by an appropriate ethics review board? <i>Yes/No/Not applicable</i>
Participant consent	Did the authors report participants consent? <i>Yes/No/Not applicable</i>
Nature of the compensation	Did the authors report if there was or not a compensation for the participants? <i>Yes/No/Not applicable</i>
<i>Materials and methods/products</i>	
Description*	Did the authors report relevant information about the food products? (brand, recipe, composition, etc.) <i>Yes/no/not applicable.</i>
Type	What was the type of product? (deduced from product description) <i>Commercial/model (prepared by the experimenter)/not reported/not applicable.</i>
Food category	What was the category of the food products? (deduced from product description) <i>Wine/chocolate/not applicable/etc.</i>
Physical state	What was the physical state of the food products? (deduced from product description) <i>Solid/semi-solid/liquid/not applicable.</i>

Serving conditions	Did the authors report the relevant information (portion size, container, temperature, light, etc.) about the serving conditions? <i>Yes/no/not applicable.</i>
Information given to the participants	What was the information given to the participants? None if blindly evaluated. <i>None/Brand/Allegation/Price/Package/etc.</i>
Number*	How many different products/samples were evaluated? <i>As reported by the authors/not reported/not applicable.</i>
<i>Materials and methods/attributes</i>	
Selection	How were the attributes selected for the study? <i>By the panel/by another panel/literature/not reported/not applicable/etc.</i>
Description*	What were the names of the attributes? <i>As reported by the authors/not reported/not applicable.</i>
Sensory modalities	What were the sensory modalities evaluated? (deduced from description) <i>Basic taste/flavour/texture/mouthfeel/not applicable/etc.</i>
Definitions	Did the author report if attribute definitions were presented to the participants? <i>Yes/no/not applicable.</i>
References	Did the author report if attribute references were presented to the participants? <i>Yes/no/not applicable.</i>
Number	How many attributes were evaluated? <i>As reported by the authors/not reported/not applicable.</i>
<i>Materials and methods/research design</i>	
Object(s) of comparison	What was the object of interest? <i>Method/product/intake/subject/not applicable/etc.</i>
Temporal unit	What was the temporal unit of the measures? <i>Within-intake/between intakes/etc.</i>
Study design*	What was the study design used for comparing the experimental units? <i>Within balanced/within unbalanced/between/factorial/not applicable/not reported/etc.</i>
Product order	What was the experimental design defining the rank of presentation of the samples? <i>Balanced/randomized/not reported/not applicable/etc.</i>
Attribute order	What was the experimental design defining the rank of presentation of the attributes? <i>Balanced/randomized/not reported/not applicable/etc.</i>
<i>Material and methods/data collection</i>	
Temporal method(s)*	What was the name of the temporal sensory method/variant involved? <i>TDS/TDS intensity/TI/TCATA/TCATA fading/etc.</i>
Other measures	Was there other information collected with the temporal sensory data? <i>Liking/physicochemical measures/none/etc.</i>
Training	How were the participants introduced to the method? <i>Familiarization/training (hours)/not reported/not applicable.</i>
Type of panel*	What was the type of panel? (depending on the training of the participants, Barton et al., 2020) <i>Trained/semi-trained/expert/consumer/not reported/not applicable.</i>

Instructions	Did the authors report the instructions given to the participants? <i>Yes/no/not applicable.</i>
Warm-up	Did the tasting include a warm-up product prior to the evaluation of the samples? <i>Yes/no/not applicable.</i>
Software	Which was the software used for temporal sensory data collection? <i>As reported by the authors/not reported/not applicable.</i>
Number of evaluations	How many times were the samples evaluated (replicates)? <i>As reported by the authors/not reported/not applicable.</i>
Standardization of the tasting	How was the tasting standardized? <i>As reported by the authors/not reported/not applicable.</i>
Duration of the tasting	How long did the standardized tasting last? <i>Fixed duration/free duration (time)/not reported/not applicable.</i>
<i>Material and methods/data analysis</i>	
Justification of data selection	Did the authors justify data selection (subjects, products, attributes), if any? <i>Yes/no/not applicable.</i>
Data transformation	How was data transformation (if any) carried out? <i>Time standardization/periods/none/etc.</i>
Variables*	What were the analysed variables? <i>Durations of dominances/citation rates/intensities/not reported/etc.</i>
Statistics*	What were the statistical analyses used? <i>Curves/PCA/ANOVA/not reported/etc.</i>
Alpha	Was the level of expected statistical probability reported (previously to the results)? <i>Yes/no/not applicable.</i>
Software	What was the software used for data analysis? <i>R/SAS/etc.</i>
<i>Results/Synthesis</i>	
Characterization of data	Did the authors provide a characterization of the data? (missing data, descriptive statistics including n, mean, standard deviations, etc.) <i>Yes/no/not applicable.</i>
Inferential statistics	Did the authors report inferential statistics parameters when comparing objects of interest (products, methods, etc.)? (p-values, df, mean square effects, mean square error, etc.) <i>Yes/no/not applicable.</i>
<i>Results/Findings</i>	
Main findings*	What were the main findings related to temporal sensory analysis? <i>As reported by the authors/not reported.</i>
<i>Results/Quality of measurements</i>	
Validity	Did the authors report validity evidence supporting the use of the method? <i>Yes/no/not applicable.</i>
Reliability	Was a measure of reliability included? <i>Yes/no/not applicable.</i>

<i>Discussion</i>	
Answer to research question	Did the authors provide an answer to their research question? <i>Yes/no.</i>
Connection to prior works	Did the authors report connections to prior works? <i>Yes/no.</i>
Interpretation*	Did the authors provide an interpretation of the results in relation to the study purpose? <i>Yes/no.</i>
Limitations	Did the authors report elements that could have biased the result or influenced the interpretation? <i>Yes/no.</i>
Contribution to the field	Did the authors discuss the contribution to the field (generalizability of the findings and/or implications for future researches)? <i>Yes/no.</i>
<i>Transparency</i>	
Declaration of interest	Were the potential sources of influence on study conduct and conclusions reported? <i>Yes/no.</i>
Funding	Were the sources of funding and other support reported? <i>Yes/no.</i>
Contributions	Were the contributions of the authors included? <i>Yes/no.</i>

Items marked with an asterisk (*) are mandatory (except for statistical oriented articles) for the article to be considered for quality appraisal (see section 2.4.5)

2.4.4. Evaluation of completeness of data form

Once the data extraction for an article is filled, its completeness will be evaluated. Answers “no” and “not reported” will count as 0 while other answers will count as 1 for the purpose of summing across topics of Table 5. Nine completeness indicators will be computed in this way: “Introduction” (0 to 4), “Materials and methods/participants” (0 to 10), “Materials and methods/products” (0 to 7), “Materials and methods/attributes” (0 to 6), “Materials and methods/research design” (0 to 5), “Material and methods/data collection” (0 to 10), “Material and methods/data analysis” (0 to 6), “Results” (0 to 5), “Discussion” (0 to 5).

The evaluation of the completeness will not require any subjective judgment as it will be based on consensual answers of Table 5. Thus, the completeness indicators will be automatically computed based on the data extraction form thanks to an Excel formula.

2.4.5. Critical appraisal

Only peer-reviewed articles will be included in the final scoping review. Nonetheless, some authors stress the importance of a thorough quality assessment in scoping reviews (Carroll & Booth, 2015; Verhage & Boels, 2017). To achieve this objective, seven quality indicators (QI) were chosen according to (Dixon-Woods et al., 2004). It should be noted that only the articles that have the information identified as mandatory (see Table 5) will be evaluated at this stage. The articles that do not comply with this condition will be retained but marked as “not evaluated”.

To help the reviewer in their assessments, the QI derived from Dixon-Woods et al. (2004) were divided in topic-related questions summarized in Table 6.

The reviewers’ involvement will be added for the sake of transparency about the neutrality of the appraisal (it will be answered “yes” if at least one the reviewers is or was personally involved in the work or with the authors, considering this as potential bias in the quality appraisal).

Table 6. Indicators of quality of articles.

Quality indicator (QI)	Topic-related questions and related items
QI1 – Clear research question?	<p>QI1a – Was the introduction detailed enough to give an overview of the problem comprehensive for a competent but non-expert reviewer? (“Review of scholarship”).</p> <p>QI1b – Was the necessity of the study justified from the perspective of a competent but non-expert reviewer? (“Relevance”).</p> <p>QI1c – Did the stated objectives set readers’ expectations for the methods, findings and discussion? (“Objective(s)”, “Area of knowledge”).</p>
QI2 – Appropriate participants?	<p>Were the participants appropriate for answering to the research question? (“Selection criteria”, “Determination of sample size”, “Number”, “Location”, “Demographics”).</p>
QI3 – Appropriate design & data collection?	<p>QI3a – Were the research design parameters appropriate to answer the research question? (“Object(s) of comparison”, “Temporal unit”, “Study design”, “Product order”, “Attribute order”)</p> <p>QI3b – Was the choice of all the methods justified by literature and/or appropriate to answer the research question? (“Temporal method(s)”, “Other measures”)</p> <p>QI3c – Was the implementation of the method appropriate to answer the research question? (“Training”, “Type of panel”, “Instructions”, “Number of evaluations”, “Standardisation of the tasting”, “Duration of the tasting”)</p> <p>QI3d – Was the data collection process described in a way that makes the experiment reproducible?</p>
QI4 – Appropriate data analysis?	<p>QI4a – Was the analysis process made according to the literature or justified in case of a new approach? (“Data selection”, “Data transformation”, “Variables”, “Statistics”, “Alpha”, “Software”)</p> <p>QI4b – Was the data analysis process described in a way that makes it reproducible?</p>
QI5 – Claims supported by evidence?	<p>QI5a – Did authors make an appropriate synthesis of the results, including a judicious use of tables and figures allowing to characterize raw data and statistics? (“Characterization of data”, “Inferential statistics”)</p> <p>QI5b – Did authors report sufficient data and substantial evidence to support the findings? (“Main findings”, “Validity”, “Reliability”)</p>
QI6 – Integrated interpretations and conclusions?	<p>QI6a – Did authors appropriately discuss their findings in relation to the study objectives and prior works? (“Support of original hypotheses”, “Connection to prior works”, “Interpretation”)</p> <p>QI6b – Did authors objectively report sufficient elements pro and against their arguments? (“Limitations”)</p>
QI7 – Useful contribution?	<p>Did authors appropriately discuss the generalizability of their findings and/or implications for future research? (“Contribution to the field”)</p>

The two reviewers (R1 and R2) will independently answer to all topic-related questions, each having three possible answers, “yes”, “no” and “can’t tell” (Critical Appraisal Skills Program, Long et al., 2020). Then, the final judgement will be obtained as follows:

Step 1. For each topic-related question:

- R1 “yes” + R2 “yes” = “yes”
- R1 “yes” + R2 “can’t tell” = “probably yes”
- R1 “can’t tell” + R2 “can’t tell” = “can’t tell”
- R1 “no” + R2 “can’t tell” = “probably no”
- R1 “no” + R2 “no” = “no”
- R1 “yes” + R2 “no” = “disagreement”

Reasons for giving “no” as an answer will have to be justified.

Step 2. For the Quality Item the least favourable evaluation among the topic-related final evaluations will be retained.

For example, if reviewer 1 answered “yes” to QI6a while reviewer 2 answered “can’t tell”, QI6a final evaluation will be “probably yes”. If the two reviewers answered “can’t tell” to QI6b, QI6a final evaluation will be “can’t tell”. Considering the answers to parts a and b of QI6, its final evaluation will be “can’t tell”.

No overall score quality evaluation (taking into account the seven QI) will be made, and all articles will be included whatever their score. Indeed, these scores will only reflect a quality level related to the research question of this review, and the final appraisal will be to the discretion of the future readers of the scoping review. The evaluation of each article could thus be summarized as a 3-steps procedure (read below):

- Extraction of information (63 items) using the extraction form
- Automatic evaluation of the completeness of the reporting (9 indicators)
- Critical appraisal (7 indicators)

This evaluation process was refined and tested on the 50 test articles. The modifications to this evaluation process, if any, will be described in the final scoping review.

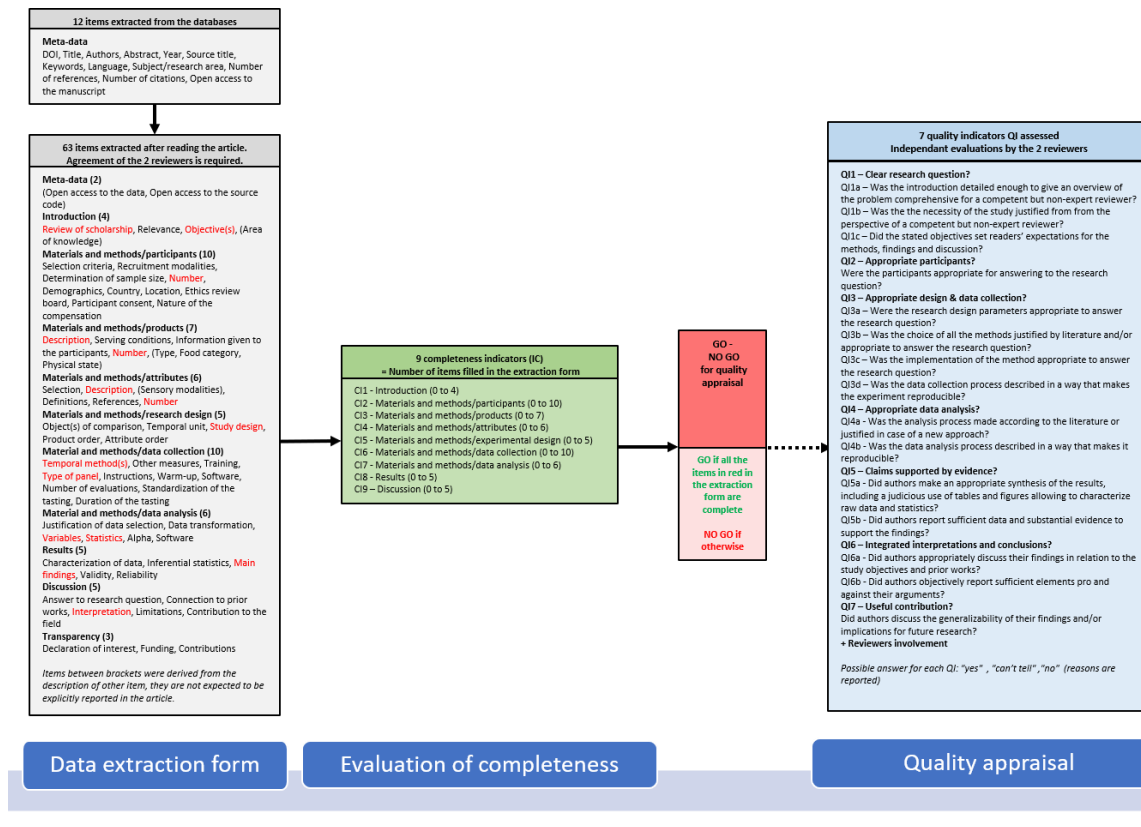


Figure 4. Summary of the process of evaluation of the articles.

3. Results

Tables and figures will be produced to summarize most of the extracted data presented in Table 5. A narrative synthesis organized into themes reflecting the scoping review objectives will also be presented. Both descriptive statistics and narrative synthesis will be supported by a qualitative analysis.

4. Discussion

The proposed scoping review resulting from the present protocol aims at summarizing how multi-attribute temporal descriptive methods have been implemented, used and compared in sensory analysis. This review will help the sensory analysts choose the appropriate method according to their needs (e.g. product description, product discrimination, oral processing behaviour) and adopt the best practices. Moreover, this review will allow to identify the areas where additional research and/or validation is necessary.

This protocol presents a detailed methodology for conducting the review. The publication of this protocol presents several advantages. The protocol has been reviewed thanks to expert feedbacks, ensuring its transparency and validity. The data extraction form and the quality

appraisal flowchart can be adapted and reused for other areas of research, particularly in a young science such as sensory and consumer science. This protocol could also be used as checklists to ensure no important information has been forgotten when writing new articles in this field. Indeed, the training exercise on the first 50 articles showed a great heterogeneity in the way the details of the studies are reported. This protocol would help standardizing the way the results of research are reported, which is important with the emergence of open science. Moreover, the authors would also point out that despite the fact that only peer-reviewed articles have been included, several studies do not meet the quality requirements that should be expected. This could be due to a lack of information or validated sources in the field. This protocol could also be used as a guideline for reviewing research manuscripts, or at least to point out the need for defining consensual criteria among journals. Finally, this protocol will promote the use of systematic reviews in science to inform the debate and improve the quality of future research.

Limitations in this protocol have to be reported. The research strategy largely depends on the name of the temporal methods. Therefore, the authors may have missed new or little-known temporal methods. Moreover, in order to limit the number of retrieved articles by the database queries, only specific subject areas (Scopus) or research areas (WOS) have been investigated. Again, it could result in missing articles published in annex research fields such as mathematics or computer science. However, these two limits are not that serious, because it is likely that the missed articles cite or have been cited by at least once one of the articles retrieved through the database queries. The additional references identified this way will be reported in the PRISMA diagram. The research has been limited to articles published in English language. A first look on the geographic distribution of the included studies suggests that it is not a concern. Only peer-reviewed articles have been considered for inclusions, which is debatable. This choice was made for the sake of feasibility and reproducibility. Moreover, it is unlikely that the excluded articles could have met the quality criteria, and the corpus of included publications is supposed large enough to answer to the research questions. Finally, only two databases have been queried. In a first time, Google Scholar was considered, but except for grey literature this database did not bring additional references. This seems to suggest that it was sufficient, considering all the limitations mentioned above.

I.3.3. Multi-attribute temporal descriptive methods in sensory analysis applied in food science: a systematic scoping review

This article was published in Comprehensive Reviews in Food Science and Food Safety: <https://doi.org/10.1111/1541-4337.13294>

M.V. Galmarini: Conceptualization, Data curation, Writing - original draft, Visualization; M. Visalli: Conceptualization, Data curation, Writing - original draft, Visualization.

The review includes articles published until December 2022, so some articles presented in this research work are part of it

Among descriptive sensory evaluation methods, temporal methods have a wide audience in food science because they make it possible to follow perception as close as possible to the moment when sensations are perceived. The aim of this work was to describe 30 years of research involving temporal methods by mapping the scientific literature using a systematic scoping review. Thus, 363 research articles found from a search in Scopus and Web of Science from 1991 to 2022 were included. The extracted data included information on the implementation of studies referring to the use of temporal methods (details related to subjects, products, descriptors, research design, data analysis, etc.), reasons why they were used and the conclusions they allowed to be drawn. Metadata analysis and critical appraisal were also carried out. A quantitative and qualitative synthesis of the results allowed the identification of trends in the way in which the methods were developed, refined and disseminated. Overall, a large heterogeneity was noted in the way in which the temporal measurements were carried out and the results presented. Some critical research gaps in establishing the validity and reliability of temporal methods have also been identified. They were mostly related to the details of implementation of the methods (e.g. almost no justification for the number of consumers included in the studies, absence of report on panel repeatability) and data analysis (e.g. prevalence of use of exploratory data analysis, only 20% of studies using confirmatory analyses considering the dynamic nature of the data). These results suggest the need for general guidelines on how to implement the method, analyze and interpret data, and report the results. Thus, a template and checklist for reporting data and results were proposed to help increase the quality of future research.

1. Introduction

1.1 Rationale

Sensory qualities of food have been reported worldwide as the most important factor in food choice (Caltabiano & Shellshear, 1998; Glanz et al., 1998) and purchase (Allès et al., 2017; Glanz et al., 1998; Honkanen & Frewer, 2009; Januszewska et al., 2011; Milošević et al., 2012). This makes the study of the sensory response to food and beverage central to the understanding of eating behaviors (Forde, 2016) and developing sustainable foods (Aschemann-Witzel et al., 2019; Hoek et al., 2017; Knaapila, 2022).

The transformation of food in the mouth results in complex mixtures of information involving the senses of smell, taste, trigeminal, ear and touch. Each sense perceives and translates the information independently and dynamically integrating it into a continuous perceptual output (Forde, 2016). Measuring this continuous perceptual output has been an objective since the early development of sensory science. Indeed, even the most advanced analytical techniques of quantification of physicochemical properties of foods cannot replicate the complexity of the human sensory perception (Forde, 2016). Thus, all sensory evaluation methods rely on subjects' self-reported measurements (Torrico, 2021) which are expected to be representative of the sensory properties of products (Lahne, 2018).

Providing data that meaningfully reflects a complex real-world experience using a task that subjects are able to perform successfully is challenging (Castura, 2018). Over the years, numerous sensory evaluation methods have been developed attempting to dynamically measure perception during the tasting of food products getting as close as possible to the moment when sensations are perceived. Entire books devoted to such temporal sensory evaluation methods were written (Castura, 2018; Kemp et al., 2017) but there are still gaps and needs for guidelines over different aspects that have not been covered.

To date, only narrative reviews have been carried out on temporal sensory evaluation methods (Cliff & Heymann, 1993; Devezeaux de Lavergne et al., 2017; Di Monaco et al., 2014; Dijksterhuis & Piggott, 2000; Fiszman & Tarrega, 2018; Foster et al., 2011; Keefer et al., 2023; Schlich, 2017), particularly on those focused on product description using more than one attribute (multi-attribute descriptive methods). Unlike narrative reviews, systematic scoping reviews are structured based on a rigorous methodology (Visalli & Galmarini, 2022) and

include a risk of bias assessment, which makes the results transparent, reproducible, and objective.

1.2 Objectives

Based on this methodology, the aim of this article is to draw up an exhaustive and objective inventory of the methods available for multi-attribute temporal descriptive sensory evaluation of food products. We chose not to consider single-attribute temporal descriptive sensory evaluation methods such as Time-Intensity (Lee & Pangborn, 1986) because of their specific use due to the limitation of measuring only one attribute at a time. Moreover, since Time-Intensity is simpler (only one attribute) and was developed decades before the multi-attribute temporal methods, it was already well documented (Cliff & Heymann, 1993, Dijksterhuis & Piggott, 2000) with a high agreement in the sensory community regarding its implementation and analysis. Thus, this review covers: (i) the development and use of multi-attribute temporal sensory evaluation methods, (ii) their implementation, (iii) the analysis of temporal sensory evaluation data, (iv) methods comparison and (v) the diffusion of results. Beyond informing actual practices and disseminating research findings, the ultimate objective is to identify research gaps in the existing literature and draw recommendations for future research.

2. Methodology

2.1 Protocol and registration

2.1.1 Summary of the original protocol

This scoping review was conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses guidelines for Scoping Reviews (PRISMA-ScR, Tricco et al., 2018). It was registered and published in PLoS ONE (Visalli & Galmarini, 2022). Hereafter is a summary of the original reviewed and published protocol (Figure 5).

The main criteria for inclusion were: research articles, available in Web of Science and/or Scopus, referencing multi-attribute temporal descriptive sensory evaluation methods, related to the evaluation of food and beverages (non-foods were considered out of scope) or to methodological/data analysis issues. Further details and research equations can be found in Visalli & Galmarini (2022).

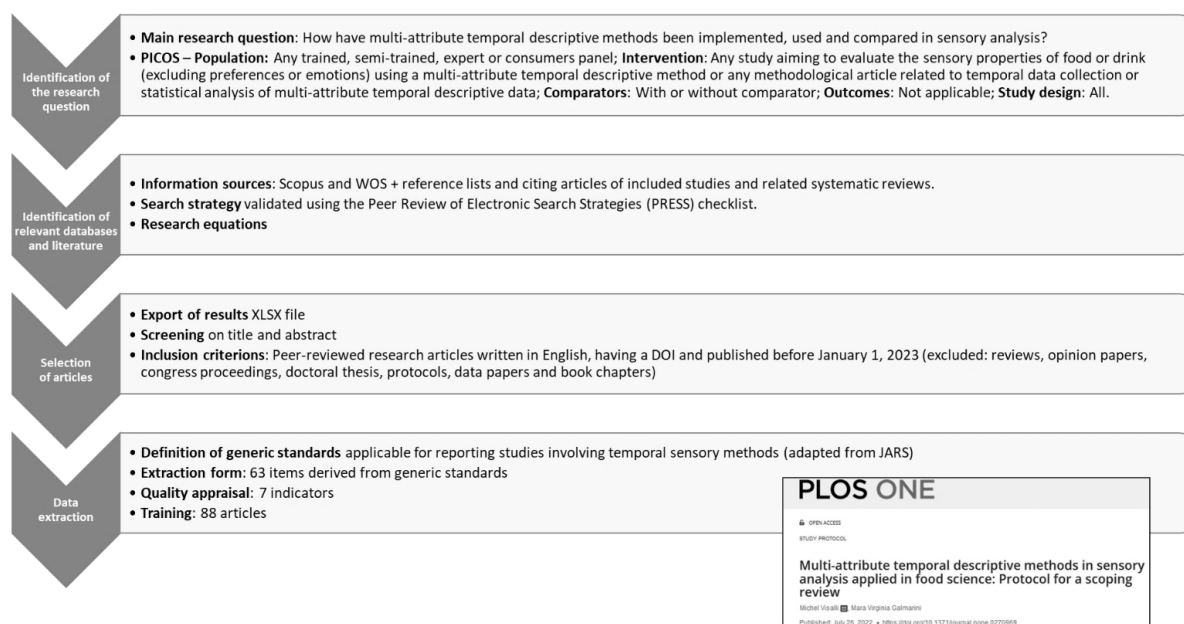


Figure 5. Main steps of the reviewing process.

2.1.2 Changes made to original protocol

Compared to the original protocol, some changes have been made. The period for article inclusion was extended to include all those published in 2022. The “Number of citations” was replaced by Field-Weighted Citation Impact (FWCI, Purkayastha et al., 2019). This metric is the ratio of the total citations actually received by the article and the total citations expected based on the average of the subject field. A FWCI lower than 1 means that the article is cited less than expected, while a FWCI greater than 1 means that the article is cited more than expected.

2.1.3 Quality appraisal of the included articles

The risk of bias was evaluated for each included article. For the purpose of transparency and reproducibility, the main reasons for “no” reported by the two authors for the seven quality criteria were analyzed (Q1-Q7, see Table 7), and specific reasons were determined by consensus for deciding not to check each quality criterion. If at least one specific reason was identified by the two reviewers, the quality criterion was not checked. Then, the number of quality criteria checked was computed for each article. It is to be reminded that the quality evaluation of the papers was made only in relation to the research question on multi-attribute temporal evaluation methods. Therefore, in those works were the multi-attribute temporal

methods were only a part of the paper, the quality appraisal was done in relation to this and not the whole paper.

Table 7. Reasons for not checking quality criteria.

Quality criterion	Main reasons for criteria not to be checked
Q1: Clear research question?	No justification of the use of the temporal method Missing key bibliographic references (e.g. reference to the temporal method) No justification of the relevance of the research
Q2: Appropriate participants?	Less than 20 evaluations (subject x replicates) of the products Lack of information on consumer panel (e.g. nothing about selection criteria, demographics, etc.) Number of evaluations not reported and number of subjects lower than 20
Q3: Appropriate design and data collection?	Product and attribute presentation orders not reported Instructions given to perform the temporal tests not reported Inappropriate design for treatment comparison (e.g. unbalanced within-subject design for method comparison) TDS-I with no later use of intensity scores Unclear tasting protocol making it non-reproducible Inappropriate attributes (e.g. “aftertaste” for TDS)
Q4: Appropriate data analysis?	Conclusions based only on exploratory analysis Unclear data analysis making it non-reproducible Unjustified data selection (e.g. subject removal) Inappropriate data transformation
Q5: Claims supported by evidences?	No substantial evidences or conclusion not congruent with findings Unjustified selective reporting (e.g. results not reported for some products) Low quality figures Errors in reporting of results
Q6: Integrated interpretation and conclusion?	No limitation reported No connection with previous works reported No discussion on the relationships between all data collected in case of multiple data acquisition (e.g. link between physico-chemical and temporal sensory description of products)
Q7: Useful contribution?	No contributions to the field reported

2.1.4 Data analysis

Data included in the extraction form (see Visalli & Galmarini, 2022) were grouped, counted and analyzed with pivot grid using Excel 365. All figures were plotted using Excel 365.

2.2 PRISMA diagram

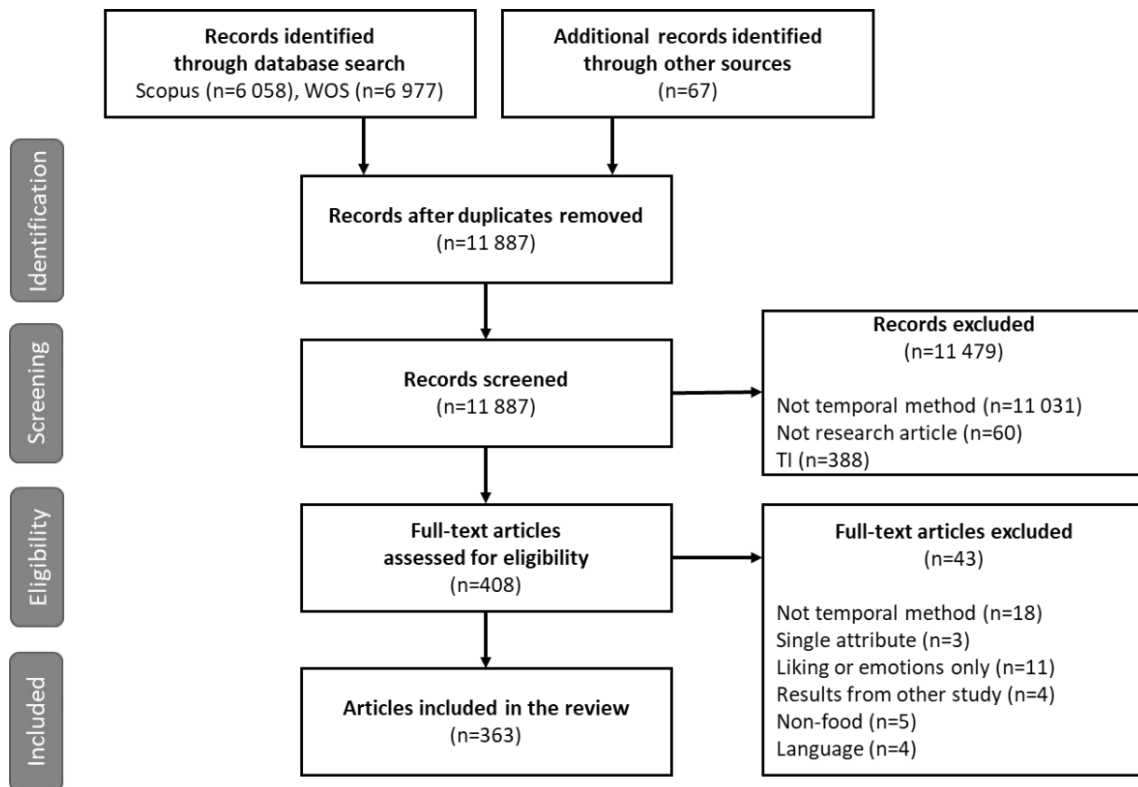


Figure 6. PRISMA diagram for article identification, selection and inclusion.

After the identification, screening and eligibility steps, a total of 363 published articles were included in this scoping review (Figure 6). It is to be noticed that the large number of studies found with the research equations was due to the use of acronyms and keywords having different meanings in the field of food science.

To contextualize the relative importance of temporal methods in the landscape of descriptive sensory evaluation methods, a research was made in Web of Science (limited to the field “Agricultural and Biological Sciences”) on Descriptive Analysis (DA) and Check-All-That-Apply (CATA, Adams et al., 2007), the most used static quantitative and qualitative methods. A total of 3247 articles were found for DA, and 637 for CATA. Even including Time-Intensity (TI, Lee and Pangborn, 1986), it can be roughly estimated that temporal methods were used in about 10% of the articles referencing descriptive sensory evaluation methods.

3. Results and Discussion

3.1 Included articles

Tables 8.A, 8.B and 8.C show the included articles, classified in three categories depending on their main objective: methodological (new method or variant, comparison of method, details of implementation), data analysis (including data processing) and other (all articles having the objective to gain knowledge on products, food oral processing, fundamentals of perception, etc.)

Table 8.A. Articles included in the present scoping review having a main objective related to methodological aspects, sorted by year. The type of panel (C – Consumers, E – Experts, T – Trained/semi-trained subjects, *Out of lab), the temporal method(s) and the temporal unit explores (B – Between intakes, W – Within intakes) are also indicated.

Reference	Objective of study	Product category	Type of panel ¹	Temporal method(s) ²	Temporal unit ³
Zwillinger & Halpern, 1991	Compare the temporal patterns revealed by taste-quality tracking with those previously indicated by taste-intensity tracking.	Model solutions	T	TQT	W
Jack et al., 1994	Develop appropriate sensory methodology to account for mastication allowing examination of textural transformations.	Cheese	T	PP	W
Duizer et al., 1997a	Compare DATI and TI on the same parameters; compare the perception of sweetness and peppermint; assess the effect of different combinations of sweetness and peppermint by means of DATI.	Chewing gum	T	DATI	W
Zimoch & Findlay, 1998	Determine the effectiveness of DATI for assessment of the temporal changes in meat texture, where the data collection time is relatively short; compare the relative effectiveness of DATI to SATI in assessing temporal differences in perceived tenderness and juiciness of meat.	Meat	T	DATI	W
Le Révérend et al., 2008	Evaluate level of similarity and assess potential application of two dynamic methods.	Beverage	T	TDS-I	W
Labbe et al., 2009	Compare the sensory characterization of nine gels done according to the sensory profiling method which was acquired in our previous study (Labbe, Gilbert, Antille, & Martin, 2007) and the TDS method; explore over time the impact of the odorant type (mint and peach) on taste and trigeminal perceptions.	Gel	T	TDS-I	W
Pineau et al., 2009	Describe how to represent TDS data (TDS dominance curves); compare the results given by TDS with those provided by TI.	Dairy product	T	TDS-I	W
Teillet et al., 2010	Evaluate and compare the contribution of classical methodologies in understanding the basics of the taste of water; propose a new sensory methodology well suited to the determination of the taste of water.	Water	T	TDS-I	W
Albert et al., 2012	Measure the dominant sensation at each point of the mastication process using the TDS technique and to compare the information obtained by TDS with that obtained by key-attribute sensory profiling of the same two commercial brands of prefried fish sticks cooked by three different cooking procedures.	Fish	T	TDS-I	W
Pineau et al., 2012	Investigate how the panellists use the attribute list during a TDS experiment, and whether the attribute list characteristics impact the panellists' response, in particular the number, the type and the position of the attributes in the list.	Not applicable	T	TDS	W
Paulsen et al., 2013	Investigate if TDS provides additional insight about drivers of liking in a preference mapping context compared to DA.	Combination of foods (salmon + sauce)	T	TDS-I	W
Kuesten et al., 2013	Explore the feasibility and usefulness of data collection using MATI; provide guidance to optimize the flavour and texture of the taffy based on the descriptive and consumer MATI results.	Candy	T	MATI	W
Vázquez-Araújo et al., 2013	Determine the suitability of TDS and TI as routine methods to evaluate beer; compare the obtained data with the data from a third technique: drinking/eating profile (drinking profile in this study: DP).	Beer	T	TDS ; PP	W ; B
Hutchings, Foster, Grigor, et al., 2014	Compare and validate the use of the TDS technique for food texture perception between a younger group and an older group; gain insight into how dynamic texture perception differs with age.	Nuts	T	TDS	W
Jager et al., 2014	Investigate the temporal aspects of food-evoked emotions during consumption of chocolates using temporal dominance of emotions (TDE); determine relationships between dynamic sensory and emotion perceptions by comparing the temporal dominance profiles of the sensory (TDS) and emotional (TDE) attributes.	Chocolate	C	TDS	W
Cheong et al., 2014	Understand the effect of completing the TDS task on basic oral processing behaviour with minimal training; determine the number of untrained assessors required to give a representative pattern of dominance and to assess if this number agreed with the numbers of subjects typically used in oral processing trials of food properties.	Biscuit	T ; C	TDS	W

Hutchings, Foster, Hedderley, et al., 2014	Evaluate differences in the use of TDS with food texture between a young (21–29) and an older (55–70) age group across a range of different foods (cheese, chocolate, gelatin gel and shortbread).	Cheese; Chocolate; Gels; Biscuit	T	TDS	W
Devezeaux de Lavergne et al., 2015	Compare QDA, TDS and PP in the assessment of dynamic texture of emulsion filled gels varying in fracture stress (low/high), fracture strain (low/high) and oil release (oil droplets bound/unbound to the gel matrix).	Gel	T	PP ; TDS-I	W
Thomas et al., 2015	Define the moment when the liking is formed in the mind of the taster during the tasting of a product and to develop a method based on TDS and liking data obtained from the same subjects; define “Temporal Drivers of Liking” (TDL) as sensory attributes inducing a positive or negative trend of liking when they become dominant.	Cheese	C	TDS	W
Ares et al., 2015	Investigate the TCATA method, through its application to products of varying complexity (yogurt, salami, cheese, orange juice, French bread, and marinated mussels) using consumers and trained panellists as assessors; deliver new methodological insights by comparing TCATA to TDS.	Yogurt; Deli meat; Cheese; Fruit juice; Bread; Mussel	T ; C	TDS ; TCATA	W
Castura, Antúnez, et al., 2016	Present the TCATA method, along with results from a TCATA study involving a trained panel.	Yogurt	T	TCATA	W
Galmarini et al., 2016	Compare the temporary descriptive results obtained by a reference method such as TI done by a trained panel to those acquired by a PP done by regular consumers in in-home conditions.	Chewing gum	C*	PP	W
Visalli et al., 2016	Compare TDS results obtained using a mouse or a touchscreen as pointing device for the selection of dominant attributes.	Chocolate	C	TDS	W
Rodrigues et al., 2016	Evaluate the behaviour of different panels with three different profiles: consumers, a selected panel based on their good sensory ability, and a selected and “familiarized” panel (introduced to the attributes involved in the analysis and conducting three preliminary tests before the final test), for TDS analysis of chocolates.	Chocolate	T	TDS	W
Ares et al., 2016	Present TCATA Fading; compare it to TCATA.	Bread; Deli meat; Cheese; Chocolate; Mussel	T ; C	TCATA ; TCATA-F	W
Lorido et al., 2016	Investigate whether the TDS method provides value in obtaining relevant temporal sensory information compared to the TI technique.	Deli meat	T	TDS	W
Jaeger et al., 2017	Evaluate the influence of a familiarization step on results from TCATA tasks with consumers.	Dried fruit; Nuts; Chocolate	C	TCATA	W
Ares et al., 2017	Compare TDS and TCATA in terms of their ability to identify the influence of the dynamic sensory profile of food products on consumer overall liking scores.	Bread; Chocolate	C	TDS ; TCATA	W
Alcaire et al., 2017	Compare static and dynamic sensory product characterization based on CATA questions with consumers.	Fruit juice; Yogurt; Dessert/cream/mousse	C	TCATA	W
Thomas et al., 2017	Record TDS and temporal liking simultaneously on the same screen from the same consumers in a single session during three intakes of a Gouda cheese.	Cheese	C	TDS	W ; B
Esmerino et al., 2017	Determine, based on consumer perception, the temporal aspects of different fermented dairy products (yogurt, fermented milk beverage, and fermented milk) using TDS, PP and TCATA with a comparison offered between methods.	Yogurt	C	TDS ; PP ; TCATA-F	W
Hutchings, de Casanove, et al., 2017	Understand changes in the responses of TDS panellists with milk protein hydrolysates as the same group of panellists progressed through different levels of training: as an untrained panellist, familiarized panellist, and trained panellist; examine the effect of increasing levels of a sweetener (sucralose) on the suppression of sensory attributes in milk protein hydrolysates during manipulation in the mouth and after swallowing using TDS.	Protein beverage/ONS	T ; C	TDS	W
Jaeger et al., 2018	Evaluate the influence of list length on TCATA and TCATA Fading tasks performed by consumers.	Fruit; Cracker	C	TCATA ; TCATA-F	W
Varela et al., 2018	Explore how trained assessors and consumers conceptualize dominance.	Bread	T ; C	TDS	W
Nguyen et al., 2018	Compare these three temporal methods (TDS, TCATA and M-TDS) based on detailed criteria consisting of dynamic profile, product trajectory and panel performance.	Yogurt	T	TDS ; TCATA ; M-TDS	W

Rodrigues et al., 2018	Evaluate the implications of analysing texture and taste sensations both simultaneously and separately on TDS descriptions for Prato cheese; compare the TDS descriptions obtained through both evaluations (TDS with different attributes modalities - taste and texture - evaluated in a same list and TDS with the attributes modalities evaluated separately) with the instrumental physico-chemical characterization.	Cheese	T	TDS ; M-TDS	W
Oliver et al., 2018	Compare the TDS methodology using an untrained consumer panel to the results obtained via QDA with a trained sensory panel.	Fruit	C	TDS-I	W
Silva et al., 2018	Explore the sensitivity of TDS, TDE, and TL, to differentiate two similar tasting wines.	Wine	C*	TDS	W; B
Jeltema et al., 2020	Develop a complementary method for working with consumers to understand their experiences as they ate food and to understand differences in experience among mouth behaviour (MB) groups, this includes (a) how texture differs over time for different MB groups and (b) how differing perceptions over time drive product acceptance.	Biscuit; Candy; Snack	C	Quessence	W
van Bommel, Stieger, Schlich, et al., 2019	Compare the performance of the Hold-down TD with the classic TD methodology for dynamic sensory and emotion profiling; identify ND duration periods in TD evaluations in an example product category (dark chocolates).	Chocolate	C	TDS ; TDS-HD	W
van Bommel, Stieger, Boelee, et al., 2019	Investigate the temporal dynamics of sensations and hedonic perceptions using multiple-intake assessment employing TDS and Alternated Temporal Drivers of Liking.	Yogurt	C	TDS	W; B
Vidal et al., 2019	Expand the methodological comparison of CATA and TCATA questions, focusing on the sensory characterization of solid products that require oral processing before swallowing.	Dried fruits; Nuts; Cheese; Bread; Chocolate	C	TCATA	W
Rizo et al., 2020	Assess the influence of fading duration on the TCATA Fading task and results.	Deli meat; Dessert/cream/mousse; Fruit juice	T	TCATA-F	W
Visalli et al., 2020	Compare the results from AEF and TDS studies	Chocolate	C	TDS ; AEF	W
Velázquez et al., 2020b	Evaluate the feasibility of using two dynamic methods (TCATA and TDS) for sensory characterization with children.	Dessert/cream/mousse	C*	TDS ; TCATA	W
Mahieu et al., 2020	Investigate whether consumers can successfully conduct an FC-AEF protocol at home and whether it enables the temporal characterization and discrimination of a set of products.	Chocolate	C*	FC-AEF-A	W
Weerawarna et al., 2021	Determine if data from consumer and expert TCATA panels for the same products are interchangeable; investigate if extending TCATA to multiple sips provides additional insights into the dynamics of the consumption experience.	Milk beverage	T ; C	TCATA-F	W; B
Crepalde et al., 2021	Integrate the sensory profile obtained by the two methodologies: ODP (optimized descriptive profile) and TDS, seeking to reduce the time required to obtain a more complete and concise sensory characterization	Fruit juice	T	TDS	W
Carrillo et al., 2021	Study how puréeing food changes flavour recognition and the sequence of flavour perception; rescue the flavour recognition, intensity, and sequence by adding flavours directly to the purées or isolated flavours incorporated in different emulsions; study the feasibility (setting and analysis) of the Free-TOS method.	Pasta; Pizza; Potato salad; Puree	C	F-TOS	W
Dinnella et al., 2022	Compare remote in-home to in-lab evaluations.	Chewing gum	T*	TDS	W
Visalli et al., 2022	Compare the sensory descriptions of commercial products collected from consumers using either a concurrent or a retrospective data collection method.	Beer	C*	AEF-A ; D-TCATA	W; B
Hutchings et al., 2022	Investigate the effect of definitions with three different groups of consumers having no prior TDS experience.	Chocolate	C	TDS	W
Keefer et al., 2022	Investigate the validity and sensitivity of the TR method; compare TR to the TCATA method.	Protein beverage/ONS	T	TR ; TCATA	W
Weerawarna et al., 2022	Study the effects of cup versus straw sipping on TCATA sensory profiles of vanilla milkshake in terms of citation proportions and product discrimination at the first sip; investigate the consistency of sipping method effects on citation proportions and product discrimination over multiple sips.	Milk beverage	T	TCATA	W; B

Table 8.B. Included articles having a main objective related to data analysis, sorted by year.

Reference	Objective of study	Product category	Type of panel ¹	Temporal method(s) ²	Temporal unit ³
Lenfant et al., 2009	Investigate whether differences and/or common patterns among the sensory trajectories of different products could be highlighted.	Cereal	T	TDS	W
Meyners & Pineau, 2010	Propose an appropriate and valid set of statistical tests that allow investigating overall differences as well as differences between pairs of products, and within each framework next to a general test, tests per attribute, per time point and per attribute by time point are suggested.	Cereal	Not reported	TDS	W
Meyners, 2011	Develop valid methods to study this important aspect of panel performance, i.e. disagreement with regard to differences between products.	Cereal	T	TDS	W
Dinnella et al., 2013	Propose ANOVA models on frequency values to estimate differences in attribute dominance among products; test the effect of collecting intensity ratings, along with TDS evaluations, on the frequency with which attributes were selected as dominant.	Coffee	T	TDS ; TDS-I	W
Lepage et al., 2014	Testing a protocol for assessing panel performance in TDS experiments together with the corresponding data analyses and keys for interpretation.	Bar	T	TDS	W
Meyners, 2016	Apply some of the methods we have proposed earlier for the analysis of TDS data (Meyners & Pineau, 2010) and amend them with additional analyses based on an aggregation of TDS data for certain time intervals into check all-that-apply (CATA) format data.	Cheese	T ; C	TDS ; TDS-I	W
Castura & Li, 2016	Use TDS sequences to characterize temporal changes in products; investigate TDS data, which requires that dominances be reported sequentially, for evidence that the sensations were at times perceived concurrently; link TDS dyads with TL data, and explore relationships between the TDS and liking data; to describe gaps and reiterations that exist in TDS and TL data that arise due to new test protocols.	Cheese	C	TDS	W
Rodrigues, Condino, et al., 2016	Explore methods for comparing products based on their TDS curves that simultaneously takes into account information about dominance rates for all the attributes over the time of evaluation.	Chocolate	T	TDS	W
Castura, Baker, et al., 2016	Propose exploratory data analysis and data visualization that can be used to gain insights and to test or generate new hypotheses.	Wine	T	TCATA	W
Galmarini et al., 2017	Propose an inferential approach based on individual total duration of dominant attributes, including also the multiple intakes as factors.	Wine	C	TDS	W; B
Vidal et al., 2017	Evaluate the influence of using data imputation to fill the gaps on results from TCATA fading.	Bread; Deli meat; Cheese; Chocolate; Mussel	T ; C	TCATA-F	W
Meyners & Castura, 2018	Propose a hypothesis tests for determining whether TCATA curves for a product of interest are significantly higher or lower than the chance line.	Wine	T	TDS ; TCATA	W
Lecuelle et al., 2018	Model TDS data using a more suitable stochastic approach.	Chocolate	C	TDS	W
Kuesten & Bi, 2018	Provide a novel Temporal Drivers of Liking technique which is based on some advanced statistical technologies and is used for MATI data.	Candy	T ; C	MATI	W
Kawasaki et al., 2019	Propose dominance-highlighted TCATA (dTCATA) curves. dTCATA curves provide visual information regarding temporal differences between the consumer panel TDS curves and the R&D panel TCATA curves.	Soup/Sauce	T ; C	TCATA ; TDS	W
Meyners & Castura, 2019	Explore the possibility of context-dependent significance thresholds which are customized for and aligned with research objectives.	Chocolate	Not reported	TDS ; M-TDS	W

Okada et al., 2022	Develop a method of causality modelling using perceptual and affective information that dynamically changes from the moment that a stimulus is experienced.	Fruit	C	TDS	W
Cardot et al., 2019	Introduce new statistical models based on finite mixtures of semi-Markov chains to describe data collected with the temporal dominance of sensations protocol, allowing different temporal perceptions for a same product within a population.	Cheese	C	TDS	W
Beaton & Meyners, 2020	Present two alternative CA-based methods: Canonical CA and Conditional CA.	Fruit juice	Not reported	TDS ; TCATA	W
Meyners, 2020	Provide more detailed power comparisons based on the time continuum instead of aggregating data into arbitrary time intervals, by refraining from using the significance lines for interpretation as recently discouraged by Meyners and Castura (2018a), and by directly testing statistically for differences between results from various methods rather than exploring (dis-)similarities between curves.	Yogurt	T	TDS ; TCATA ; M-TDS	W
Castura, 2020	Apply a graph theoretic approach to investigate perception dynamics in the data.	Yogurt	T	TDS ; M-TDS ; TCATA	W
Berget et al., 2020	Compare the TDS and TCATA by using multivariate methods to compare the data structures directly.	Yogurt; Cheese; Bread	T	TDS ; TCATA	W
Okamoto et al., 2022	Validate the prospect of principal motion analysis (PMA) of TD data.	Fruit	C	TDS	W
Kurata et al., 2022	Obtain the transition process of attributes (tastes and mouthfeels) in the oral cavity, to express the tendency of dominance durations of attributes, and to specify factors (such as sex, age, food preference, dietary habits, and sensitivity to a particular taste) affecting dominance durations, simultaneously.	Milk beverage	C	TDS	W
Dietz, Yang, et al., 2022	Compare the impact of raw, non-processed data with different approaches to time standardize the data on the temporal sensory profiles obtained from a preceding TCATA study.	Beer	T	TCATA	W
Okamoto, 2021	Propose a bootstrap resampling method for estimating the uncertainties of TDS curves, demonstrate the method's validity, and calculate the necessary sample size.	Not applicable	C	TDS	W
Peltier et al., 2022	Propose a methodology to relate in vivo PTR-MS data and temporal sensory data obtained by TDS or TCATA including pre-treatment taking breathing into account and statistical analyses based on RGCCA.	Chocolate	T	TDS ; TCATA	W
Frascolla et al., 2022	Introduce a statistical testing procedure based on the likelihood ratio between two semi-Markov processes.	Chocolate; Cheese	C	TDS	W
Castura et al., 2022	Evaluate new and existing methods to explore the uncertainty in the PCA scores.	Wine	T	TCATA	W
Derks et al., 2022	Undertake a time-discrete bucketing approach by sacrificing temporal (time) resolution in order to gain the signal (counts) resolution.	Yogurt	T	TCATA	W
Shimaoka et al., 2022	Investigate the temporal evolution of the primary-sensory and multi-sensory responses toward strawberries and statistically linked these responses by Canonical Correlation Analysis.	Fruit	C	TDS	W

Table 8.C. Included articles having a main objective related to knowledge on product, food oral processing, fundamentals of perception, etc.

Reference	Objective of study	Product category	Type of panel ¹	Temporal method(s) ²	Temporal unit ³
Pierguidi et al., 2021	Describe the temporal evolution of the intensity of dominant sensations in alcoholic cocktails prepared with different aperitif liquors (same typology, different brand) considered as a case study of products with subtle sensory differences.	Alcoholic cocktail	T	TDS ; DTI	W
Martínez-García et al., 2020	Determine physicochemical characteristics and the dominant attributes of amaranth bars made with different sweeteners as substitutes for the mixture of brown sugar and sugar commonly used as binding agents in the manufacture of commercial amaranth bars.	Bar	T	TDS	W
Keefer et al., 2020	Characterize the influence of sweetener and protein source on flavour, texture, and shelf life of high-protein low-carbohydrate bars.	Bar	T	TCATA	W
Medoro et al., 2016	Compare the scientific profiles obtained with the description given by an expert beer taster in order to check for the correspondence between the trained panel perceptions and the expert description.	Beer	T	TDS	W
Missbach et al., 2017	Assess the most dominant flavour attributes across beers with varying alcohol content.	Beer	T	TDS	W
Simioni et al., 2018	Assess the use of a multiple-sip TDS method and acceptance test to evaluate four different specialty beers with the aim of contributing to a greater understanding of consumer acceptability of products over repeated consumption.	Beer	C	TDS	W; B
Ramsey et al., 2018	Evaluate the influence of ethanol on consumer liking of lager and perception of its sensory characteristics; determine if particular time points during temporal liking related to overall liking; investigate the relationship between the temporal sensory profile of beer and temporal liking data identifying critical attributes driving consumer acceptance of beer in relation to ethanol concentration.	Beer	C	TCATA	W
Mitchell et al., 2019	Use TCATA to determine how beer perception varies with thermal taste status and investigate the effects of serving temperature and a concurrent auditory cue on TCATA responses and how these may interact with TTS.	Beer	T	TCATA-F	W
Silva et al., 2019	Investigate how dynamic sensory and emotion profiles interact and how this relates to temporal liking.	Beer	C*	TDS	W; B
Wakihira et al., 2020	Establish a new method to measure “the will of drinking”, which can be practically and effectively used in a wide range of consumer tests in ad libitum consumption settings; clarify the factors that affect beer drinkability, which could be achieved by investigating the effects of manipulating explanatory variables such as ingredients, recipes and brewing technologies.	Beer	E ; C	TDS	W; B
Ramsey et al., 2020	Explain the orthonasal and retronasal differences in consumer perception of standard and NAB.	Beer	C	TCATA	W
Dietz, Cook, et al., 2022	Establish a TCATA by modality approach for the temporal sensory evaluation of complex beverages characterized by lingering multi-modal profiles.	Bees	T	M-TCATA	W; B
Machado et al., 2023	Use dynamic sensory analysis methods to evaluate and better understand the consumer preferences for two beer styles enriched with elderberries.	Beer	C	TDS	W; B
Laguna et al., 2013	Conduct an in-depth study of the mechanical phenomena occurring in mouth during the oral processing of biscuits using TDS; study the impact of fibre addition and fat reduction on the perception of texture over time during in-mouth handling; achieve a better understanding of the factors affecting consumer acceptance.	Biscuit	T	TDS-I	W
Young et al., 2013	Investigate the links between the evolution of the changing biscuit bolus structure (at the macroscopic scale) throughout mastication, the associated changes in perceived texture, and the associated changes in mechanical and rheological properties.	Biscuit	C	TDS	W
Young et al., 2016	Investigate the subjective nature of food texture perception and the threshold for swallowing by following changes in perceived sensory texture, bolus structure, and bolus MC throughout oral processing.	Biscuit	T	TDS	W
Rizo, Jimenez-Pérez, et al., 2019	Assess the impact of TDS evaluations of texture and flavour on oral activity, measured with a 3D motion capture recording system, during the consumption of two different chocolate-chip biscuits.	Biscuit	T	M-TDS	W
Le Calvé et al., 2019	Assess the sensory impact of fat reduction in biscuits.	Biscuit	T	TDS	W

Panouillé et al., 2014	Understand bread bolus formation and texture and salty perception from three breads that varied in structure and/or composition.	Bread	T	TDS	W
Jourdren et al., 2016	Identify the respective contributions of bread and bolus properties to dynamic texture perceptions during consumption.	Bread	T	TDS ; PP	W
Gao et al., 2017	Understand the impact of bread structure and its dynamic transformation on people's chewing behaviour and texture perception.	Bread	T	TDS	W
Nguyen et al., 2017	Explore the role of texture of solid foods in consumers' perception and expectations of satiation and satiety, in particular the role of dynamic perception during oral processing, with barley bread as a case study.	Bread	T	TDS	W
Jourdren et al., 2017	Better understand the role of oral processing in volatile release and dynamics of aroma perception; clarify how bread structure and volatile profiles influence the dynamics of perceived aroma.	Bread	T	PP	W
Jourdren et al., 2017	Gain insight into the effect of bread crumb and crust structure on VOC release and aroma perception over the course of oral processing.	Bread	T	TDS	W
Gao et al., 2018	Reveal the impact of changes in bread structure and texture on the eating quality of baked bread and steamed bread.	Bread	T	TDS	W
Antúnez et al., 2018	Explore the effect of partially replacing NaCl with KCl on the sensory characteristics and hedonic perception of toast bread.	Bread	T	TCATA	W
Pu et al., 2019	Investigate the aroma release and perception from white bread during oral processing by GC-IMS and dynamic sensory evaluation of TDS.	Bread	T	TDS	W
Puerta et al., 2020	Describe the oral trajectory (fragmentation and bolus properties) of gluten-free commercial breads in relation to the dynamics of sensations perceived during its consumption, to better understand the factors involved in the perception of different texture sensations.	Bread	T	TDS	W
Pu, Duan, et al., 2021	Characterize the dynamic texture perception and study the mechanisms occurring in bolus from chewing to swallowing during white bread oral processing.	Bread	T	PP	W
Pu, Zhang, et al., 2021	Investigate the process of taste release and perception during chewing bread using high-performance liquid chromatography and sodium ion-selective electrode analysis combined with dynamic sensory evaluation.	Bread	T	TDS ; PP	W
Puerta et al., 2021	Study how structural changes in gluten-free breads, caused by different baking conditions (level of dough hydration and fermentation time), impact oral processing (bread breakdown and bolus formation) and sensory perception.	Bread	T	TDS	W
Menegali et al., 2020	Investigate the effect of pink pepper extract and different types of packaging on the oxidative stability, colour, pH and temporal and hedonic sensory perception of chicken burger.	Burger	C	TDS	W
Rios-Mera et al., 2020	Investigate changes in the instrumental texture and temporal sensory profile, and to determine the TDL of beef burger, due to the reduction of the content and size of NaCl.	Burger	C	TCATA-F	W
Barker & McSweeney, 2022	Determine the consumer acceptance of HMBs made from chicken and yellow field peas.	Burger	T	M-TCATA-F	W
De Souza et al., 2013	Determine the equivalent amount of different sodium chloride replacements (potassium chloride, monosodium glutamate, potassium phosphate, magnesium chloride, calcium chloride, calcium lactate and potassium lactate) required to promote the same degree of ideal saltiness in butter; study the sensory profile of sodium chloride and the substitutes using TDS.	Buttes	T	TDS	W
Gonçalves et al., 2017	Use acceptance test and temporal sensory technics (TI and TDS) to evaluate sensory margarines with reduced sodium content prepared using a salt mixture consisting of NaCl, KCl and monosodium glutamate at different concentrations.	Butter	C	TDS	W
Souza et al., 2022	Evaluate cocoa bean shell powder (CSp) in terms of physical and chemical characteristics; use CSp to make a cake, followed by evaluation of its physical, chemical and sensory characteristics.	Cake	T	TDS	W
Mayhew et al., 2017	Quantify the temporal contribution of stickiness-relevant texture attributes to overall stickiness perception in a caramel system; relate sample Tg to oral and tactile texture properties in a caramel system.	Caramel	T	TDS	W
Mayhew et al., 2018	Relate consumer and trained panellist perceptions of stickiness in a caramel system; identify texture attributes that contribute to stickiness perception.	Caramel	T	TDS	W
Peyron et al., 2011	Analyze the physical properties of the bolus at various time points in the chewing process.	Cereal	C	TDS	W

Deegan et al., 2013	Investigate the sensory and chemical consequences of Emmental cheese made using a pre-processing routine involving homogenization; investigate the potential of TDS as a sensory tool in the evaluation of cheeses made with the pre-processing routine.	Cheese	T	M-TDS	W
Da Silva et al., 2014	Determine the equivalence/salting power of different salt substitutes (potassium chloride, magnesium chloride, potassium phosphate, potassium lactate, calcium lactate and monosodium glutamate) in relation to cream cheese with sodium chloride; analyze the sensory profile of cream cheese with these substitutes.	Cheese	T	TDS	W
Rodrigues et al., 2014	Develop an accepted Mozzarella cheese with reduced sodium content, using a mixture of salts consisting of NaCl, KCl, and MSG in different concentrations, through acceptance testing and temporal sensory evaluation.	Cheese	C	TDS	W
Saint-Eve et al., 2015	Gain insights in texture perception by considering the dynamic aspects of consumption.	Cheese	T	TDS ; PP	W
Bemfeito et al., 2016	Characterize the dynamic sensory profile of artisanal Minas cheese produced in the Serra da Canastra region through TDS and sensory acceptance tests and verify the attributes that lead to product quality.	Cheese	T	M-TDS	W
Silva et al., 2018	Evaluate the effect of sodium reduction and the addition of flavour enhancers on the dynamic and static sensory profile of probiotic Prato cheese using both the TDS and QDA methods.	Cheese	T	TDS	W
Ningtyas et al., 2019	Investigate the influence of fat replacers (β -glucan and phytosterols) on reduced-fat cream cheese texture by TDS and instrumental analysis with the purpose of gaining a better understanding of the changes in cream cheese texture as perceived during oral processing.	Cheese	T	TDS-I	W
Mendes et al., 2019	Assess the effects of the β -CN A2 on sensory characteristics of Petit Suisse and Minas cheeses.	Cheese	C	TDS	W
Bord et al., 2019	Characterize the flavour properties of raw and heated cheeses using TDS and DA in order to explore the impact of heating on the taste of blue cheeses; investigate the relative benefits of using these two methods.	Cheese	T	TDS-I	W
Ramos-Gabriel et al., 2019	Evaluate the impact of fermented whey in the sensory characteristics of ripened cheeses and the consumer preference.	Cheese	C	TDS	W
Santagiuliana et al., 2020	Use TDS to assess whether the dominance of positive sensations caused by the addition of macroparticles to quark containing microparticles can be increased so that the dominance of undesired gritty sensations is suppressed.	Cheese	C	TDS	W
Shahbazi, Jäger, & Ettelaie, 2021	Reduced-fat casein-based emulsion stabilized by acetylated MCC can be effectively printed in a 3D fashion with stable structural properties.	Cheese	T	TDS-I	W
Aguayo-Mendoza et al., 2021	Determine the influence of physical-chemical properties of heterogenous foods on oral processing behaviour, bolus properties and dynamic sensory perception.	Cheese	T	TDS	W
Rodrigues et al., 2022	Identify the main texture and flavour sensations that drive the Prato cheese preference using the TDS and CATA methodologies.	Cheese	T	TDS	W
Duizer et al., 1996	Determine the release time characteristics that produced the longest duration and maximum intensity of both tastes, with greatest reliability.	Chewing gum	T	DATI	W
Pedrotti et al., 2019	Study the impact of ethnicity, gender and physiological parameters, on flavour perception and release of mint chewing gum through in nose PTR-MS approach coupled to DTI sensory analysis and collection of physiological parameters like stimulated salivary flow, volume of oral cavity and fungiform papillae density.	Chewing gum	T	PP	W
Goza et al., 2022	Characterize the ability of a starch-limonene IC flavour release system to alter temporal flavour perception in a chewing gum using human participants; examine how individual differences in sAA and sFR impact temporal flavour perception of these starch-guest ICs.	Chewing gum	C	TCATA ; D-TCATA	W
Batista et al., 2016	Investigate the effects of a mixed starter culture of yeasts (<i>S. cerevisiae</i> , <i>P. kluyveri</i> , and <i>H. uvarum</i>) on the microbial communities and volatile compounds of cocoa fermentation; evaluate the volatile compounds and TDS of resulting chocolate.	Chocolate	T	TDS	W
Menezes et al., 2016	Evaluate the chemical and sensorial characteristics of chocolate produced from four different cocoa varieties (CCN51, PS1030, FA13, and CEPEC 2004) inoculated with <i>S. cerevisiae</i> CA11.	Chocolate	T	TDS	W
Miguel et al., 2017	Compare the sensory properties of the inoculated and spontaneous fermentation processes of cocoa beans using TDS.	Chocolate	T	TDS	W

Moreira et al., 2017	Use a cocktail of microorganisms as a starter culture on the fermentation of the ripe cocoa pods from PH15 cocoa hybrid; evaluate its influence on the microbial communities present on the fermentative process, on both the volatile and non-volatile compounds produced during the fermentation; perform the chocolate sensorial characterization.	Chocolate	T	TDS	W
Oberrauter et al., 2018	Explore the kinetics of dominant sensations of polyphenol-associated sensory attributes bitter taste and astringency as well as sweet taste as covering agent in commercial dark origin and non-origin chocolates with different cocoa contents (%cc) over the course of evaluation.	Chocolate	T	TDS	W
Januszewska et al., 2020	Investigate the sensory characteristics of white and milk chocolate with three origins of vanilla (Madagascar, Indonesia, Papua New Guinea) using a multi-analytical approach.	Chocolate	C	TDS	W
Kiumarsi et al., 2021	Evaluate the changes in instrumental measurements and dynamic sensory properties of chocolate as affected by the partial and total substitution of sucrose by dodecanyl-succinylated inulin upon storage; compare the results of TDS of low-calorie chocolates with the sensory profile obtained by QDA to understand if these methods can complement one another in profiling chocolates.	Chocolate	T	TDS-I	W
S et al., 2021	Formulate low GI and low-calorie chocolates using natural sugar substitutes; evaluate various physicochemical, sensory, and glycemic studies were evaluated for the formulated products.	Chocolate	T	TDS	W
Martinez et al., 2021	Characterize the fermentation microbiologically and physicochemically.	Chocolate	T	TDS	W
Sethupathy et al., 2021	Provide insights into the development of novel products with modified texture, composition, and digestibility, particularly when the GI is the focus.	Chocolate	T	M-TDS	W
Khemacheevakul et al., 2021	Investigate how layering order of chocolates with different concentrations of sugar would influence the temporal sensory attribute profile, and if changes in the sensory profile influenced perceived sweetness intensity and overall acceptance of sugar-reduced and non-sugar-reduced 3D printed chocolates.	Chocolate	C	TDS	W
Bikos et al., 2022	Investigate the stages of oral processing of two varieties of micro-aerated chocolate and a solid variety, from the first bite until the thin film formation in the oral cavity.	Chocolate	E	TDS	W
Tanaya & Nagai, 2022	Propose a method to clarify the relationship between food texture and the geometric features of air bubbles.	Chocolate	T	TDS	W
Olegario et al., 2020	Evaluate the influence of formulations of different product categories on their dynamic sensory profile and consumer acceptability.	Chocolate; Coffee	T	TDS	W; B
Moss et al., 2021	Evaluate the differences between dry and sweet ciders.	Cider	T	TCATA-F	W
Evangelista, da Cruz Pedrozo Miguel, et al., 2014	Evaluate the effects of four yeasts as starter cultures during semi-dry coffee processing through microbiological analysis by denaturing gradient gel electrophoresis (DGGE), biochemical analysis by high performance liquid chromatography (HPLC) and headspace solid-phase microextraction/gas chromatography (HS-SPME/GC) and sensory analysis by TDS.	Coffee	E	TDS	W
Evangelista, Silva, et al., 2014	Evaluate the potential of yeast strains as starter cultures for dry fermentation of washed and non-washed coffee beans.	Coffee	E	TDS	W
Charles et al., 2015	Get better insight into the flavour perception processes by clarifying the link between the evolution of descriptors of flavour dominant at each moment during the consumption of a complex real product (coffee) and the release kinetics of aroma compounds.	Coffee	T	TDS	W
Evangelista et al., 2015	Evaluate the sensory profile of the final coffee beverage.	Coffee	E	TDS	W; B
Vietoris et al., 2015	Investigate the relationship between the sensory assessment of the varietal coffee samples and their aromatic profile measured using an electronic nose.	Coffee	C	TDS	W
Ribeiro et al., 2017	Use two sensory techniques (the cup taste and TDS analysis) to evaluate the effect of inoculating two coffee varieties [Ouro Amarelo (OA) and Mundo Novo (MN)] produced by the semi-dry method, with three distinct starter cultures (<i>Saccharomyces cerevisiae</i> CCMA 0543 and CCMA 0200 and <i>Torulaspora delbrueckii</i> CCMA 0684, respectively).	Coffee	T	TDS	W
Ribeiro et al., 2018	Identify the bacteria present during coffee fermentation of three varieties OA, MN, and Catuai Vermelho (CV); chemically characterize coffee beans by liquid and gas chromatography; evaluate final quality of the final beverage varieties by cupping taste.	Coffee	Not reported	TDS	W

Barbosa et al., 2019	Evaluate the dynamics of the dominant sensory perceptions for arabica coffee brews, comparing brews produced with two different coffees and with different roasting degrees.	Coffee	T	TDS	W
Martins et al., 2020	Evaluate the influence of different altitudes on the epiphytic microbiota of coffee beans and on sensorial and chemical quality of coffees grown at 800, 1000, 1200, and 1400 m.	Coffee	E	TDS	W
Vimercati et al., 2020	Develop and characterize the physicochemical, rheological, microbiological and sensory properties of coffee flavoured kefir; determine an optimal formulation.	Coffee	C	TDS	W
McCain-Keefer et al., 2020	Develop a lexicon for cold brew coffees; document the sensory properties of cold brew coffees, including temporal sensory properties; determine consumer preferences for cold brew coffee.	Coffee	T	TCATA	W
Galmarini et al., 2021	Evaluate the dynamic impact of two different sound fragments on coffee tasting.	Coffee	C	TDS	W
Pereira et al., 2022	Investigate the influence of fermentation under anaerobiosis and its microbial diversity, chemical and sensory profiles of coffees grown under different environmental conditions.	Coffee	E	TDS	W
Portela et al., 2022	Evaluate the dominant sensory characteristics of cold brews with different brewing conditions (coffee particle size and water temperature) and using different coffee species (<i>C. arabica</i> and <i>C. canephora</i>).	Coffee	T	TDS	W
van Eck, Hardeman, et al., 2019	Investigate the effect of toppings on oral processing behaviour and dynamic sensory perception of carrier foods when consumed as composite foods.	Combination of foods (bread + topping)	C	PP	W
van Eck, Fogliano, et al., 2019	Investigate how dynamic and static sensory perceptions of foods change when they are combined with condiments.	Combination of foods (bread or vegetable + sauce)	C	TDS	W
Gonzalez-Estanol et al., 2022	Investigate the effect of carrier addition (bread, wafer) on dynamic sensory perception of chocolate hazelnut spreads varying in composition using TDS and TCATA.	Combination of foods (carrier foods)	C	TDS ; TCATA	W
Paz et al., 2021	Analyze the interactions of pairing different types of chocolates with nuts, from professionals' and consumers' perspectives, to determine the best pairs depending on the dominant organoleptic characteristics of the different combinations; determine the different sensory perceptions in the pairing between chocolate and nuts from trained panelists; identify what makes a better match between chocolate and nuts according to consumers' perspective towards liking, the influence of balance in the pairs, and the predominant attributes; create pairing guidelines between chocolate and nuts from consumers' perspectives.	Combination of foods (chocolate and nuts)	T	TDS	W
Dinnella et al., 2012	Assess the impact of two Italian extra-virgin olive oils with different sensory properties on the perceived profile of pureed beans and tomatoes.	Combination of foods (olive oil + puree)	T	TDS	W
Sato & Kinugasa, 2019	Examine the changes in perception when tasting green tea and Japanese food alternately.	Combination of foods (tea + Washoku)	T	TDS	W
Galmarini et al., 2016	Present a different use of TDS data acquisition and analysis to evaluate the impact of one food product on another.	Combination of foods (wine + cheese)	C	TDS	W; B
Nygren et al., 2017	Explore consumer liking for wine and cheese pairs; explore whether either wine or cheese dominates a combination; understand the relationship between consumer liking and the dynamic taste experience.	Combination of foods (wine + cheese)	T	TDS	W; B
Galmarini et al., 2017	Evaluate a reverse protocol in which the effect of wine on cheese perception was studied.	Combination of foods (wine + cheese)	C	TDS	W; B
Galmarini et al., 2018	Identify which attributes impacted the dynamic liking of cheese and wine individually, as well as when consumed together over several sips and bites.	Combination of foods (wine + cheese)	C	TDS	W; B
Wilkin et al., 2021	Better understand consumer perception of taste and texture through TDS, effects of texture on the food product, and ultimately digestibility of seaweed species as a fortification product into simple food products (crackers).	Cracker	C	TDS	W
Saita et al., 2021	Determine the associations between onomatopoeic expressions and crispness/palatability of "kakinotane" rice crackers using the TDS and TDL methods.	Cracker	C	TDS	W
Sparvoli et al., 2021	Verify how some nutritional characteristics of the bean flours were influenced by different formulations and/or cooking conditions; assess the sensory properties and related acceptance level of bean-based products.	Cracker; Biscuit; Dessert/cream/mousse	T	TDS	W

Hutchings, Horner, et al., 2017	Investigate the effect of modifying the aftertaste of potato crisps on (1) temporal sensory perception and (2) appetite using three mouthwash conditions (no mouthwash, a water mouthwash, and a menthol mouthwash).	Crisp	T	TDS	W
Nguyen & Wismer, 2022	Compare temporal sensory profiles between commonly consumed salty food products available in the marketplace and their sodium-reduced counterparts, and compare the TDS and TCATA temporal sensory profiles of these products, over single and multiple intakes.	Crisp; Vegetable; Deli meat; Soup	T	TDS ; TCATA-F	W; B
Lorido et al., 2018	Evaluate the influence of NaCl replacement by KCl (15, 20 and 25%) on the sensory characteristics of dry-cured loins.	Deli meat	T	TDS	W
Rizo, Peña, et al., 2019	Characterize quality and oral processing characteristics of pork ham prepared with three culinary methods; establish links between oral processing characteristics, in-mouth sensory perceptions and textural properties depending on culinary method employed.	Deli meat	T	TDS	W
Lorido et al., 2019	Identify the emotions evoked during consumption of dry-cured hams, the connection of these emotions with dominant sensations during consumption and the possible influence of major ham components such as intramuscular fat.	Deli meat	T	TDS	W
Saldaña, Soletti, et al., 2019	Evaluate the sensory aspects of bacon smoked with reforestation woods and liquid smoke using the TDS methodology.	Deli meat	C	TDS	W
Saldaña, Martins, et al., 2019	Evaluate the impact of the type of smoking and fat content on the dynamic sensory perception of bacon by Brazilian consumers using TDS.	Deli meat	C	TDS	W
Auriema et al., 2022	Evaluate the effect of partial and total replacement of chicken skin (fat) by GBB on nutritional composition, technological properties, and dynamic sensory characterization of mortadella; determining the drivers of liking.	Deli meat	C	TCATA-F	W
Djekic, Ilic, et al., 2021	Characterize quality and oral processing characteristics of pork ham prepared with three culinary methods; establish links between oral processing characteristics, in-mouth sensory perceptions and textural properties depending on culinary method employed.	Deli meat	T	TDS	W
de Oliveira Paula et al., 2021	Evaluate the effects of different salt content and pork quality meats (PSE and RFN) on the sensory characteristics of restructured cooked hams using the acceptance and CATA descriptive tests, and its subsequent impact on the timely perception by the TDS test.	Deli meat	C	TDS	W
González-Mohino et al., 2021	Identify the drivers of liking and disliking as well as the attributes that define the ideal dry-cured loin using CATA questionnaires; identify the dominant sensory attributes during dry-cured loin consumption were identified by using the multiple intake TDS.	Deli meat	T	TDS	W; B
Morais et al., 2014	Analyze the TI profile and TDS in chocolate dairy dessert in the traditional version (with sucrose and traditional ingredients) and with prebiotic and light version (sweetened with different sweeteners and light ingredients).	Dessert/cream/mousse	T	TDS	W
Morell et al., 2015	Analyze the relevance of some rheological and texture-related cues to the expected satiation elicited by milk-based desserts prepared with different levels of HPMC.	Dessert/cream/mousse	T	TDS	W
Palczak et al., 2019	Sensory characterize some bi-layered chocolate desserts and find the sensory determinants of their perceived complexity.	Dessert/cream/mousse	T*	SP	W; B
Palczak et al., 2020	Investigate the influence of perceived complexity on consumers' preferences for some dairy desserts.	Dessert/cream/mousse	T	SP	B
Velázquez et al., 2020a	Apply cross-modal interactions (taste-odour-texture) for sugar reduction in products targeted at children, on both hedonic response and sensory perception and evaluating them by TDS and TCATA.	Dessert/cream/mousse	C*	TDS ; TCATA	W
Délérís et al., 2011	Investigate the impact of swallowing on aroma release and perception.	Distilled alcoholic beverage	T	TDS	W
Fiches et al., 2016	Better understand the origin of temporal perception of French grape brandies.	Distilled alcoholic beverage	T	TDS	W
Simioni et al., 2018	Perform a cross-cultural comparison of the dynamic profile of the attributes perceived in the cachaças, and the sensorial acceptance of the samples.	Distilled alcoholic beverage	T	TDS	W
Harwood et al., 2020	Evaluate the influence of ethanol concentration on temporal sensory perception of dark rums using TCATA.	Distilled alcoholic beverage	T	TCATA	W
He et al., 2021	Characterize the pungency during Baijiu consumption based on the dynamic temporal methods of TI and TDS, considering both pungency intensity and pungency sub-qualities.	Distilled alcoholic beverage	T	TDS	W

Silva et al., 2020	Evaluate the effects of the application of OH to milk (2, 4, 6, 8, or 10 V cm ⁻¹) on the sensory profile of DL evaluated using PAE and TCATA.	Dulce de leche	C	TCATA	W
Borgogno et al., 2017	Investigate the effect of replacing part of diet proteins with <i>H. illucens</i> on sensory properties of rainbow trout (<i>Oncorhynchus mykiss</i> Walbaum, 1792); describe and compare the physical-chemical characteristics and sensory properties of rainbow trout fed regular diet with FM as exclusive source of protein or fed diet including <i>H. illucens</i> in partial replacement of FM.	Fish	T	TDS	W
Dong et al., 2022	Evaluate the dynamic sensory perception of the fresh and roasted salmon and quantify the release of free amino acids (FAAs) and 5'-nucleotides in saliva during the chewing process.	Fish	T	TDS	W
Rocha et al., 2020	Evaluate the temporal profile of four different flavour enhancers: (i) the MSG; (ii) the IMP; (iii) the GMP; and (iv) the MAG.	Flavour enhancer	T	TDS	W
Ng et al., 2012	Discuss the benefits of using a sequential approach of QDA and TDS in measuring commercial blackcurrant squashes and explore the impact of the sample composition on taste and flavour perceptions in blackcurrant squashes.	Fruit juice	T	TDS-I	W
Zorn et al., 2014	Evaluate the influence of sucrose replacement by low-calorie sweeteners on the dynamic sensory profile of orange juice using multiple-sip TDS.	Fruit juice	T	TDS	W; B
Rodrigues et al., 2018	Use the TDS in conjunction with consumer acceptance to identify the drivers of liking of orange juice and to verify the influence of the preservation method in the sensory profile of this product as well as their acceptability.	Fruit juice	T	TDS	W
Wu et al., 2019	Shed light on the added value of TDS and TCATA to the dynamic profiles of simplified lemonade; examine the impact of citric acid level on the dynamic perception of citrus retro-nasal aroma in lemonade; investigate sensory interactions between the type of sweetener and citric acid levels in their effects on the dynamic perception of other tastes, lemon flavour (retro-nasal aroma) and tactile sensations.	Fruit juice	T	TDS ; TCATA	W
De Paula et al., 2022	Evaluate the interference of maceration time during 96 to 168 h over the antioxidant activity, volatile compounds profile and temporal dominance of sensations of fermented jabuticaba.	Fruit juice	T	TDS	W
Charles et al., 2017	Understand the influence of aroma perception on taste and texture perception in a real food matrix such as the apple, considering sensory perception both as a static and dynamic process; evaluate the suitability of TDS for studying sensory interaction.	Fruit	T	TDS	W
Gonçalves et al., 2017	Evaluate the appearance, texture, and flavour qualities of pasteurized strawberry pulp, using TDS and acceptance testing.	Fruit	T	TDS	W
Ávila et al., 2022	Evaluate bioactive compounds and sensory parameters of blood orange in order to overcome the obstacles and prejudices related to its consumption in Brazil and thus generate information that helps in trade and marketing strategies, exposing its functional benefits.	Fruit	Not reported	TDS	W
Gonçalves et al., 2020	Determine the sensory and physicochemical profiles of frozen mangaba pulp following long-term storage.	Fruit	C	TDS	W
Lima et al., 2020	Investigate the effect of indoor temperature on dominant sensations during intake of a beverage containing non-nutritive sweeteners.	Functional beverage	C	TDS	W
Saint-Eve et al., 2011	Investigate the role of texture on aroma and taste perception dynamics by using two sensory methods: a reference method sensory profile and a temporal method: TDS.	Gel	T	TDS-I	W
Frank et al., 2015	Examine pure agar gels and also the effect of fat addition on oral breakdown, volatile release, and perception.	Gel	T	PP	W
de Lavergne et al., 2016	Understand the oral processing and dynamic texture perception of composite foods by using emulsion filled gels with mechanical contrast.	Gel	T	TDS-I	W
Di Monaco et al., 2016	Characterize complex gels (sweetened, acidified, flavoured, and coloured gels) by using TDS in order to develop food models like jellies.	Gel	T	TDS	W
Patterson et al., 2021	Examine the sensory perception of textural complexity, as any notable impact on satiation is valid only if the designed levels are congruently perceived under human judgement.	Gel	C	TDS	W
Zhou et al., 2022	Investigate the effect of SFC in fat droplets on the creamy mouthfeel of acid milk gel.	Gel	T	TDS	W
Varela et al., 2014	Study the effect of different formulations, varying the proportion of stabilizing components (cream, egg yolk and a hydrocolloid mix) on the temporal sensory properties of the ice cream.	Ice cream	T	TDS-I	W

Kantono et al., 2016	Measure temporal changes in taste perceptions while chocolate gelati is consumed in the presence of music that varied in liking.	Ice cream	T	TDS-I	W
Kantono et al., 2018	Understand how different environmental settings (laboratory, immersive, and a natural eating environment) containing music varying in valence (disliked, liked, neutral) influences the temporal flavour perception of chocolate gelati.	Ice cream	C*	TCATA	W
Xu, Hamid, Shepherd, Kantono, & Spence, 2019	Investigate temporal changes in the perceived flavour of chocolate ice cream when consumed in a laboratory, café, university study area, and a city bus stop, and further examined how emotion and electrophysiological measures were influenced by these environments.	Ice cream	C*	TDS-I	W
Doyennette et al., 2019	Compare different methodologies to characterize oral behaviours applied during consumption of ice cream, and to understand how oral processing behaviour influences dynamic sensory perception of ice creams.	Ice cream	T	TDS	W
Xu, Hamid, Shepherd, Kantono, Reay, et al., 2019	Document perceptions of ice-cream consumed under different café soundscape conditions overlaid with either bird, machine, or forest soundscapes.	Ice cream	C	TDS	W
Kantono, Hamid, Shepherd, et al., 2019	Explore how emotions (measured both subjectively and electrophysiologically) covary with perceptions of chocolate gelato while listening to music differing in liking.	Ice cream	C	TDS	W
Lin et al., 2019	Understand the effects of environmental sounds (park, food court, fast food restaurant, cafe, and bar sounds) on the perception of chocolate gelato (specifically, sweet, bitter, milky, creamy, cocoa, roasted, and vanilla notes) using the TCATA method.	Ice cream	T	TCATA	W
de Medeiros et al., 2021	Characterize the temporal profile of low-calorie chocolate ice cream in traditional and lactose-free formulations, and to evaluate the preference of consumers.	Ice cream	T	TDS	W
Lin et al., 2022a	Provide an understanding of how a mixture of musical and non-musical sounds varying in pleasantness can influence temporal changes in food perception.	Ice cream	T	TDS-I	W
Rodrigues, Andrade, et al., 2016	Characterize the temporal profile of miracle fruit and assess its sugar substitute power in sour beverages.	Lemonade	T	TDS	W
Andrade et al., 2019	Assess the effect of different miracle fruit quantities on the temporal sensory profile and overall liking of lemonade and green tea, evaluating its effect on bitter beverages and determining the optimal quantity to be used as a sweetener in these products.	Lemonade; Tea/infusion	T	TDS	W
Shahbazi, Jäger, Chen, et al., 2021	Explore the effects of different biosurfactant types on the printing performance, textural properties, thermal behaviors, crystalline patterns, morphological features, and oral tribology of 3D printed constructs.	Meat analogue	T	TDS-I	W
Cordelle et al., 2022	Study the acceptance and sensory perceptions of plant protein meat substitutes, including the optimized experimental variant, compared to meat, in the context of a full meal to be as close as possible to real consumption habits; identify sensory drivers to improve the acceptance of plant protein meat substitutes.	Meat analogue	C	M-TDS	W; B
Shahbazi et al., 2022	Fabricate a reduced-fat 3D-printed meat analogue having a therapeutic property using the functionality of a modified bioactive MCC.	Meat analogue	Not reported	TDS-I	W
Ma et al., 2016	Evaluate the effects of chilled and frozen–thawed pre-treatments prior to PEF processing on the flavour and sensory characteristics of cooked lamb meats.	Meat	T	TDS-I	W
Watanabe, Motoyama, et al., 2019	Assess the sensory perception of Wagyu beef using TDS; check the integrity of the TDS by comparing the results with a physicochemical analysis.	Meat	T	TDS	W
Watanabe, Ohmori, et al., 2019	Assess the relative contribution of sensory characteristics on the dynamic sensory perception of pork loin by the TDS.	Meat	T	TDS	W
Kantono, Hamid, Oey, et al., 2019	Study the effects of PEF processing and storage (0 and 7 days) on the physicochemical and sensory characteristics of semitendinosus (ST) and biceps femoris (BF) beef muscles.	Meat	T	TDS-I	W
Pinto da Rosa et al., 2021	Check sensory differences in chicken cuts bred in three different systems of production (conventional, natural and organic); derive lexicons from emotional, functional or abstract concepts for the chicken meat category; investigate the consumer's perception of the products available in the poultry market.	Meat	C	TDS	W

Solomando et al., 2021	Evaluate the sensory characteristics, acceptability and purchase intention of cooked and dry-cured meat products as affected by the enrichment in EPA and DHA with fish oil microcapsules, by QDA, TDS and hedonic sensory techniques.	Meat	T	TDS	W
Djekic, Ilić, et al., 2021	Characterize oral processing characteristics of grilled pork meat after hot sauce coatings; evaluate the effects of pungency sensation during mastication of the samples on the perception of sensory stimuli as well as emotions; analyze pungency stimuli after swallowing.	Meat	T	TDPS	W
Kantono et al., 2021	Determine how chemical changes in terms of amino acids, fatty acids, and volatile composition with PEF processing and storage influenced temporal sensory perception of cooked meat during consumption.	Meat	T	TDS	W
Ilic et al., 2022b	Examine the impact of boiling, grilling, and sous-vide cooking methods on wild boar meat's textural, oral processing, and sensory qualities.	Meat	T	TDS	W
Djekic et al., 2022	Characterize sensory and oral processing characteristics of pork and poultry meat as food with different physical properties under different grilling temperatures grilled at three temperatures and identify parameters of interest for simulating mastication from the first bite; evaluate the mechanical properties of the samples and explore the potential of simulating the first bite using the finite element method.	Meat	T	TDS	W
Oliveira et al., 2015	Study the influence of sugar reduction on the dynamic sensory profile and consumers' liking of probiotic chocolate-flavoured milks; evaluate two alternatives (vanilla flavour and thaumatin) to attenuate the sensory changes caused by sugar reduction.	Milk beverage	T ; C	TCATA	W; B
Tomadoni et al., 2018	Study how enrichment with soluble fibers and protein affects the dynamic perception profiles of a strawberry-and-milk-based shake; study how the addition of a strawberry flavouring agent further affects these profiles; investigate the relationship between the changes in the formulation and the dynamic perception of sensations, consumer liking, and the expectation of satiety of the product.	Milk beverage	T	TDS	W
Zhu et al., 2020	Study the effect of milk protein compositional changes and hydrocolloids addition on the dynamic oral perception of the textural and mouthfeel attributes of the chocolate milk from a trained panel; investigate the role of saliva in the texture aspects of chocolate flavoured milk.	Milk beverage	T	TDS-I	W
Lin et al., 2022b	Understand how the sounds influenced the temporal sensory perception of flavours of different food types.	Milk beverage; Ice cream	T	TDS	W
Pionnier et al., 2004	Assess flavour perception of the model cheese using this modified TI procedure and investigate the potential relationships between inter-individual variations in sensory assessments, in oral parameters and in flavor release.	Model food	T	A-TI	W
De Loubens et al., 2011	Quantify the breakdown of model dairy products in terms of contact areas independently of the individual variability of consumers.	Model food	Not reported	TDS-I	W
Tang et al., 2017	Develop model foods of varied and controlled structural and textural complexity, in order to determine quantitative instrumental measures of “complexity” related to sensory assessment.	Model food	T	TDS	W
Toko, 2022	Visualize the changes in the taste of food during the chewing process using as a model food with a three-layer structure monaka (a traditional Japanese sweet) composed of three ingredients: a salty gel, mashed potato, and the wafer.	Model food	T	TDS	W
Feltrin et al., 2015	Determine the amounts of different sodium chloride replacements, including potassium chloride, monosodium glutamate, potassium lactate, calcium lactate and potassium phosphate, required to promote an equivalent degree of ideal saltiness in aqueous solution; study the sensory profile of sodium chloride and its substitutes using TDS analysis.	Model solution	T	TDS	W
Kawasaki et al., 2016	Investigate whether TDS curves of basic taste solutions reflect the concentrations of three basic tastants (NaCl, lactic acid, and MSG) and also assessed whether TDS curves of binary tastant mixtures reflect the coexistence of these basic tastes.	Model solution	T	TDS	W
Appelqvist et al., 2016	Determine the impact of a spoon to spoon multiple ingestion on selected flavour and texture attributes for a range of emulsions containing different levels of fat.	Model solution	T	SP	B
Medel-Marabolí et al., 2017	Evaluate the effect of different concentrations of commercial tannin on the temporality of the perception of astringency.	Model solution	T	TDS	W
Stolle et al., 2017	Classify subjects according to their salt taste sensitivity, quantitatively measure their saliva flow induced by the salty stimuli of sodium chloride, potassium chloride, ammonium chloride, and sodium gluconate; to investigate time-	Model solution	T	TDS	W

	dependent dynamic changes in the salivary proteome by means of tryptic in-solution digestion of whole saliva samples prior to and after chemosensory stimulation, followed by protein quantitation using isobaric tags for relative and absolute quantitation (iTRAQ) and nano-liquid chromatography and tandem mass spectrometry (nano-LC-MS/MS).				
Zhang et al., 2017	Investigate the DTs, perceptual characteristics, and temporal migration of hydroxyl-sanshool compounds dissolved in ethanol-water, saccharose, NaCl, and MSG using 2-AFC, TI and TDS methods.	Model solution	T	TDS	W
Yang et al., 2021	Explore the impact of capsaicin on aroma release and perception simultaneously.	Model solution	T	SP	W
Pu et al., 2022	Evaluate the saltiness enhancement of umami compounds with different replacement ratios in NaCl solution by sensory evaluation methods including 2-AFC and TDS; optimize the optimal sodium-reduce ratio and the best saltiness enhancement formula by the 2-AFC test; 3 explore the saltiness enhancement mechanisms by multiple instrumental analytical methods, including characterization of the rheology and particle size of different salt solutions, and the effect of umami compounds and food gel on the binding properties between Na ⁺ and mucin.	Model solution	T	TDS	W
Rosenthal & Share, 2014	Examine the breakdown path of peanuts and peanut products, and to put them in the context of Hutchings and Lillford (1988) model to illustrate the oral breakdown path.	Nuts	T	TDS	W
Rodrigues et al., 2019	Survey the volatile composition, chemical and sensory characteristics of VOO from eight olive cultivars grown in the southeast region of Brazil, providing information for comparisons with VOOs from other regions and help in future studies on a controlled designation of origin or geographical indication.	Olive oil	Not reported	TDS	W
Cliceri et al., 2021	Investigate the variability in perception of bitterness, pungency and astringency in EVOO over time and whether this might result in differences in flavor perception.	Olive oil	T	TCATA	W
Laleg et al., 2017	Study the impact of the percentage (0%–100%) of enrichment using legume protein (faba) in pasta and the impact of drying temperature (low temperature, LT, vs. very high temperature, VHT) on pasta structure and its resulting textural and cooking properties.	Pasta	C	TDS	W; B
Marcano et al., 2015	Discover how each individual minor ingredient of the cheese pie affects: a) dynamic texture perception, b) instrumental texture parameters, and c) dynamic liking.	Pies	T	TDS	W
di Monaco et al., 2016	Evaluate the effect of storage condition, cooking method and serving temperature on dominant attributes over time for polenta sticks; evaluate dynamic liking by consumers by using an innovative approach; find correlations between TDS data and consumer liking scores.	Polenta	T	TDS	W
Makame et al., 2019	Characterize the texture of selected indigenous Complementary Porridge typically used for feeding infants and young children aged 6–24 months in African communities during Oral Processing (therefore dynamic), in order to make recommendations for optimizing their oral texture to improve nutrient intake.	Porridge	T	TCATA	W
Methven et al., 2010	Investigate the effect of consumption volume on the sensory profile and liking of ONS; investigate if modifications of sweetness and metallic levels could improve the hedonic liking of ONS.	Protein beverage/ONS	T	SP	W; B
Withers et al., 2014	Test the following hypotheses: (1) that casein within protein-fortified beverages of neutral pH could contribute to perceived mouth drying, and that this may result from the presence of γ -CN; (2) that whey proteins within protein-fortified beverages of neutral pH could contribute to perceived mouth drying; (3) that increased sweetness could suppress the perception of mouth drying from protein-fortified beverages; (4) that increasing viscosity could reduce perceived mouth drying from protein-fortified beverages, perhaps by reducing the interaction between proteins and the oral epithelium; and (5) that increasing fat could reduce perceived mouth drying from protein-fortified beverages, perhaps by increasing lubrication.	Protein beverage/ONS	T	SP	W; B
Thomas et al., 2016	Expand the TDL concept by recording TDS and dynamic liking data in the same session and during the consumption of the full portion of a product.	Protein beverage/ONS	C	TDS	W; B
Bull et al., 2017	Explore the relationship between denaturation of whey proteins and sensory attributes related to mouth drying.	Protein beverage/ONS	T	SP	W; B
Thomas et al., 2018	Extend the A-TDL protocol with a TDS and dynamic liking pairing at the lab twice a day (morning and afternoon sessions) with older adults, research could reproduce the full daily experience of ONS consumption.	Protein beverage/ONS	C	TDS	W; B
Wagoner et al., 2018	Evaluate the effects of texture on sweetness dose-response profiles of whey protein solutions sweetened with sucrose, sucralose, stevia, or monk fruit extracts, and to assess temporal progression of dominant attributes.	Protein beverage/ONS	T	TDS	W

Parker et al., 2018	Profile the sensory properties of ready-to-mix (RTM) whey protein beverages sweetened with the natural sweeteners stevia, monk fruit, and crystalline fructose and blends of these sweeteners using 3 temporal methods: TI, TDS, and TCATA.	Protein beverage/ONS	T	TDS ; TCATA	W
den Boer et al., 2019	Investigate how altered levels of sweetness intensity (low versus high) and thickness (thin versus thick) could affect intake of ONS, as a strategy to improve ONS intake; investigate the descriptive sequential profile of each ONS to determine how this may impact intake.	Protein beverage/ONS	T	SP	W; B
Wagoner et al., 2020	Evaluate how different systems alter temporal perception during consumption.	Protein beverage/ONS	T	TDS	W
Cosson et al., 2020	Understand better how product composition affected the sensory perception of pea protein-based beverages using three different sensory profiling methods.	Protein beverage/ONS	T	TDS-I	W; B
Delompré et al., 2020	Study the dynamics of perception of the taste and retro-nasal odour of ODTs to better control and limit their off flavours.	Protein beverage/ONS	T	TDS ; TCATA	W
Lester et al., 2021	Investigate differences in the temporal consumption experience (comprising sensory perception, in-mouth aroma release and subjective appetite) of a clinically relevant portion of ONS, for groups differing in SFR.	Protein beverage/ONS	T	SP	B
Harwood & Drake, 2021	Evaluate the application of penalty analysis to temporal evaluation of vanilla-flavoured protein beverages sweetened with different sweeteners.	Protein beverage/ONS	C	TCATA	W
Agudelo et al., 2015	Assess three hydrocolloid systems (tapioca starch, modified waxy corn starch and a tapioca starch-pectin blend) in reformulated fruit fillings, using two relatively novel sensory techniques: TDS and CATA; link these sensory results with product liking.	Puree	T	M-TDS	W
Sharma & Duizer, 2019	Investigate the effect of starch, xanthan, and their blends on dynamic texture perception in a hydrocolloid thickened pureed food; determine if TDS and TCATA provide similar or complementary information in terms of texture evolution and product differentiation considering the aspect of dynamics of sensory perception.	Puree	T	TDS ; TCATA	W
Rodrigues et al., 2014	Prepare a seasoning and garlic salt with a low sodium content using various mixtures of salts, including NaCl, KCl and monosodium glutamate; perform sensory tests and evaluations of temporal acceptance (TI and TDS).	Rice	T	TDS	W
Nascimento et al., 2020	Evaluate the effect of induced germination and various stresses on the technological, nutritional, and sensory parameters of brown rice.	Rice	T	TDS	W
Anandharamakrishnan et al., 2021	Assess the overall picture of oral processing of three different rice varieties considering mastication behaviour, bolus properties, oro-sensory perception, and starch hydrolysis.	Rice	T	M-TDS	W
Filho et al., 2020	Determine the dynamic sensory profile of tomato sauce formulated with flavor enhancers, a bitter blocker and 50% sodium content reduction.	Sauce	C	TDS	W
Paulsen et al., 2014	Determine the effects of different salt substitutes on the overall sensory perception of sausages.	Sausage	T	TDS	W
Devezeaux de Lavergne et al., 2015	Determine the influence of naturally preferred eating duration on dynamic texture perception of sausages; explain differences in dynamic texture perception between subjects having a different eating duration by their chewing behaviour and bolus properties.	Sausage	T	TDS	W
Braghieri et al., 2016	Assess the effect of preservative addition on sensory properties of Lucanian dry sausage using both the TDS method and the conventional QDA.	Sausage	T	TDS-I	W
Paglarini et al., 2020	Evaluate the influence of salt and fat reduction on Bologna sausages with inulin-based emulsion gels in the dynamic sensorial perception by using TDS and TCATA methods considering overall liking.	Sausage	C	TDS ; TCATA	W
Câmara et al., 2020	Assess the potential use of chia mucilage (in the form of a powder and gel) as a substitute for phosphates in low-fat Bologna sausages through the evaluation of its functional properties, such as water-holding capacity, gel and emulsion formation, texture, and physicochemical and sensory properties.	Sausage	T	TDS	W
Aguayo-Mendoza et al., 2020	Determine the effect of age, gender and ethnicity on consumption time, bolus properties and dynamic sensory perception of sausages.	Sausage	C	TDS	W
Simunovic et al., 2022	Evaluate the impact of nitrite reduction on geometrical, colour, chemical, microbiological, and sensory traits of dry sausage traditionally produced with red paprika powder.	Sausage	T	TDS	W

Furey et al., 2022	Develop nutritious convenient spreadable seafood products using roe in combination with milt and fish liver; determine the products eating quality and acceptance among teenagers; survey teenager's attitudes towards seafood consumption.	Seafood pâté	C*	TDS	W
Ramos-Diaz et al., 2019	Identify the spatial location of major food components in extrudates containing amaranth, quinoa and kañiwa through the application of NIR hyperspectra imaging; and study the effect of such food components on specific sensory attributes through the application of chemometrics.	Seeds	T	TDS	W
Menis-Henrique et al., 2019	Provide information about dominant sensory sensations, which may be related to lipid content.	Snacks	T	TDS	W; B
Panzarini et al., 2020	Assess the effects of salt and monosodium glutamate on the sensory characteristics of expanded snacks of cheese-flavoured corn grits, aiming to reduce the sodium content of these products.	Snack	T	TDS	W
Emorine et al., 2021	Follow the progressive release of salty and odorant stimuli and the dynamics of perception during the consumption of flans under the action of mastication and saliva.	Snack	T	TDS	W
Tang et al., 2020	Examine whether savoury enhancement in a beef broth has an impact on its perceived calories, expected fullness and prospective portion selection.	Soup	T	TCATA-F	W
Liang et al., 2022	Explore the novel umami or umami-enhancing peptides derived from chicken soup and understand the umami mechanism of the peptides.	Soup	T	TDS	W
Zhang et al., 2018	Investigate the pungent intensity, perceptual characteristics, and time-related aspects of <i>Z. bungeanum</i> .	Spice	T	TDS	W
Ramírez-Rivera et al., 2021	Develop the sample testing conditions and the vocabulary for sensometric characterisation of habanero chilli; to validate the sample testing conditions comparing the habanero chilli against other chillis consumed in Mexico using QDA and TDS.	Spice	T	TDS	W
Hawthornthwaite et al., 2015	Quantify the sticky sensations implicated in peanut butter as the dominant sensation prior to deglutition.	Spreadable	C	TDS	W
Di Monaco et al., 2014	Compare the sensory performance of a sweet protein, MNEI, with aspartame, saccharin and sucrose.	Sweetener	T	TDS	W
Withers et al., 2016	Modify sequential profiling to increase the total ingestion volume and include a wider range of attributes in a case study of sugars and sweeteners.	Sweetener	T	SP	B
Reyes et al., 2017	Characterize and compare the dynamic properties of non-nutritive sweeteners using TCATA.	Sweetener	C	TCATA	W
Tan et al., 2019	Profile and compare the temporal sweetness and qualitative differences of 15 sweeteners.	Sweetener	T	TCATA-F	W
Junge et al., 2022	Investigate how citric acid alters the temporal sweet taste profile of sucrose and d-allulose and MFE; investigate the Sweetening System-Acidity interaction in three-time intervals reflecting the attack-evolution-finish of perception; examine perceptual differences between Sweetening Systems, and Sweetening Systems-Acidity combinations, relative to sucrose, over time.	Sweetener	C	TCATA-F	W
Kobue-Lekalake et al., 2012	Determine the effect of sorghums with varying tannin content on bitterness and astringency simultaneously using DATI.	Tea/infusion	T	DATI	W
Liu et al., 2021	Characterize the sensory properties of cold-brewed infusions of Taiwanese teas.	Tea/infusion	C	TCATA	W
Rodrigues, de Souza, et al., 2016	Evaluate the effect of reducing the particle size in a low-sodium salt mixture (sodium chloride, potassium chloride, and monosodium glutamate) on the salting potency, TI profile, and TDS in shoestring potatoes, with the aim of obtaining a low sodium content product with sensory quality.	Vegetable	T	TDS	W
Dupas de Matos et al., 2019	Evaluate the sensorial effects of its use as the substitute of vinegar in vegetable preserves by using pickled cucumbers as a model.	Vegetable	T	TDS	W
Ilic et al., 2021a	Inspect how boiling, steaming, and sous-vide impact the physical properties, oral processing, and dynamic sensory perception of celeriac, and what are the relations among them.	Vegetable	T	TDS	W
Ilic et al., 2021b	Explore the influence of boiling, steaming, and grilling on potato mechanics, mastication, and sensory perception.	Vegetable	T	TDS	W
Ilic et al., 2022a	Uncover the influence of cooking on the eggplant and zucchini appearance, mechanics, mastication, and sensory perception.	Vegetable	T	TDS	W

Ilic et al., 2022c	Uncover the influence of two cooking methods, steaming and sous-vide, on the eggplant and zucchini colour, mechanics, mastication, and sensory perception.	Vegetable	T	TDS	W
Pacheco et al., 2018	Determine the static and dynamic sensory profile of different bottled mineral water samples; evaluate the consumer's liking of the products.	Water	T	TDS	W
Meillon et al., 2009	Describe the impact of partial dealcoholisation by reverse osmosis on the sensory perception of Merlot and Syrah wines.; investigate the use of TDS method to describe subtle differences in wines.	Wine	T	TDS-I	W
Meillon et al., 2010	Investigate the effect of partial alcohol reduction on the perceived complexity, temporality of sensations and liking of a Syrah wine.	Wine	T	TDS-I	W
Sokolowsky & Fischer, 2012	Compare the evaluation of bitter taste in commercial wines, which strongly varied in their bitterness, with DA, TI and TDS.	Wine	T	TDS	W
Sokolowsky et al., 2013	Evaluate the sensory impact of commonly used skin contact on two cool climate varieties namely, Riesling and Gewürztraminer, by applying DA, TI and TDS analyses to the same set of wines; correlate wine composition with orally perceived intensities recorded by DA and parameters extracted from TI and TDS curves; investigate which complementary knowledge could be gained by each applied sensory technique.	Wine	T	TDS	W
Rébénaque et al., 2015	Gain insight into the sensory properties of subpopulations of proanthocyanidic tannins of different molecular sizes; determine whether a correlation exists between the sensory and the nephelometric data.	Wine	T	TDS-I	W
Etaio et al., 2016	Describe the evolution of the sensations perceived in red wines from Rioja Alavesa by using TDS; compare wines made with the two winemaking procedures used in Rioja Alavesa: carbonic maceration and destemming.	Wine	T	TDS-I	W
del Barrio-Galán et al., 2016	Study the effect of the ageing on lees, with or without oak wood chips, with a commercial inactive dry yeast and previously fermented with different <i>Saccharomyces cerevisiae</i> yeast strains, which exhibit different capabilities for polysaccharide release, on the quality of Chardonnay Chilean white wines.	Wine	T	TDS	W
Lytra et al., 2016	Evaluate the impact of ester and DMS concentrations evolution on variations in wine fruity aroma during sensory analysis.	Wine	E	SP	B
Vidal et al., 2016	Obtain a dynamic description of the astringency of Tannat red wines using TDS.	Wine	T	TDS	W
Baker et al., 2016	Examine the application of a novel temporal method and its ability to describe the nuances of wine finish, beyond just total duration, as a function of ethanol concentration.	Wine	T	TCATA	W
McMahon et al., 2017	Describe the sensory aspects associated with carbonation in sparkling wines using both static (DA) and dynamic (TCATA) sensory evaluation methods, with a comparison offered between these two methods.	Wine	T	TCATA	W
Frost et al., 2017	Evaluate the effect of tannin, acid and ethanol concentration on the temporal perception of taste and mouthfeel.	Wine	T	TDS	W
Vazallo-Valleumbrocio et al., 2017	Chemically characterize eleven commercial oenological tannins that are marketed for oenological use and to evaluate their effects on the chemical and sensory characteristics of a Carménère wine during 90 days of bottle aging.	Wine	T	TDS-I	W
Frost et al., 2018	Evaluate the effect of cap management and extended maceration technique on the chemical and sensory profile of Merlot wine.	Wine	T	TDS	W
Schumaker et al., 2019	Evaluate the influence of oak and green characteristics on the temporality of red wine finish, specifically in-mouth evaluations of wines containing 4-EP, 4-EG, and IA.	Wine	C	TCATA	W
Kang et al., 2019	Characterize the dynamic astringency profiles of 13 Australian commercial red wines and 2 rosés made from 11 grape varieties using modified progressive profiling.	Wine	T	PP	W
Criado et al., 2019	Determine the relationship between individual differences in saliva composition and the perception of different aroma stimuli elicited by some typical wine carboxylic esters associated to the pleasant and fruity character of many wine types.	Wine	T	PP	W
Kemp et al., 2019	Ascertain whether TCATA, when focused on specific modalities, could distinguish red wines made from the same grape variety, according to mouthfeel and texture descriptors only.	Wine	T	TCATA	W
Wang et al., 2019	Assess how the same red wine might be experienced differently when tasters were exposed to various sonic environments (two pieces of music plus a silent control condition).	Wine	C	TDS	W

Poveromo & Hopfer, 2019	Assess the dynamic changes in sensory perception as a function of varying wine matrix components, such as ethanol, glycerol, and phenolics.	Wine	T	TCATA	W
Olatujoye et al., 2020	Investigate the effect of β -lactoglobulin, in comparison with gelatin, on the perception of in-mouth attributes, particularly astringency, in red wine; ascertain if there is agreement between the chemical method applied in the assessment of astringency and the sensory evaluation of astringency.	Wine	T	SP	W; B
Kang et al., 2020	Investigate, in Shiraz wine production, the impact of the ACE technique on wine chemical composition, sensory attributes and in particular, astringency and its sub-qualities.	Wine	T	PP	W
Lyu et al., 2022	Determine the influence of tannins on intraoral release by intraoral SPME after wine spitting at different time points.	Wine	T	TDS ; PP	W
Wang et al., 2021	Understand how temporal elements of the tasting experience is related to the assessment of complexity in both novices and experts.	Wine	E ; C	TCATA-F	W
Medel-Marabolí et al., 2021	Contrast data related to sensory evaluation of six wines produced from different <i>Vitis vinifera</i> grape varieties carried out using three methods (DA, TDS and TCATA) against the phenolic composition of the wines to examine associations between secondary metabolites and sensory perception over the course of tasting.	Wine	T	TDS ; TCATA	W
Fryer et al., 2021	Determine the optimal interstimulus rinse protocol and time separation of samples during smoke attribute evaluation to mitigate any potential carryover effects.	Wine	T	TCATA ; PP	W
Pittari et al., 2022	Shed light on the impact of the addition of oenological tannins on wine perception and on in vivo aroma release before and after oxidation.	Wine	T	TDS-D	W; B
Correia et al., 2022	Analyze the sensory characteristics of red wines produced in the three Douro sub-regions by QDA® and also applying a temporal method-TDS, intending to verify if there are specific sensory profiles related to each of the sub-regions.	Wine	E	TDS	W
Bouteille et al., 2013	Confirm the link between the freshness sensation and the overall appreciation of the yoghurts and yoghurt-like products; to derive the relevant sensory attributes to characterize the complex perception of freshness for these products; to estimate the strength of the links between these sensory attributes and the freshness sensation; to derive what temporal profile of the sensory attributes would better generate the freshness sensation.	Yogurt	T	TDS	W
Bruzzone et al., 2013	Apply the TDS methodology to evaluate the dynamics of yoghurt textural perception; study the influence of fat, modified starch and gelatine concentrations on the temporal aspects of yoghurt textural perception; compare results from TDS with a traditional methodology, such as QDA.	Yogurt	T	TDS	W
Mesurole et al., 2013	Study the impact of some characteristics of fruit pieces on aroma release and perception in the case of flavored yogurts.	Yogurt	T	TDS	W
Lesme et al., 2020	Combine the M-TDS methodology with the multi-intake approach to investigate the changes of perception due to the use of innovative functional whey protein aggregates in yogurts.	Yogurt	T	M-TDS	W; B
Greis et al., 2020	Determine how mouthfeel perception alters during mastication and to understand if there is a difference in temporal texture profile between plant-based and dairy yogurts by using TDS in a consumer test; identify the temporal drivers of mouthfeel liking and disliking in the plant-based semi-solid yogurt alternatives and specially to understand to what extent the bulk properties are responsible for liking.	Yogurt	C	TDS	W
Ribeiro et al., 2020	Optimize a mix of different stevias for a high protein plain yoghurt using the mixture design model with the Napping® and response surface methodology.	Yogurt	C	TDS	W
Varela et al., 2021	Explore individual differences in drivers of liking and satiation through the combined analysis of eating behaviour, liking, eating rate, food intake and both static and dynamic sensory perception.	Yogurt	T	TCATA	W
Oliveira et al., 2021	Characterize the temporal profile, evaluating sweetness perception overtime, by TI, and the evolution of the dominant sensory perceptions, by TDS, of low-sugar yogurt samples added with strawberry and vanilla flavours in different concentrations, and evaluate the acceptance of these products.	Yogurt	T	TDS	W
Pereira et al., 2021	Evaluate the dynamic sensory profile of skyr yoghurts with the addition of mango, prebiotics, and natural sweeteners.	Yogurt	T ; C	TDS	W
Olegario et al., 2022	Explore the impact of lactose free version on dynamic sensory perception using the multiple-intake TDS approach.	Yogurt	T	TDS	W; B
Chadha et al., 2022	Investigate how sugar reduction in yogurt can influence temporal sensory perception and liking; investigate the effects of sweetener type on healthiness, satisfaction, and purchase intent after the consumption of yogurt.	Yogurt	T	TDS	W; B
Greis et al., 2022	Determine the influence of aroma compounds possibly congruent with sweetness on perceived sensory profile, particularly on the mouthfeel properties.	Yogurt	T	TDS	W

3.2 Development and use of temporal sensory evaluation methods

3.2.1 History of development of temporal sensory evaluation methods

The first descriptive method that considered the temporal aspect was TI. The reference research paper on the method is from 1986, however, this method was first implemented in 1945 (Sjöström, 1954), and was more efficiently used later thanks to the development of computerized systems.

After TI, many methods were developed aiming to describe products in a temporal fashion using more than one descriptor. Table 9 shows a total of 23 methods which were developed and published after TI up to December 2022.

It took many years to come up with a method which could allow assessors to perform a wider description using more than one attribute. Following the TI paradigm, the older multi-attribute methods (PP, A-TI, DATI, SP, and MATI, see Table 9 for acronyms) were intensity-based. However, these methods were not widely used: only a total of 38 published articles were found, representing 10% of the articles of the present scoping review. This raises concern about the methods' difficulty of use for panellists (Meiselman et al., 2022). Moreover, it could be possible that, at the moment when these methods were introduced, researchers expected to get information on many attributes (influenced by the static Descriptive Analysis method, which has a wide list of attributes) or with a high level of detail as in TI and this could have impacted on the way the experiments were designed and analysed.

It was only in 2008 (publication date of first research paper though the method was presented before at the 5th Pangborn Symposium as Pineau et al., 2003) that a change of paradigm was introduced with TDS (Table 9). This method proposed a description based on the tracking of dominant attributes stating one dominant attribute at a time. To date, the most used definition of a dominant attribute is “the attribute that triggers the attention (i.e. most striking perception at a given time), not necessarily the most intense” (Pineau et al., 2009), but several definitions can be found in the literature (Hutchings et al., 2022; Varela et al., 2018). Originally, subjects were also asked to rate the intensity of the dominant attributes (TDS-I) as it was done with all the methods before, but the rating task was deemed too difficult (Schlich, 2017) and the intensity rating was no longer recommended, and TDS-I became TDS.

Table 9. Descriptive temporal sensory evaluation methods, acronyms used, date and reference of first publication in a research paper and number of articles in which the method has been used until December 2022. (several methods can be referenced in a same article).

Method name and acronym	Reference paper	Number of articles
TI: (Single-attribute) Time-Intensity	Lee & Pangborn, 1986	414 ¹
Multi-attributes intensity-based methods		38
PP ² : Progressive Profile	Jack et al., 1994	18
DATI: Dual-Attribute Time-Intensity	Duizer et al., 1996	4
A-TI: Alternated Time-Intensity	Pionnier et al., 2004	1
SP: Sequential Profile	Methven et al., 2010	12
MATI: Multi-Attribute Time-Intensity	Kuesten et al., 2013	2
TDS-I: Temporal Dominance of Sensations with intensity	Le Révérend et al., 2008	35
TDS and variants		255
TDS ³ : Temporal Dominant of Sensations	Pineau et al., 2009	241
M-TDS: TDS by modality	Agudelo et al., 2015	13
TDS-HD: TDS hold down	van Bommel et al., 2019	1
TDS-D: Dual TDS	Pittari et al., 2022	1
TCATA and variants		70
TCATA: Temporal Check All That Apply	Castura et al., 2015	54
TCATA-F: TCATA fading	Ares et al., 2016	15
D-TCATA: Discrete time TCATA	Visalli et al., 2022	2
M-TCATA: TCATA by modality	Dietz et al., 2022	1
M-TCATA-F: TCATA fading by modality	Barker et al., 2022	1
Other qualitative methods		7
TQT: Time-Quality Tracking	Zwillinger & Halpern, 1991	1
AEF-D: Attack-Evolution-Finish with Dominance	Visalli et al., 2020	1
FC-AEF-A: Attack-Evolution-Finish with Applicability with free comment	Mahieu et al., 2020	1
AEF-A: Attack-Evolution-Finish with Applicability	Visalli et al., 2022	1
Quessence	Jeltema et al., 2020	1
F-TOS: Free Temporal Order of Sensations	Carrillo et al., 2021	1
TR: Temporal ranking	Keefer et al., 2022	1

¹This number includes 388 articles referencing only TI (out of scope of this review), plus 26 referencing TI and other temporal methods (in the scope of this review); ²Progressive Profile was found also as “Dynamic profile”, “Fixed-Time Profile Intensity”, “Discrete Time-Intensity”, but in all the cases the principle was the same. ³TDS was found as “Temporal Dominance of pungency Sensations” in one publication.

In 2015, TCATA was developed (Table 9) looking to overcome the limitation inherent to the concept of dominance that imposes subjects to choose and report only one attribute at each time. TCATA enables the tracking of all perceived (applicable) sensations. In this way, the path of temporal methods continued to evolve in the direction of qualitative evaluations. TCATA also assumes that when a sensation is no longer perceived, assessors will uncheck the corresponding attribute. This last assumption was difficult to prove (Ares et al., 2016) given

that the cognitive task of “unchecking” (stating absence of sensation) is quite different from “checking” (being vigilant of the new perceived sensations). Trying to cope with this, the fading alternative was introduced (TCATA-F, Table 9) considering that after a certain time the sensation will no longer be applicable and thus the attribute is automatically unchecked, freeing the panellist from this task. The time elapsed from check to automatic uncheck (fading time) is determined by the experimenter and not by the evaluator.

New variants of TDS and TCATA were developed afterwards (Table 9). These include TDS evaluations with the possibility of indicating two attributes at a time (instead of only one) given that they corresponded to different sensory modalities (Dual TDS). Another variant considers the possibility of having periods of “non-dominance” (TDS-HD). Discrete time variants of TCATA have been proposed, still trying to avoid the limitation that comes from (not) unchecking attributes (D-TCATA). To limit the number of attributes simultaneously tracked and avoid favouring the sensory modalities easier to identify (e.g. texture vs. flavours), both TDS and TCATA proposed variants consisting in evaluating successively the different sensory modalities (M-TDS, M-TCATA).

Recently, new qualitative methods (Table 9) summarizing the perception in several periods were proposed (AEF-D, AEF-A, FC-AEF-A, Quessence, F-TOS). The rationale behind these retrospective evaluation methods (by opposition to concurrent evaluations in continuous or discrete time) is to simplify data collection and analysis by sacrificing temporal resolution. Finally, TR (Table 9) is in-between qualitative and quantitative measurements, as the method asks subjects to rank the perceived attributes according to their intensity at each time (TR).

These first mentioned results show there may be a delay between the first presentations of methods (often at congresses) and their validation by peers (publication). Some temporal methods have never been published in a peer-review journal (e.g. Temporal Order of Sensations, Pecore et al., 2011; Pick-3-And-Rank, Vandeputte et al., 2011; etc.) and thus probably never reused by other people. It can also be noticed that some identical methods were named in different ways, probably because the authors “reinvented” them without knowing it. Finally, ongoing developments of temporal methods suggest that there are needs that are still not being met and that there is room for improvement in the field.

3.2.2 Main characteristics and differences of temporal sensory evaluation methods

Based on the previous description on the evolution of multi-attribute temporal methods, it is clear that they are not all based on the same paradigm, nor do they integrate time in the same fashion. In fact, temporal methods can be grouped based on two main characteristics: the type of measurement which can be associated with qualitative (dominance and applicability) or quantitative (intensity) concepts, and the moment of the evaluation. Figure 7 presents the different sensory methods resulting from the combination of these characteristics.

		Moment of evaluation				Attribute indicated using	
		Concurrent		Retrospective			
		Continuous time	Fixed time	By period	Global		
Type of measurement and concept	Qualitative	Dominance	TDS, M-TDS, Dual-TDS, TDS-HD, TDPS		AEF-D	Button	
		Applicability			FC-AEF-A, F-TOS, Quessence	FC ¹	Free comment
			TQT, TCATA, TCATA-F, M-TCATA, M-TCATA-F	D-TCATA	AEF-A	CATA ²	Checkbox
	Quantitative	Dominance + Intensity	TDS-I				Discrete/Continuous scale
		Applicability + Intensity				RATA ³	
		Applicability + Ranking	TR				
		Intensity	TI, DATI	SP, PP, MATI, A-TI		DA ⁴	

High ← Temporal resolution → Low

Figure 7. Main characteristics of descriptive sensory evaluation methods. Gray text: static methods: ¹Free Comment; ²Check-All-That-Apply; ³Rate-All-That-Apply, ⁴Descriptive analysis. See Table 9 for other acronyms.

When products are evaluated globally in a retrospective manner with no particular consideration of the temporality of perception, the descriptive method is considered static (e.g. Free Comment, CATA, Rate All That Applies, Descriptive Analysis – all of them out of scope of the present article). On the other hand, an evaluation can be retrospective but considering and recalling the sensations perceived during specific temporal periods of the evaluation. For example, in AEF (and its variants, see Table 9) the periods are attack, evolution and finish (giving origin of its name) while in F-TOS the three first sensations are considered.

The so-called dynamic methods ask subjects to report their perception concurrently to the tasting, either at predefined discrete times (D-TCATA, PP, MATI, A-TI) or over a continuous window of time (TDS, TDS-I, TCATA, TQT, TR). In dynamic methods, the subjects' reaction time

is key, this is why attributes are chosen beforehand by the experimenter and are presented as a list (see section 3.3.4 for details on attribute choice). It is evident that the challenge of keeping the attribute list manageable for the assessors and yet detailed enough for sample description and discrimination, makes attribute selection a key step when designing a TDS experiment. It should be taken into account that an incomplete or not representative list can lead to dumping effect that, in the case of TDS, would produce an illusory enhancement in the choice of one attribute as dominant only because assessors have a restrictive list from which to choose. As in traditional sensory profiling, dumping effect is especially important when a conspicuous attribute that varies across the samples was omitted (Lawless & Heymann, 2010). Retrospective measurements lose in temporal resolution but, as an advantage, they allow for a different choice of descriptors, such as free comment instead of predefined list of attributes, which could reduce the dumping effect that could result from a poor list of attributes.

3.2.3 Use of temporal sensory evaluation methods over years

All the presented multi-attribute temporal methods have not been used with the same frequency over time. Figure 8 describes the number of articles that have used each multi-attribute descriptive method by year of publication. TI is presented for the purpose of comparison.

It can be observed that TDS (and all the variants that do not include intensity rating) is the most frequently used method followed by TCATA (and related variants). Their use increased over time, showing their adoption by the sensory community. After a peak in 2021 (72 articles), a slight slowdown is observed in 2022, which seems to be confirmed in 2023 (38 articles published from January to August, not represented on Figure 8). Nonetheless, multi-attribute temporal methods did not replaced TI which has been used in a constant manner over the past three decades. Despite the recommendations given on the rating of dominance intensity (Schlich, 2017), TDS-I continued to be used after this date.

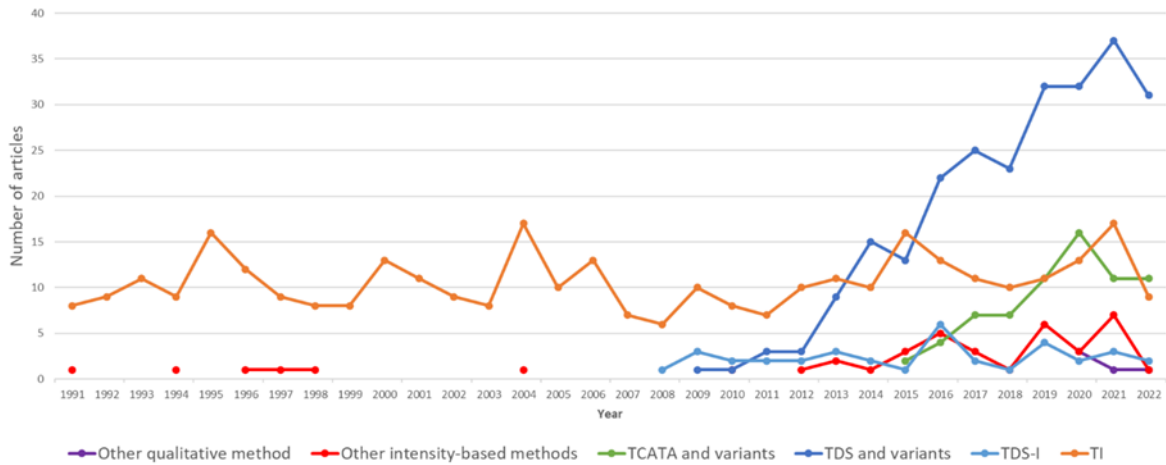


Figure 8. Number of articles published by year and type of method used.

3.3 Implementation of temporal sensory evaluation methods

3.3.1 Aim of the research

Multi-attribute temporal descriptive methods were present in papers with different objectives.

Figure 9 summarizes the main applications of temporal measurements of perception in food science, showing the central role of multi-attributes temporal evaluation methods. Most studies (75%) were interested in assessing the impact of food properties (physico-chemical, structural and microbiological properties process, composition, association of foods, etc.) on measured sensory properties, expected representative of individual sensory perception (taste, flavor, texture/mouthfeel, trigeminal sensations). The influence of several other factors on these measured sensory properties was also studied: successive consumption and quantity consumed, subjects' physiological state (saliva composition, sensitivity) and characteristic (age, gender, culture, knowledge and familiarity with food, etc.), in-mouth mechanisms (food oral processing: food breakdown, saliva flow, bolus formation; flavor release; cross-modal interactions; etc.), context of tasting (location, environmental sound, external information, temperature of the room, etc.), the details of implementation of the sensory evaluation method (type of method, training, number of subjects/attributes, definition of the task, etc.) or of the data analysis method (including data processing and interpretation of outputs). The impact of the sensory properties on other measurements was also studied: on affective properties triggered by food (liking, wanting, emotions), on perceived sensory complexity (not represented on Figure 9), on physiological state (satiety), food choices and food intakes.

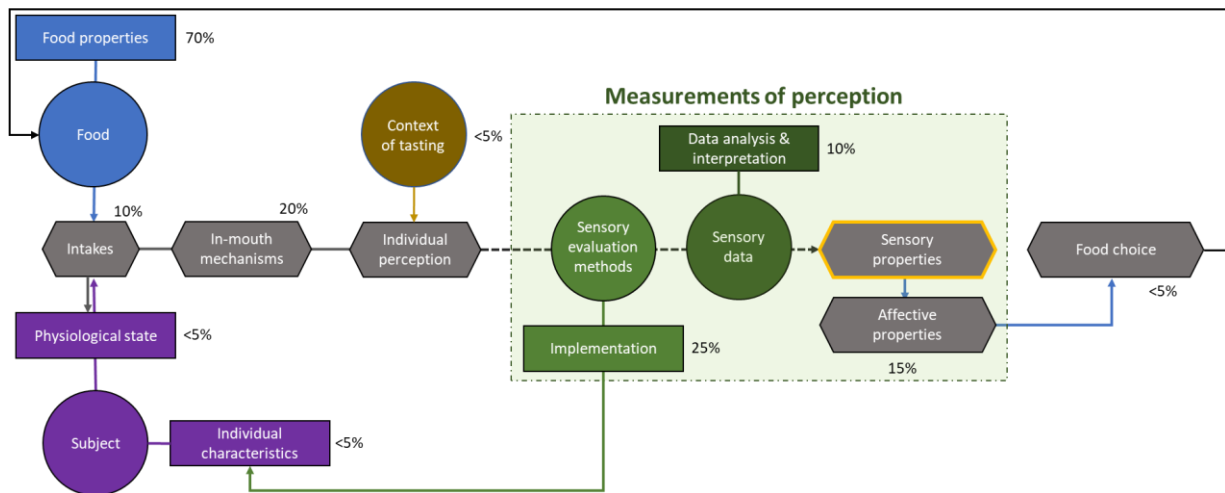


Figure 9. Main objectives of the use of temporal methods (as reported in the introduction of the articles). Numbers indicate the percentage of articles having related descriptive temporal sensory measurements with other objects of interest. An article can have multiple objectives.

It should be noticed that authors' use of keywords did not add information which could contribute to the identification of the area of knowledge studied. In most cases, keywords repeated elements of the title or were too generic. The ten most frequently used keywords were: "Consumers", "Sensory", "Sensory characterization", "Temporal methods", "Sensory analysis", "Time-intensity (TI)", "Texture", "Oral processing", "Temporal check-all-that-apply (TCATA)", "Temporal dominance of sensations (TDS)". Thus, keywords should be used as recommended by editors, avoiding repetition of words present in the title, and including relevant and controlled vocabulary (Ishida et al., 2020).

About 20% of the articles (presented in Table 8.A plus other articles having a methodological question as a secondary objective) reported methodological developments (either new methods, modifications of existing ones or methods comparison). In the past five years (2018-2022) this percentage decreased only to 15%, remaining as an area of interest showing that temporal measurements have not reached methodological maturity yet.

Regardless of the area of knowledge, in 45% of the evaluated papers, the temporal method was not the primary focus of the article but secondary to other sensory evaluation methods or to non-sensory measurements.

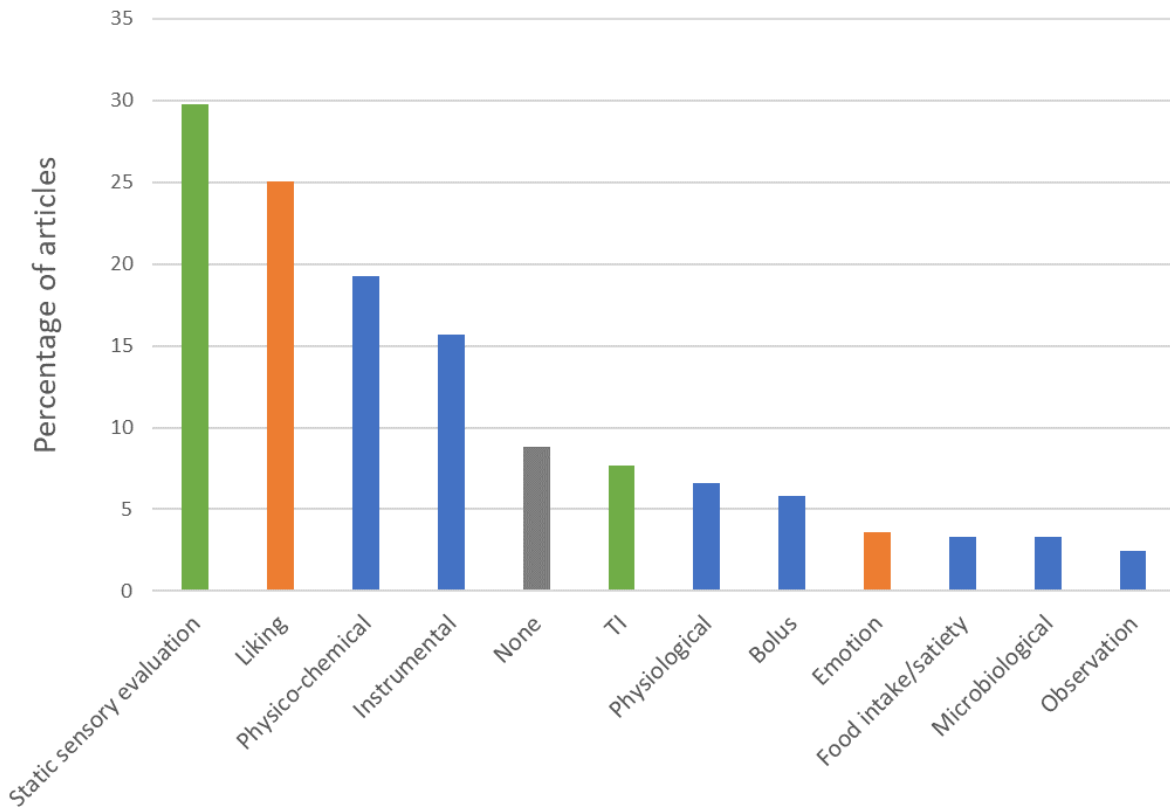


Figure 10. Other measurements found in articles which had multi-attribute temporal descriptive measurements. Green: other sensory evaluation measurement, orange: affective measurement. Different measurements can be presented in one article.

Figure 10 presents the most frequent observed measurements done together with multi-attribute temporal measurements. Most product-oriented research used also instrumental analysis (e.g. rheology), physico-chemical (e.g. chromatographic profiling) or microbiological characterization, while most subject related measurements involved surveys or observational methods. Other measurements of perception include other explicit measurements (sensory evaluation methods such as temporal or non-temporal descriptive, discriminative or hedonic tests), but also implicit measurements (EEG or face reading).

The use of such varied complementary measurements, together with the different aims of the works, shows that multi-attribute temporal methods are used for many different purposes and, probably, by sensory scientists with different levels of training and knowledge. Moreover, when possible, providing detailed information about products' composition (ingredients, nutritional facts, etc.) would allow the reusing of data for other purposes (e.g. investigating relationships between formulation and perception).

3.3.2 Type of products

The information collected on the evaluated samples included: whether there was at least minimal information characterizing them, if information on the samples was given to the assessors and whether there was information regarding the serving conditions (possible answers: yes/no/not applicable), the origin of the samples (commercial or model), their physical state (liquid, semi-solid, solid), and the global product category (e.g. chocolate, strawberries, gouda cheese).

The product categories were first transcribed as presented in the research paper and were then regrouped into more general categories (e.g. “dealcoholized wines” and “sparkling wines” were regrouped as “wines”). In this way, some detail on the products was lost but it allowed a better global representation of the information.

Figure 11 shows that among the 58 categories of products, solid (45%) and liquid products (40%) were evaluated almost with equal frequency while semi-solids (mainly dairy products) represented a minority (15%). About 1/3 were non-commercial samples including model solutions and products specifically designed for research purposes. The most frequently evaluated solid products were chocolate, cheese, bread, and fresh and deli meats. As for liquid products, those with a higher presence were wines, protein beverages, coffee, and beer. Surprisingly, while some food combinations have been studied, few composite prepared meals have been evaluated using temporal methods.

Figure 12 shows the distribution of the number of products per study. The most frequent number of products were 4 and 6. It should be noted that there were 13 studies where only one product was evaluated.

In the evaluated research articles, almost all products were evaluated in blind conditions (only four gave information on the products to the assessors). No information about the serving conditions (such as portion size, product temperature, type of light, blind/informed conditions, number of products per session, etc.) was presented in about 10% of the papers despite the importance of this information specially for the reproducibility of the experiment. Moreover, even though information on sample composition was presented, this was not as complete as expected (composition, ingredients, etc.) considering that it is better to have as much detail as possible specially for data reuse for other purposes.

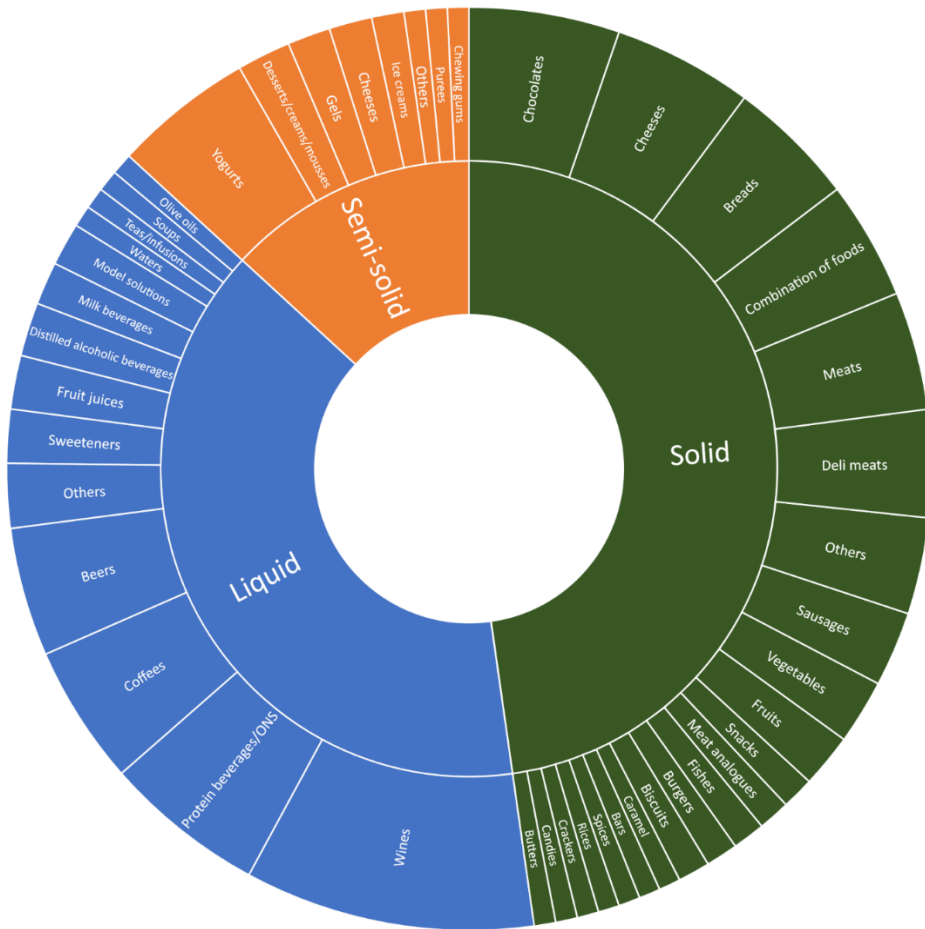


Figure 11. Categories of products evaluated using multi-attribute temporal methods. Only categories of products referenced in articles interested in the impact of food properties on sensory properties are included.

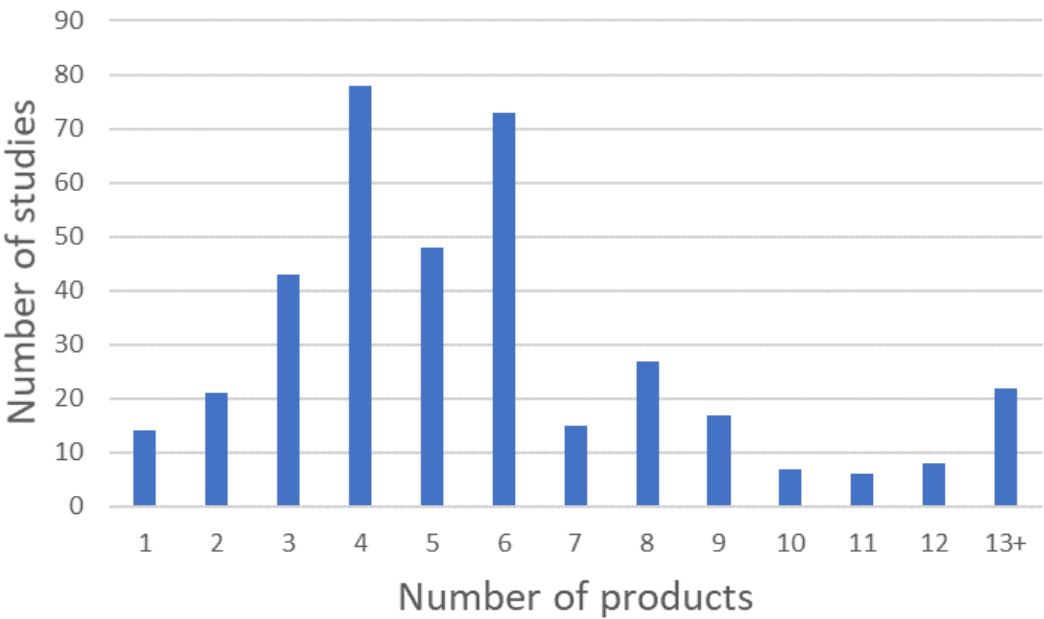


Figure 12. Distribution of number of products per study. An article can reference several studies involving different number of products.

3.3.3 Type of panels

To better understand the characteristics of the assessors participating in experiments with multi-attribute temporal methods the following information was registered from the 363 articles: whether authors had presented the criteria for subject selection, the basis for the choice of the panel size and the number of participants involved, the recruitment modalities, demographic information on the final panel, the nature of compensation (if any), the type of panel considered (consumer, semi-trained, trained, expert) and the training received.

Figure 13 shows the evolution over the last ten years of the type of panels used to carry out the multi-attribute temporal descriptive measurements. The timeline begins in 2013 because before that date almost 100% of studies used trained panels. Regardless of the type of method used, it can be observed that most evaluations were performed by assessors with some kind of training (trending from 80 to 60% in the last 10 years). Consumers panels were implemented in around 30% of the evaluations, while the use of expert panels remained marginal. As a general trend, the use of trained/semi-trained panels slowly decreases over years, while the use of consumer panels increases. However, the use of trained panels remains the norm, except with the newly introduced qualitative methods specifically designed to be used by consumer panels (in particular AEF variants). In details, overall intensity-based methods and TDS-I were used with trained/semi-trained panels in 85% of studies, TDS and variants in 70%, TCATA and variants in 60%.

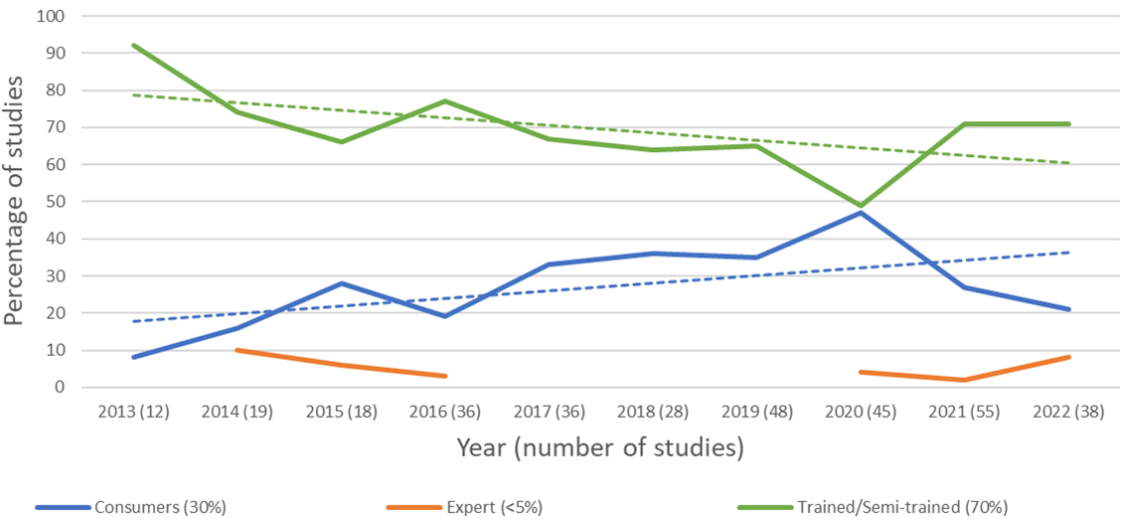


Figure 13. Types of panel involved in multi-attribute temporal descriptive measurements over the years, expressed as a percentage of the total number of studies published each year. An article can reference several studies involving different types of panels. Dashed lines represent the trend lines.

The fact that panels were considered trained, semi-trained or consumers is related precisely to the level and type of training and to what authors mentioned as the type of panel they used; however, this information was not expressed in a unified fashion across papers (Figure 14).

The definition of the panel (trained or semi-trained) was sometimes given by the authors but without being supported by the information on the type of training. In other cases, the training task was mentioned but there was no reference to the time devoted to it. This made it difficult to objectively classify and distinguish between trained and semi-trained. This is why, even though these two types of panels are different, they were grouped in the present scoping review.

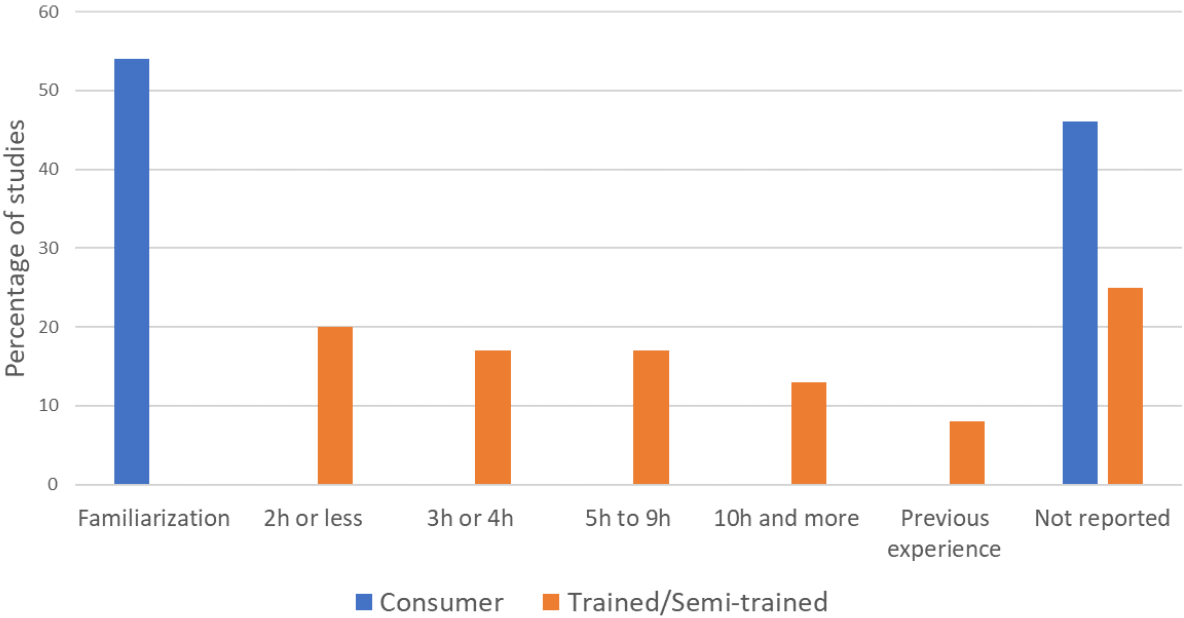


Figure 14. Details of training by type of panel.

In terms of the training, 25% of the articles working with trained or semi-trained panels did not mention any information. The papers that did report training mentioned durations that could range from very short (2h or less) to long periods (more than 10 hours). As could be expected, there were disparities between the different types of methods. For intensity-based methods and TDS-I, the most frequent training duration was more than 10 hours. For TDS and TCATA, about 50% of studies (among those having reported the information) included a training of two to four hours. Almost 10% of trained/semi-trained panels were defined as such based on previous experience with other descriptive analysis, different from the temporal method which was actually carried out. Only 15% of articles (not represented on Figure 14)

involving trained or semi-trained subjects reported results on panel repeatability. It would seem that replicates were collected to “artificially” increase the sample size and not to check panel consistency (as it is the case in static descriptive methods).

More than half of consumer panels (55%) received a familiarization training (i.e. a short introduction to the attributes and some preliminary tests before the final test, Jaeger et al., 2017; Rodrigues, de Souza, et al., 2016) before evaluating the products. However, details on the duration of this process were not presented. It is to be noted that two studies investigated the impact of familiarization on the capacity of subjects to discriminate between products, but their results were contradictory (Jaeger et al., 2017; Rodrigues, de Souza, et al., 2016).

Other than training, there are certain practices that are usually incorporated in descriptive measurements to improve the quality of the obtained data. These include giving a warm-up or dummy sample to evaluators for them to get acquainted with the method before the evaluation, using physical references to better understand and increase agreement on attribute perception and description, and giving definitions of the used attributes also to improve consensus among subjects. These practices were checked in all the evaluated papers and it was found that references were reported in 25% of studies (35% with trained/semi-trained panels), definitions in 45% (55% with trained/semi-trained panels) and warm-up in only 15%. These values are surprisingly low for descriptive methods, but it could be argued that these practices are held regularly but are not detailed in the research papers. It could be considered an activity which is part of a “previous training”. However, it is a good practice to use references and definitions to get the panel agreement and this should always be reported in order to stimulate this use in all future research.

About 10% of consumer studies were done out of laboratory. A few studies investigated the influence of data collection settings on temporal measurements (Dinnella et al., 2022; Kantono et al., 2018; Xu et al., 2019). Measuring perception in natural settings is desirable to evolve towards more ecologically valid data. However, these uncontrolled settings potentially introduce new bias that have to be identified, and replication studies are needed before generalizable conclusions can be drawn.

Figure 15.A shows the distribution of the number of subjects by study depending on the type of panel. The means of the number of subjects are about 70 for consumers, 9 for experts, and 15 for trained/semi-trained subjects. The choice of the number of assessors, regardless of the

type of panel, was not explained or justified in about 95% of the published papers. In less than 5% of the articles, the choice was made based on literature recommendations. Finally, a tiny minority (<1%) justified their choice based on power calculation.

Figure 15.B shows the distribution of the number of evaluations by study depending on the type of panel. The means of the number of evaluations are about 75 for consumers (mode = 1 replicate), 15 for experts (mode = 3 replicates), and 38 for trained/semi-trained subjects (mode = 3 replicates). The wide range evidences the lack of agreement on the most adequate number of participants/evaluations, which is more evident when working with consumers. For these, the number 70 was probably chosen in reference to the minimum number of tasters recommended for hedonic tests (Hough et al., 2006; Mammasse & Schlich, 2014) or other descriptive tests involving consumers (Ares, Tárrega, et al., 2014).

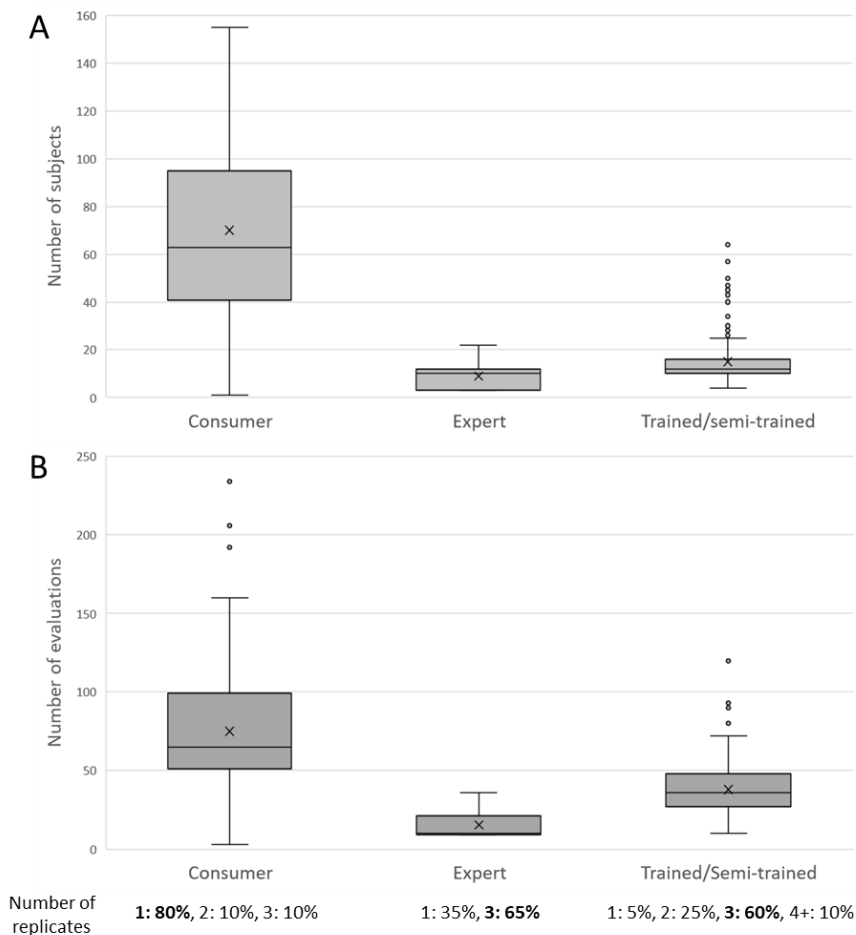


Figure 15. Distribution of the number of subjects and evaluations by type of panel. A: Number of subjects. B: Number of total evaluations (subjects by replicates). An article can reference several studies involving different numbers of subjects and evaluations.

Recommendations are scarce in the literature regarding the number of subjects needed for the evaluations. Pineau et al. (2012) recommended for TDS about 16 trained subjects and two or more replicates to have at least 30 evaluations. Relating the use of replicates with the total number of evaluations instead of the consistency of the panel might be one of the reasons why experimenters do replicates with a different goal than in static descriptive methods: looking to increase the evaluations but not checking for consistency. Cheong et al. (2014) suggested for TDS at least 15–20 untrained panellists evaluating samples in triplicate. Okamoto (2021) showed that standard error around citation rates is correctly estimated in TDS curves with samples of sizes larger than 100 (samples of size 50-100 being acceptable). Again, it seems that the conclusions on the best number of subjects and evaluations depend on various factors (sensory complexity of the product, size of the differences between products, expected temporal resolution for product characterization, etc.). In any case, it seems reasonable to consider that compared to a static sensory evaluation (in which no temporal measurements are collected), a larger number of panelists is required to draw solid conclusions related to temporal aspects of perception.

Finally, the criterion for subject selection, the recruitment modalities, the demographics of subjects and the nature of the compensation were reported in 80% (75% with trained/semi-trained panels and 90% with consumer panels), 30% (20% and 60%), 80% (70% and 90%) and 25% (20% and 45%) of articles, respectively.

The results collected from trained and consumer panels have been compared in a few studies (Hutchings, de Casanove, et al., 2017; Rodrigues, de Souza, et al., 2016; Weerawarna et al., 2021). However, the conclusions seem to depend on the method and/or the product category under consideration.

These results suggest the need to deeper investigate the question of the subjects to formulate guidelines for the choice of the type of panel, the number of subjects they should include, the type and duration of training, and the way for reporting it. Meanwhile, it is important to explicitly report information about recruitment, selection modalities, training, demographics and retribution of the subjects as well as settings of the experiment, all factors that have an impact in the repeatability of the test.

3.3.4 Attributes choice and use

The following information on attributes was gathered: how were attributes selected for the study, the name of the attributes, the sensory modalities represented, the total number of evaluated attributes, and their order of presentation. The presence of definitions and physical references was also registered, since these are also related to subjects training, results were already presented in Section 3.3.3.

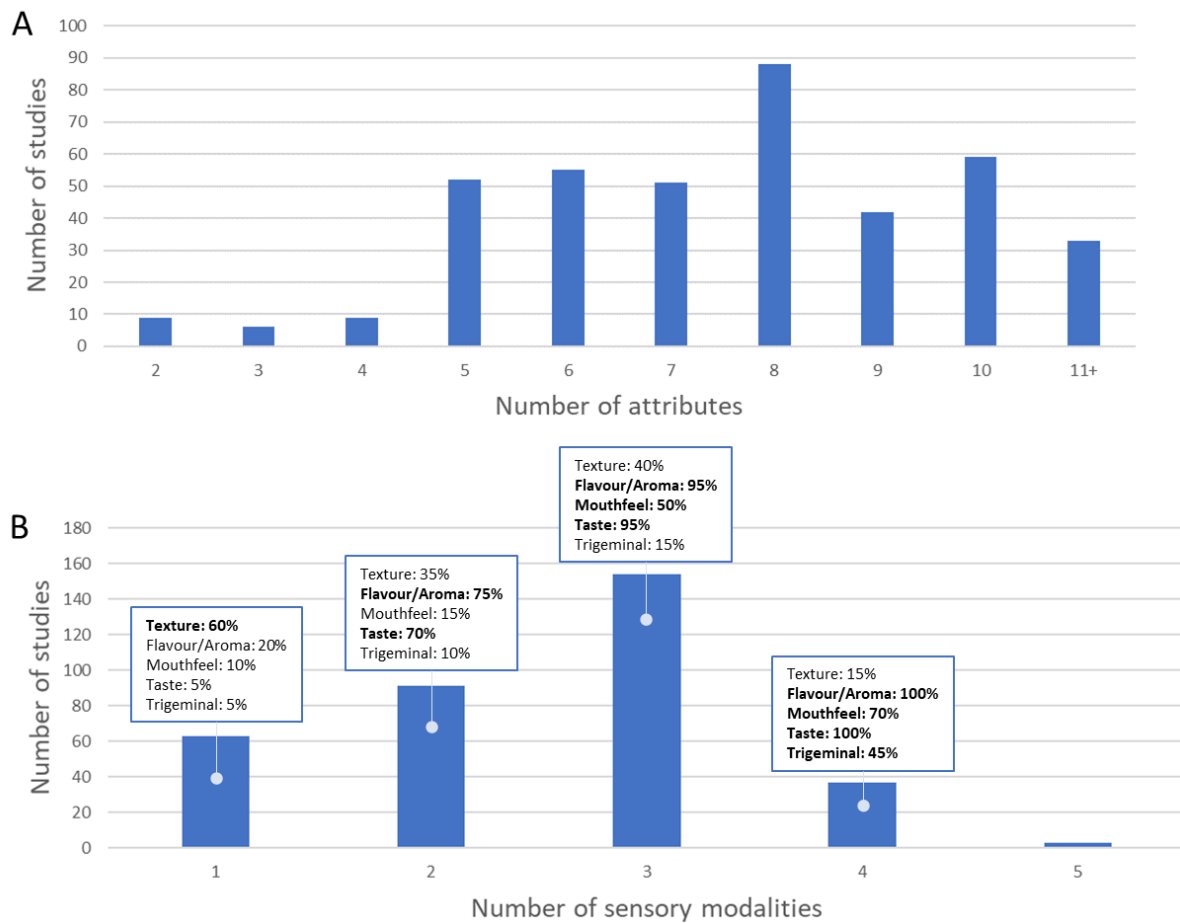


Figure 16. Use of attributes with temporal evaluation methods. A: Distribution of the number of attributes by study. B: Distribution of the number of sensory modalities by study. An article can reference several studies involving different numbers of attributes and sensory modalities.

Figure 16.A shows that most studies included five to ten attributes, the mode being eight. This is in-line with classic recommendations about the number of attributes for TDS evaluations by Pineau et al. (2012). Taking into account all the methods (Figure 16.B), the most frequent observation was the use of attributes belonging to three sensory modalities by study: flavor, taste and texture (for solid products) or mouthfeel (liquid products). Most studies involving a single modality were dedicated to food oral processing research related to texture.

The most frequently proposed attributes correspond to basic tastes (sweet in 65% of articles, bitter in 50%, sour and salty in 30%), then mouthfeel or texture attributes (astringent, soft, dry, creamy, hard, sticky, juicy, firm, metallic, melting, smooth, all between 10 and 25%). Flavor and aroma attributes are more specific and varied making it more difficult to group when analysing all the papers.

Table 10 shows the most frequently used attributes in the main product categories. A large diversity is observed in the number of attributes by product category resulting also from the number of studies which evaluated the product category; but it can give an overall idea of the complexity of the product categories. As noticed above, it has been recommended to limit the number of sensory attributes to keep the task feasible for subjects (up to 10 attributes for TDS - Pineau et al., 2012; up to 15 for TCATA - Jaeger et al., 2018). Thus, as explained in Section 3.2.2., for the most complex product categories (e.g. chocolates, cheeses, wines, etc.), the selection of the attributes can be critical. This is why in some studies (very few) the authors added an attribute “other”. An attribute “no taste” has also been added in approximately 10% of TDS studies, which amounts to allowing panellists to declare periods of non-dominance as with TDS-HD (see Table 9). Even though it represented a minority of the studies, some hedonic terms were also found as part of the descriptive list (e.g. “bad taste”, Santos Gonçalves et al., 2017). Thus, as for subjects, it is important to explicitly report information about the basis for choosing the sensory attributes, and the presence (or absence) of definitions and references to subjects.

For studies involving trained and semi-trained panels, attributes were mostly (40%) chosen by the subjects conforming the panel. Other ways of selecting the attributes involved references from literature, choice done independently by the experimenter or by another panel (15% each). In 15% of the articles, the basis for the choice were not reported. The presentation order of the attributes was reported only in 45% of the articles. Within them it was found that about 30% were randomized and 15% balanced (in accordance with the recommendations of Pineau et al., 2012 for TDS), and less than 5% fixed.

Table 10. Attributes most frequently used (used at least twice) in main product categories (those evaluated in at least five articles) evaluated by multi-attribute temporal methods. Attributes are presented in alphabetical order, not by modality.

Product category	Main attributes ¹
Beers	Alcohol, astringent, bitter, carbonated citrus, chocolate, citric, coffee, estery, floral, fruity, full-body, grassy, herbal, honey, hoppy, lemon, malty, refreshing, sour, spicy, sweet, tingly, toasted, toffee, warming
Biscuits	Buttery, crispy, crumbly, crunchy, dry, hard, sticky
Breads	Aerated, bitter, butter, cardboard, bread, chewy, coarse, compact, crispy, crumbly, crunchy, dense, doughy, dry, fermented, grilled, hard, hydrated, metallic, off-flavour, roasted cereals, rough, salty, smooth, soft, sour, spongy, sticky, sweet, toasted, wet flour, wheat
Cheeses	Semi-solid: Astringent, bitter, buttery, cheese, (cooked herbs), cream, creamy, crumbly, dry, firm, fresh (herbs), (garlic), gummy, grainy, off-flavour, peppery, pungent, salty, sharp, soft, sour, spicy Hard: Bitter, brittle, buttery, cheese, creamy, crumbly, dry, fatty, firm, fresh herbs, fruity, garlic, grainy, greasy, gummy, hard, melty, milky, mouth-coating, off-flavour, pungent, rancid, rubbery, salty, sharp, smooth, soft, sour, spicy, sticky, sweet, thick
Chewing gums	Bitter, fresh, (mint), (peppermint), sweet
Chocolates	Adhesive, astringent, bitter, brittle, buttery, caramel, chewy, chocolate, cocoa, coffee, creamy, crispy, crumbly, crunchy, dairy, dry, fruity, gooey, grainy, hard, melting, milky, mouth-coating, nutty, off-flavour, powdery, roasted, smooth, soft, sour, springy, sticky, sweet, vanilla, woody
Coffee	Acidic, almond, astringent, bitter, burnt, caramel, chocolate, cocoa, coffee, fruity, herbaceous, nutty, roasted, sour, sugar cane, sweet, tobacco, woody
Deli meats	Bitter, chewy, crunchy, cured, fatty, fibrous, firm, greasy, gummy, ham, hard, juicy, meat, off-flavour, pungent, rancid, salty, smoky, soft, spicy, succulent, tacky, tender, umami
Milk desserts	(Caramel), cream, creamy, milky, off-flavour, soft, sweet, thick, (vanilla)
Distilled beverages	Alcohol, burning, caramel, fruity, green, sweet, vanilla, woody
Fruit juices	Acid, bitter, astringent, bitter, (grape), off-flavour, (orange), sweet
Fruits	Astringent, bad taste, bitter, crunchy, fermented, fruity, hard, juicy, metallic, refreshing, soft, sour, sweet, tasteless
Gels	Bitter, creamy, crumbly, elastic, grainy, melting, moist, refreshing, smooth, sticky, sweet
Ice creams	Bitter, (cocoa), cold, creamy, icy, milky, roasted, sweet, (vanilla)
Meats	Browned, dry, fatty, fibrous, firm, juicy, livery, meaty, oily, oxidized, smooth, soft, sweet, tender, though/hard, umami
Milk beverages	Astringent, (cocoa), creamy, licorice, milky, mouth-coating, sweet, thick, (vanilla)
Protein beverages & Oral Nutritional Supplements	Astringent, bitter, (caramel), cardboard, (coffee milk), cooked, creamy, drying, filming, metallic, mouth-coating, (praline), salty, sweet, thick, (vanilla)
Sausages	Chewy, dry, fatty, firm, grainy, hard, juicy, meaty, salty, slippery, smooth, soft, spicy
Sweeteners	Bitter, chemical, drying, licorice, metallic, sour, sweet
Vegetables	Adhesive, astringent, bitter, firm, juicy, pungent, salty, smooth, sweet, vegetable
Wines	Acid, alcohol, adhesive, animal, apple, astringent, banana, bitter, black fruit, burnt, citrus, dark fruit, dry, drying, earthy, floral, fruity, grainy, green, grippy, heat, herbaceous, hot, mouthcoating, pineapple, puckery, pungent, red fruit, rose, smoky, sour, spicy, sweet, vegetal, woody, yellow fruit
Yogurts	Acidic, artificial, astringent, bitter, (caramel), cloying, cold, cream, creamy, fatty, fermented, (lemon), licorice, melting, metallic, milky, off-flavour, sour, sticky, (strawberry), sweet, thick, thin, (vanilla), viscous

¹Between brackets: attributes depending on particular flavouring within the product category.

3.3.5 Experimental design

To explore the uses of multi-attribute temporal methods in terms of experimental design the following information was checked: the temporal unit (within intake, e.g. evaluation of changes in perception during a sip, or between intake, e.g. evaluation of changes in perception over successive intakes), the duration of the tasting and the way in which it was standardized (e.g. free or fixed duration, free or fixed way of eating/drinking), and the product presentation order.

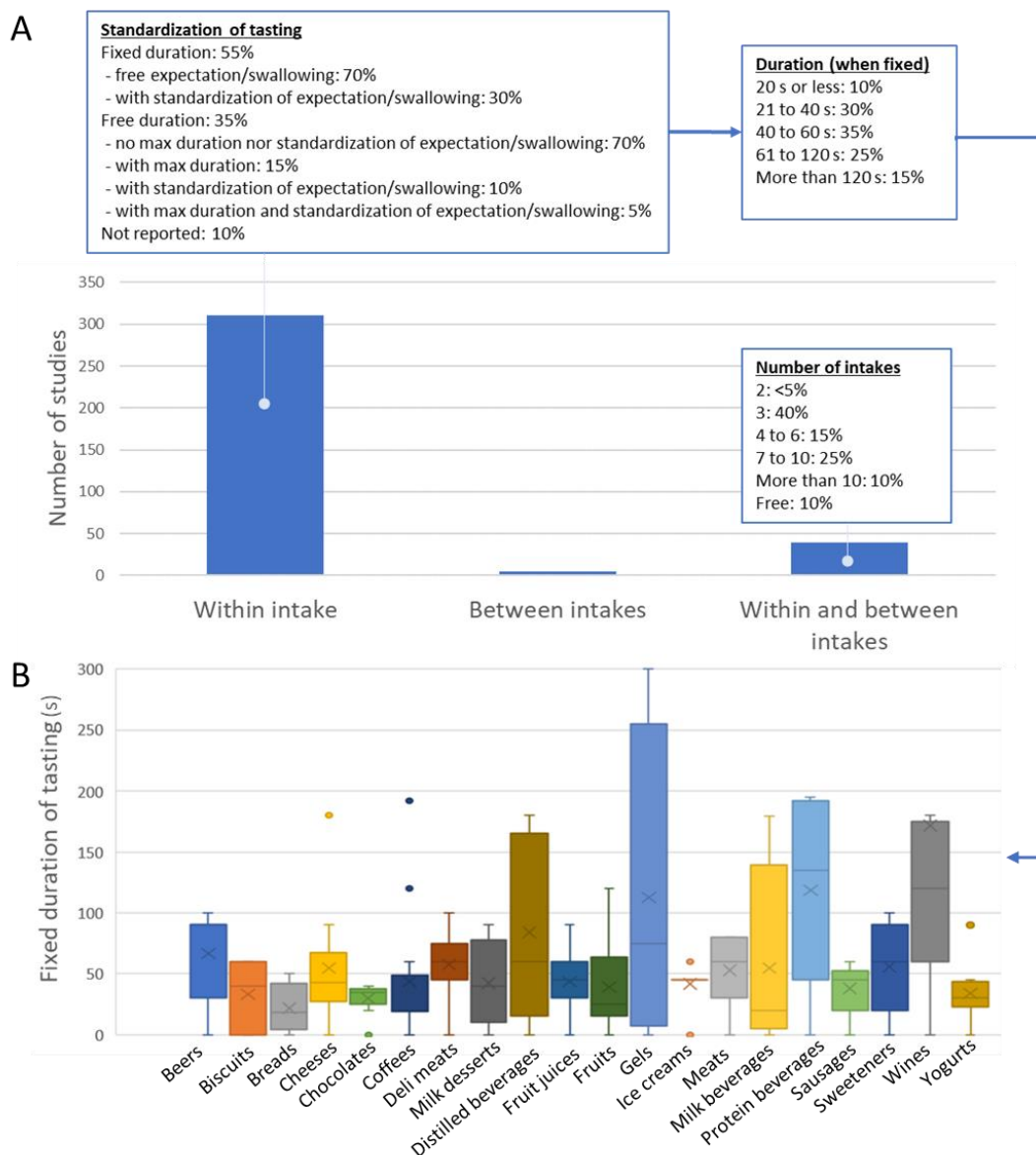


Figure 17. Characterization of the tastings. A. Characterization of the tastings by temporal unit. B. Distribution of durations of evaluation within-intake, for the main product categories.

Figure 17 describes how the temporal descriptive tastings were implemented from a practical point of view. Most articles (85%) focused on the evolution of sensations within intakes, while 10% were interested also in the multi-intake approach and less than 5% focused only on the temporal description among intakes (Figure 17.A).

Looking at those articles that evaluated the temporality within intakes, it was observed that the duration of the tasting was free (remained to the choice of the subject) in 35% of studies. Among them, a maximum duration was imposed in 15%, and the moment to swallow or expectorate was standardized in 10%. The duration of the tasting was imposed (fixed) in 55% of studies, and among them 30% standardizing the moment of swallowing or expectoration. Deciding on a fixed or free duration of the intake can be determined by different factors, each having advantages and disadvantages. A free duration of the evaluation can be chosen, for example, when looking for a way of consumption closer to natural conditions. However, it will require certain posterior data transformation to unify temporal criteria among subjects. With a fixed duration, no posterior data treatment is necessary, but some information can be lost when choosing the duration. Even though this kind of choices are interesting and can have an impact on results, they were rarely explained or shared by researchers.

The fixed durations for the evaluations chosen by the researchers with an interest in the temporality within the intake, varied from very short (less than 20 s, 10% of studies) to more than six times that duration (more than 120 s, 15% of studies), but most durations (65%) were fixed between 21 and 60 s. Figure 17.B shows that very different durations have been reported for products in the same category, notably with liquid ones. It is not possible to know whether these variations correspond to differences in duration between products of the same category or to differences in the implementation of the protocol. As with the descriptors, the choice of duration is critical, especially in TDS since the last descriptor selected is considered as dominant up to the maximum duration chosen (in the absence of a STOP button).

Still considering articles that evaluated the temporality within intakes, the products were presented in a balanced or randomized order in 40% and 30% of the studies, respectively. The order was not reported in 30% of articles.

In real life settings, food products are rarely consumed on a single intake, but rather over multiple ones. This is why temporal methods also extend to study the evolution of sensory perception over consecutive intakes, looking for something more representative of natural

eating behaviour. SP (Table 9) was developed for this purpose, but almost all temporal methods have been used to characterize temporal dynamics of several intakes, even though, as previously stated, this type of temporality interested only 15% of the studies. In these studies, the number of evaluation ranges between 2 and 30, the mode being 3 (40% of studies, Figure 17.A).

Table 11 presents the details of implementation of temporal methods with products evaluated over successive intakes. Most studies focused on liquid products presenting bitter or astringent compounds known to have a build-up sensation. Overall, the implementations vary largely within and between product categories, in the number of intakes, the duration of evaluation of intakes and in the quantity of product consumed. In about 40% of articles only three intakes were evaluated (Figure 17.A), probably for practical reasons especially when working with alcoholic beverages. In this case, the tasting can hardly be considered as representative of a full product consumption. As it is plausible that a minimum quantity is necessary to observe built-up effects, it could explain why some articles (Table 11) did not conclude on the multi-intake measurements.

Table 11. Details of implementation of temporal methods with products evaluated over successive intakes (categories of products evaluated at least in three articles).

Product category	Method ¹	Type of Panel ²	Temporal Unit ³	Number of intakes	Total quantity	Duration by intake	Reference	Conclusion related to evolution of product description over multiple intakes (as reported in abstract)
Beers	AEF-A D-TCATA	C	WB	7	Free	Free	Visalli et al., 2022	No difference in perception over intakes
Beers	SP	T	B	5	Free	NA	Vázquez-Araújo et al., 2013	-
Beers	M-TCATA	T	WB	2	40 ml	90 s	Dietz, Cook, et al., 2022	Limited effects were observed between sips
Beers	TDS	C	WB	6	120 ml	50 or 90 s	Simioni et al., 2018	Increased dominance of bitterness, decreased dominance of fruity, floral, toffee, and coffee
Beers	TDS	C	WB	Free	80 ml	Free	Silva et al., 2019	-
Beers	TDS	E ; C	WB	Min 3	E: 350 ml C: 500 ml	30 s	Wakihira et al., 2020	Fewer built-up effects with less standout flavor beers
Beers	TDS	C	WB	15	330 ml	Free	Machado et al., 2023	Duration of perception gradually decreased
Protein beverages/ONS	SP	T	WB	8	40 ml	60 s	Methven et al., 2010	Built-up of mouthdrying, mouthcoating, metallic and soya
Protein beverages/ONS	SP	T	WB	30	600 ml	90 s	den Boer et al., 2019	Mouth-drying first increased, up to a consumption volume of 300 ml, and then decreased
Protein beverages/ONS	SP	T	WB	8	40 ml	20 s	Withers et al., 2014	Built-up of mouthdrying
Protein beverages/ONS	SP	T	B	8	120 ml	NA	Lester et al., 2021	Built-up of mouthdrying and higher aftertaste perception
Protein beverages/ONS	SP	T	WB	8	40 ml	60 s	Bull et al., 2017	Built-up of mouthcoating, drying and chalky
Protein beverages/ONS	TDS	C	WB	Free	Free	Free	Thomas et al., 2016	-
Protein beverages/ONS	TDS	C	WB	10	Free	Free	Thomas et al., 2018	-
Protein beverages/ONS	TDS-I	T	WB	3	120 ml	Free	Cosson et al., 2020	Built-up of fatty, decrease of beany and bitter
Wines	DA	E	B	7 over 30 minutes	50 ml	NA	Lytra et al., 2016	Evolution of fruity notes
Wines	SP	T	WB	4	20 ml	60 s	Olatujoye et al., 2020	Built-up of astringency
Wines	TDS	C	WB	Free	80 ml	Free	Silva et al., 2018	-

Wines	TDS-D	T	WB	3	30 ml	90 s	Pittari et al., 2022	-
Yogurts	M-TDS	T	WB	3	40 ml	30 s	Lesme et al., 2020	Global flavor perception of the samples varied with the number of spoons, which particularly impacted the taste attributes
Yogurts	SP	T	B	2	2 or 4 spoons	NA	Palczak et al., 2020	-
Yogurts	TDS	T	WB	3	15 g	20 s	Souza Olegario et al., 2022	Numbers of intakes presented a significant impact on temporal perception
Yogurts	TDS	C	WB	5	60 g	Free	van Bommel, Stieger, Boelee, et al., 2019	Built-up of dominance for sticky
Yogurts	TDS	T	WB	3	50 g	45 s	Chadha et al., 2022	Increased dominance of bitter and astringent

¹See Table 9 for abbreviations; ²T: trained/semi-trained, C: consumers; ³WB: evaluation within intakes and between intakes, B: evaluation only between intakes

3.3.6 Data acquisition software

Some methods are more used in certain geographic regions or with certain software (comparatively to the relative use of methods), notably TDS with SensoMaker and TimeSens, and TCATA with CompuSense. These results can probably be explained by the geographic proximity between the software distributors and their customer base. They also suggest an influence of the software on the use of the methods, either indirect (availability or non-availability of the method in the software) or direct (promotion of methods by software distributors). This can be explained by the academic competition between the few teams (including that of the first author of this review) involved in methodological and/or software development. Indeed, among the 1027 different contributors to the articles included in this review, 25% of the 363 articles were co-authored by three researchers who are directly or indirectly involved in the development or promotion of a software (TimeSens, CompuSense or Sensomaker). This entanglement between software and methods can add a bias to the choice of temporal methods in research and the conclusions reached in methodological articles (see Section 3.5).

3.4 Analysis of temporal sensory data

To have an overview of how multi-attribute temporal sensory data was analysed in the literature, the following information was registered: if data was transformed and how, which were the variables and statistical analyses used, if there were inferential statistics and values for alpha determined beforehand, and the software used for data analyses.

3.4.1 Main variables and data transformations

It was previously shown that temporal methods have evolved from the use of quantitative intensity scales to a qualitative evaluation. It seems that the methodologists constantly seek to find the best compromise between the level of detail of the data collected and the difficulty of the evaluation task for the panelists (e.g. intensity for one attribute with TI, then intensity for several attributes with DATI/MATI, then intensity for the dominant attribute with TDS-I, then only dominance, then applicability for several attributes with TCATA, then applicability with no need for uncheck with TCATA-F, then applicability per period with AEF-A, etc.).

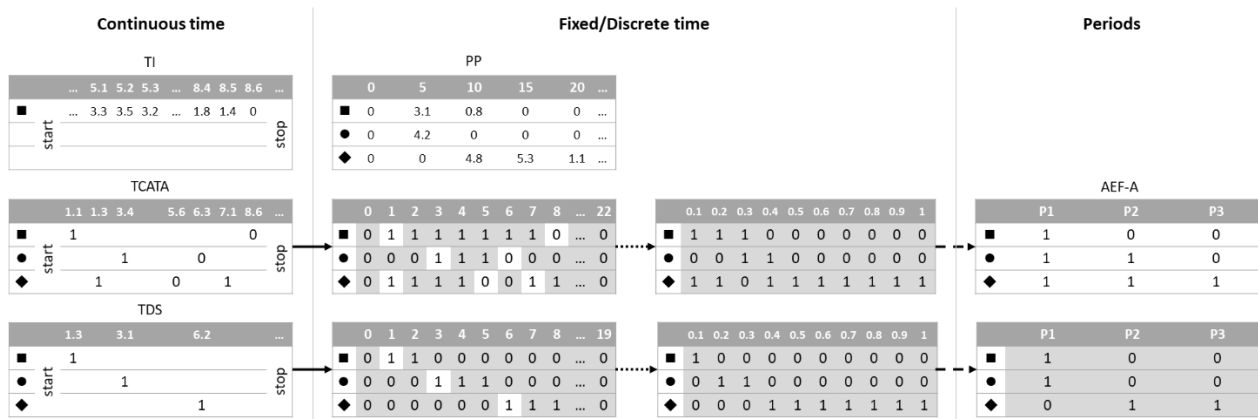


Figure 18. Examples of data collected at continuous times (intensity, dominance and applicability), fixed pre-determined moments of consumption (progressive profile) and recapitulative time periods. Plain arrows represent discretization; dotted arrows are for time standardization and dashed arrows for transformation in periods. Gray cells: imputed data. Different symbols represent sensory attributes.

Figure 18 shows what the collected data look like and how they are possibly transformed. With intensity-based methods such as TI and PP (but also DATI, SP, or MATI), all presented attributes are rated at any given time on linear scales derived from classical descriptive analysis (presented on a x-y plan in DATI). Only TI and DATI data are continuous, panelists having to move the cursor constantly. For other methods, data are fixed interval data (times of evaluation are imposed). Data are directly stored as *attribute x time* matrices of intensity scores (one by *product x subject*). The same goes for retrospective methods such as AEF-A (but also FC-AEF-A, AEF-D, Quessence and F-TOS), replacing intensity scores by 1 or 0 standing for presence or absence (0: not dominant/applicable, 1: dominant/applicable). For TDS and TCATA, it is quite different as data collection relies on stochastic processes (non-fixed interval data, times of evaluation are chosen by the panelists). Collected data consist in an ordered sequence of events constituted of two random variables: the selected attribute (dominant/applicable attribute), the time of click on the attribute, and the value corresponding to the state of the attribute (0/1 for dominant/applicable attributes, intensity of the dominant attribute with TDS-I). For the purpose of data representation and statistical analysis (see 3.4.2), these events are transformed in discrete time series assumed as continuous if the discretization step is small enough. Some variable transformations are justified by the intrinsic complexity of dynamic methods which add a new source of uncontrolled variability in data. To disregard subjects' temporal signatures or heterogeneity in times to first citations (Tfirst) and total durations (Dtot), temporal data can be standardized

by subject (Lenfant et al., 2009), i.e. transformed between 0 (corresponding to T_{first} of the subject) and 1 (corresponding to D_{tot} of the subject). Time standardization is relatively frequent in TDS (40% of studies), less in TCATA (20%). Other treatments are probably related to the scientific community's greater predilection for tests with parametric and quantitative analyses. For example, the division of time into periods (between 3 and 20, being 3 or 4 periods the most frequent) is used in 10% of TDS studies and 20% of TCATA. It is to be noticed that the transformation of TCATA and TDS in periods makes the structure of data similar to those of AEF-A, and for TDS several dominant attributes can be cited within a given period.

Table 12 shows the main variables related to data collected with temporal methods. Among these 31 variables, only nine are primary (those suffixed with a star in Table 12), that is to say that they were directly collected and can be observed in data matrices (see Figure 18). These primary variables are rarely analysed as such, but other variables are derived from them using computations either at the subject or panel level. For example, in TDS durations are not directly asked to the panelists, but computed as the difference between two successive citations of distinct attributes. At the panel level, only citation rates (CR_t) and mean intensities ($IMEAN_t$) are computed at each time. These two variables have been used in more than 90% of the articles. The variables aggregated by period require prior data transformation (Figure 18) that necessarily results in a loss of temporal resolution. The aggregation at global level is an inheritance of TI curves parameters (e.g. AUC, TMAX, DMAX, RINC, RDEC, IMAX/CRMAX, see Table 12).

Table 12. Main variables (used at least in two articles) derived from data collected with temporal methods and the abbreviations used in the present article for each by time period and globally.

Variable	By time	By period	Global
Variable collected or computed at the subject level with intensity-based methods			
Intensity, by attribute*	I_t	I_p	-
Mean intensity, by attribute	-	I_{mean_p}	I_{mean}
Max intensity, by attribute	-	I_{max_p}	I_{max}
TDS score, by attribute (duration std/no std x intensity)*	-	DI_p	DI
Duration of perception (intensity>0), by attribute*	-	D_p	D
Variable collected or computed at the subject level with qualitative methods			
Citations, by attribute* (dominant/applicable=1, not dominant/not applicable=0)	C_t	C_p	C
Rank of citation, by attribute*	RC_t	-	RC
Time to first citation*	-	-	T_{first}
Time or period to first citation, by attribute*	-	-	T
Citation, after period or static transformation, by attribute	-	C_p	C
Number of citations, by attribute	-	N_p	N
Duration (std/no std) of applicability/dominance, by attribute	-	D_p	D
Sojourn time (std/no std) of applicability/dominance, by attribute	-	S_p	S
Number of attributes cited	-	N_{att_p}	N_{att}
Number of citations	-	N_{cit_p}	N_{cit}
Variable collected* or computed at the subject level common to intensity-based and qualitative methods			
Total duration of perception*	-	-	D_{tot}
Number of intakes* (if applicable, see Implementation)	-	-	N_{int}
Variable computed at the panel level with intensity-based methods			
Mean intensity, by attribute (mean of I_{mean})	$IMEAN_t$	$IMEAN_p$	$IMEAN$
Max intensity, by attribute (max of $IMEAN$)	-	$IMAX_p$	$IMAX$
Variable computed at the panel level with qualitative methods			
Citation rate (or frequency), by attribute	CR_t	CR_p	CR
Mean citation rate, by attribute	-	CR_{MEAN_p}	CR_{MEAN}
Max citation rate, by attribute	-	CR_{MAX_p}	CR_{MAX}
Duration above significance (TDS), by attribute	-	$DSIG_p$	$DSIG$
Variable computed at the panel level common to intensity-based and qualitative methods			
Mean duration, by attribute (mean of D)	-	D_{MEAN_p}	D_{MEAN}
Mean sojourn time, by attribute (mean of S)	-	S_{MEAN_p}	S_{MEAN}
Number of transitions from one attribute to another, by attribute	-	TR_p	TR
Area under the curve (CRT or $IMEAN_t$), by attribute	-	AUC_p	AUC
Time to CR_{MAX} or $IMAX$, by attribute	-	T_{MAX_p}	T_{MAX}
Duration above 90% of CR_{MAX} or $IMAX$, by attribute	-	D_{MAX_p}	D_{MAX}
Rate of increase, by attribute	-	$RINC_p$	$RINC$
Rate of decrease, by attribute	-	$RDEC_p$	$RDEC$

*: Collected variables; std: time standardization.

Figure 15 and Table 12 show that what is analysed is different from what is actually collected during product evaluation. The primary variables undergo many transformations and/or aggregation, sometimes unnecessary and/or excessively manipulating the nature of the data, with an unclear impact on results. Time standardization distorts individual sequences (Meyners, 2020) while there is no consensual evidence on the beneficial effects of this transformation. For TDS, it has been shown that time standardization complemented analysis on raw data (Lesme et al., 2020) or highlighted most differences between products (Frost et al., 2018). For TCATA, the opposite effect was observed (Dietz, Yang, et al., 2022). For transformation into periods, the choice of the number and duration of periods was almost always arbitrary, except for Lecuelle et al. (2018) who automatically determined them. If this transformation simplifies data analysis, it is however unlikely that periods should be chosen of uniform sizes, and the choice can have an impact on conclusions (Beaton & Meyners, 2020).

3.4.2 Main statistical analyses

Table 13 shows that numerous statistical analyses (more than one hundred were found, only those used more than once are presented) have been used for gaining insights about temporal data. The data analyses have been classified into two categories: exploratory (including techniques used to investigate the data and summarize the main characteristics of the dataset) and confirmatory (including techniques based on statistical inference or parametric and non-parametric tests to decide whether or not the data support a particular hypothesis). Regarding the assessment of global differences between products, the most used analyses were univariate linear models such as ANOVAs (20% of the articles) or multivariate maps such as PCA (10%) or CVA (10%). These analyses were derived from those generally applied with static descriptive analysis methods and applied on citation rates, durations or intensities aggregated by period or as area under curves, thus ignoring the dynamic dimension in the data. “Trajectory maps” (mostly CA and PCA) were also used to represent the evolution of variables by period (15%).

With intensity-based methods, there was a prevalence of parametric procedures in the analysis of results. For TDS and TCATA the most used analyses (more than 80% of articles) were the plots of citation rates over time. These plots can be used to assess temporal evolution of the agreement among panelists regarding the dominance or applicability of a specific attribute. TDS proposes determining the significantly dominant attributes within a

product based on the comparison of citation rates of attributes in relation to a significance threshold. . As this test does not allow statistical comparisons between products or attributes to be made, it was considered as an exploratory analysis. For TCATA, there is no such significance threshold, and the significantly applicable attributes within a product are determined by comparison to all other products (as with difference plots). Difference plots of citation rates over time were used in about 20% of the articles. For intensity-based methods, plots of intensity over time were used in 20% of the articles. All other analyses were used in less than 10% of the articles, probably because they are not available in main commercial software.

These results denote that - except for the plots of citation rates/intensity over time - there is little agreement on the “must do” analyses. As with data acquisition software (and probably even more) it is likely that the choice of the data analysis depends on its availability and its easiness of use. As a result, primary variables were rarely analysed as such, thus very few articles considered individual differences in temporal perception. Likewise, few articles had statistical analysis related to subjects’ behaviour or panel performance. Half of the articles based their conclusions exclusively on exploratory data analyses (e.g. visual inspection of curves), inappropriate to draw robust conclusions on product comparisons (Meyners, 2020) and submitted to subjective interpretations. The other half used confirmatory data analyses enabling an objective interpretation based on a statistical criterion. Among them, 20% considered the sequentiality of perceptions (see Table 13, Assessment of temporal differences between products, confirmatory analysis are suffixed by a star), the main interest of temporal measurements. About 40% of articles reported the alpha risk in the data analysis section prior to present results, and almost none mentioned the size effect. The way of reporting statistical results on multi-attribute temporal methods should evolve to follow recent recommendations on better practices (Aguinis et al., 2021; Johnson et al., 2020).

Table 13. Main statistical analyses reported in the literature (cited at least twice) grouped by category.

Method category	Exploratory and confirmatory data analysis methods	Variables ¹	Number of articles ²	Selection of references ³
Checking for subjects' behaviour and performance				
Qualitative	Linear models* (ANOVAs, t-tests, etc.) for comparison of subjects' behaviour	Tfirst, Dtot, Natt, Ncit	19/333	-
Qualitative	Repeatability/ Agreement index	C, C _t	5/333	Castura et al., 2016; Dietz, Yang, et al., 2022; Fiches et al., 2016; Hutchings, Foster, Hedderley, et al., 2014; Poveromo & Hopfer, 2019
All	Linear models* for assessment of performances	I, D, C, CR, CR _p , RC, index	22/363	Dietz, Cook, et al., 2022; Dinnella et al., 2012; Hutchings, de Casanove, et al., 2017; Hutchings, Foster, Grigor, et al., 2014; Keefer et al., 2022; Kuesten et al., 2013; Lepage et al., 2014; Mesurolle et al., 2013; Nguyen et al., 2018; Palczak et al., 2019; Visalli et al., 2016
Qualitative	Plot of citation rates by subject or plot of differences* in citation rates over times to assess panellist or panel repeatability	CR _t	4/333	Patterson et al., 2021; Visalli et al., 2016; Young et al., 2013
Qualitative	Randomization tests* for assessment of performances	-	2/333	Meyners & Castura, 2018; Meyners, 2011
Assessment of temporal evolution within product				
Dominance	Plot of citation rates over time/period, with comparison to chance, by subject or attribute	CR _t , CR _p	241/282	Missbach et al., 2017; Pineau et al., 2009; Visalli et al., 2020
TDS	TDS bandplots	CR _t	24/281	Galmarini et al., 2017
TDS	Graph of transitions	TR, TR _p	3/281	Castura, 2020; Lecuelle et al., 2018
Assessment of global differences between products (not based on temporal evolution)				
All	PCA	CRMEAN, CRMAX, DMEAN, AUC, IMAX, IMEAN, DI, TMAX, DMAX	33/363	-
Qualitative	CA	CR	6/333	-
Qualitative	Canonical/Conditional CA	CR	3/333	Beaton & Meyners, 2020
All	Linear models*	D, CR, C, I, AUC, AUC-Sig, RDEC, RINC, T, IMAX, DI, TMAX, CRMAX, CRMEAN, RC, Tfirst, Dtot, Natt, Ncit	75/363	-
All	Non-parametric tests* (Cochran Q-test, Friedman test)	CR, D, I	7/363	-
Qualitative	Bootstrap tests*	CR	2/333	Okamoto, 2021; Shimaoka et al., 2022
All	PARAFAC	CR	2/363	Rodrigues, Condino, et al., 2016
All	PLS-R	CR, I	4/363	Kang et al., 2020; Pu et al., 2019

All	MANOVA*/CVA*	D, DI, I, DMAX, TMAX, CRMAX, CRMEAN	36/363	Galmarini et al., 2016
Intensity	MAM-CVA*	I	2/39	Kang et al., 2019
TDS	Semi-markov models*	S, TR, TR _p	3/281	Cardot et al., 2019; Frascaola et al., 2022; Kurata et al., 2022; Lecuelle et al., 2018
Qualitative	Randomization tests*	-	3/233	Meyners, 2020; Meyners & Castura, 2019; Meyners & Pineau, 2010)
Qualitative	mrCA* + hypergeometric test	CR	3/233	Mahieu et al., 2020
All	PCA* with partial/total truncated bootstrap*	CR _t	2/333	Castura et al., 2022
Assessment of temporal differences between products				
Qualitative	Plot of citation rates over time, with comparison to all other products	CR _t	56/333	Castura et al., 2016; Dietz, Yang, et al., 2022
All	Trajectory PCA by period	IMEAN _p , CRMEAN _p , DMEAN _p	40/363	Galmarini et al., 2016; Lenfant et al., 2009
Qualitative	Trajectory CA/MRCA by period	CR _p	13/333	Castura et al., 2016; Visalli et al., 2020
All	Linear models by period*	D _p , CRMEAN _p , C _p , I _p , CRMAX _p	15/363	-
Qualitative	Plot of differences in citation rates over times*	CR _t	69/333	Castura et al., 2016; Pineau et al., 2009
Intensity	Intensity curves	I _t	9/39	Kuesten et al., 2013; Methven et al., 2010; Zimoch & Findlay, 1998
Clustering				
All	HCA	Coordinates of PCA	5/363	Lorido et al., 2018

*: Confirmatory data analysis; ¹See Table 12 for abbreviations; ²Number of articles: actual use/potential use; ³References are reported only for analysis specifically developed or adapted for temporal sensory data.

3.5 Comparison of methods

Comparison between methods was done in 15% of the articles (for 55% of them comparison was a primary objective while for the other 45% it was secondary). To evaluate how the comparisons were carried out, seven criteria were considered (taking into account what was presented more often in the concerned papers): the overall differences (“Different”), similarities (“Similarity”) or complementarities (“Complementary”) between methods (reported or not), plus conclusion (“+” or “-“) about which method is “better” regarding their capacity to highlight temporal patterns within product (“Temporal”), to give complete description of products (“Description”), to discriminate between products (“Discrimination”), and to give consensual results at the panel level (“Agreement”). The criteria were evaluated based on the conclusions reported by the authors. Example: “TDS and TCATA provided *comparable information* for the key sensory attributes characterizing and differentiating the regular and sodium-reduced products. TDS was *more discriminative* than TCATA for single-product intakes, while TCATA generated *more consistent* profiles across multiple intakes.” (from Nguyen & Wismer, 2022) was summarized as “TCATA-F vs. TDS: Similar; Discrimination; Agreement+”.

Table 14 shows that most method comparisons involved TDS/TDS-I vs. DA, TDS vs. TI and TCATA vs. TDS. TDS and TDS-I were declared complementary to DA in about two articles out of three stating that TDS add a temporal dimension (something that could be expected without carrying out any experiment, given the nature of the methods). Similarities were also reported between TDS-I and DA in the same proportion (probably linked to the fact that both are descriptive methods). No clear conclusions emerged from comparisons between TDS and TI. TDS and TCATA were judged as complementary in about one article out of two, while one article over three said that TCATA provided a better product description. Among all the articles comparing methods, “negative” findings (e.g. less discrimination/information or unexplained differences) concerning the more recent method have been reported in less than 20% of articles, which could be symptomatic of a publication bias (Nair, 2019).

Table 14. Main conclusions reported in articles involving method comparisons (one article can include multiple comparisons).

Method 1	Method 2	Number of studies	Different ¹	Similar ¹	Complementary ¹	Temporal ²	Description ²	Discrimination ²	Agreement ²
AEF-A	D-TCATA	1		1				+: 1	
AEF-D	TDS	1		1					
DATI	TI	2					+: 2	+: 1	
M-TDS	TDS	2		2			+: 1	+: 2	-: 1; +: 1
PP	DA	1		1	1	+: 1			
PP	TI	2		1	1		-: 1		
TCATA	CATA	2		2	1	+: 1			
TCATA	DA	1		1					
TCATA	PP	1		1			+: 1		
TCATA	TDS	11		6	2	-: 1	+: 4	-: 1; +: 2	
T-CATA	M-TDS	1		1			+: 1	+: 1	
TCATA-F	TCATA	3		1			+: 1	+: 2	+: 1
TCATA-F	TDS	1		1				-	+: 1
TDS	CATA	1	1						
TDS	DA	11	2	1	8	+: 6	-: 1; +: 1	+: 1	
TDS	PP	2		1	1		+: 1		
TDS	TI	9	1	1	3			+: 1	
TDS-HD	TDS	1		1					
TDS-I	DA	10	1	6	6	+: 7	-: 1; +: 1	+: 1	
TDS-I	PP	1		1	1				
TDS-I	TI	2		2			+: 2		
TR	TCATA	1		1				+: 1	

¹Different/Similar/Complementary: number of articles concluding on differences, similarity and complementarity between Method 1 and Method 2, respectively; ²Temporal/Description/Discrimination/Agreement: number of articles concluding in superiority (+) or inferiority (-) of Method 1 compared to Method 2 regarding the capacity to highlight temporal patterns within product, to obtain complete descriptions of the products, to discriminate between different products, to reach high agreement between subjects, respectively.

This qualitative assessment is probably reductive because it relates to the main conclusions reported by the authors in the abstracts. Indeed, a lot of different criteria have been used in the articles to compare the temporal methods, and regarding the diversity of the statistical analyses performed, it was not possible to rely on specific outputs that would have enabled a more generic characterization. Moreover, only 2/3 of conclusions were supported by confirmatory analyses, and 1/3 by confirmatory data analyses considering sequentiality of perceptions (larger proportions than in Section 3.5). Thus, these results have to be considered as general trends rather than definitive conclusions, especially considering that some concepts related to temporal measurements might be beyond comparison (Meyners, 2020). However, this highlights the need for guidelines and methodology to compare results collected in different studies with different methods. In particular, it was observed that only 10% of the articles including method comparisons reported some results about reliability and/or validity. Most methodological conclusions were based on the capacity of methods to discriminate between products. However, statistical significance is not necessarily a synonym of meaningful results nor of validity (Stone et al., 2012). Such face validity is considered as the weakest form of validity, at risk for research bias (Bhandari, 2022) when people subsequently conclude based on low evidence. Discrimination should not be considered as the golden rule to validate methods, validity (even if it is difficult to establish) and reliability matter (see Moskowitz, 2008 for an old but still actual debate).

3.6 Dissemination of research involving temporal sensory evaluation methods

3.6.1 Quality appraisal

Figure 19.A shows the result of the quality appraisal of the reviewed articles (see 2.1.3). About 25% of the articles were not evaluated, mostly because there was no reference to a temporal evaluation method in the introduction. This result was quite surprising because temporal evaluation methods are not routinely used and even if they are used as a secondary measurement their choice is generally guided by specific hypotheses. For the other articles, the mean quality score was 5.3/7: 15% have the maximum score, 55% have a score greater than 4, a little less than 10% a score lower than 4. The least validated criteria (Figure 19.B) were Q6 (“Integrated interpretation and conclusion?”, 65% of articles not validating the criterion) and Q4 (“Appropriate data analysis?”, 45%).

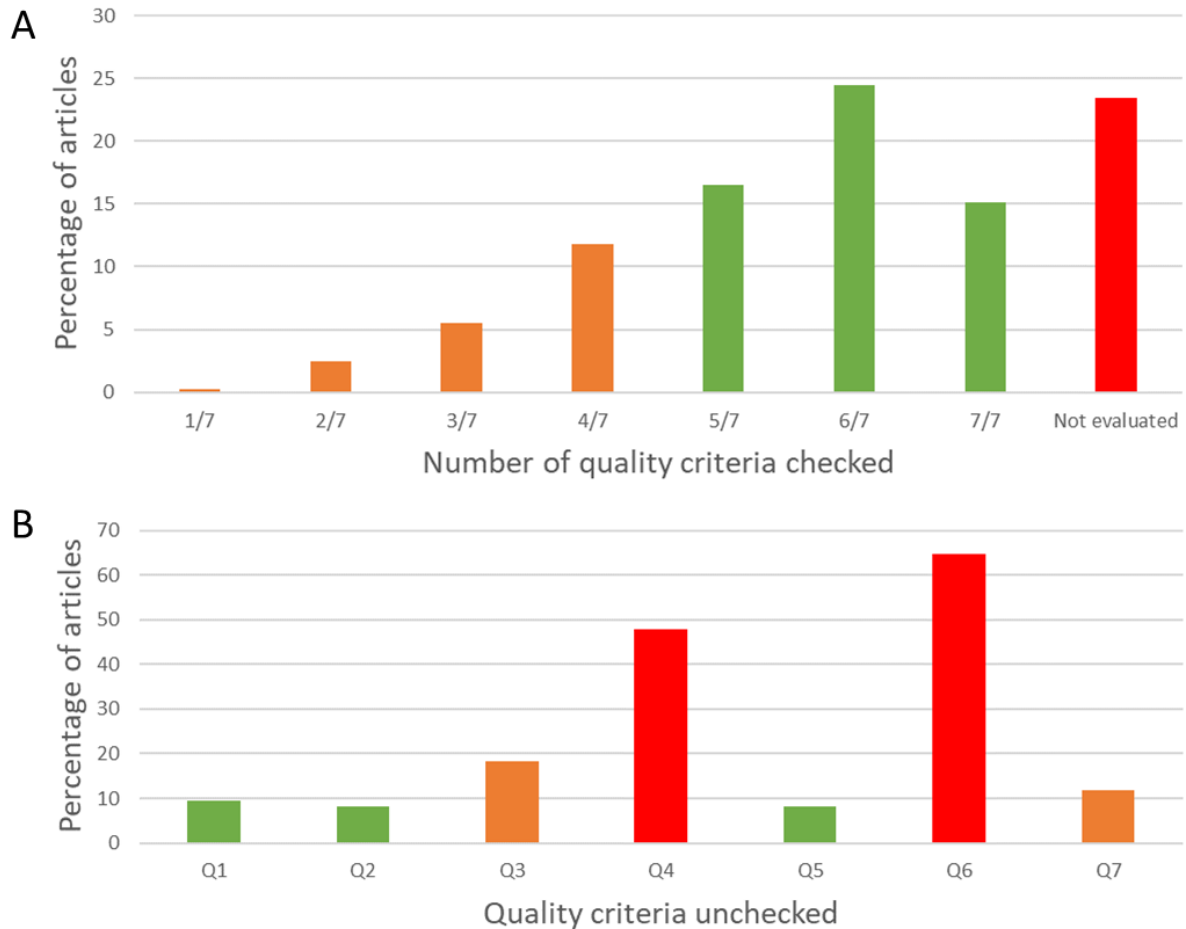


Figure 19. Quality appraisal of the included articles. A: Distribution of the quality scores of articles. B: Distribution of the percentage of articles having quality criteria unchecked, by criteria. Note: the authors were involved (directly or not) in 45 articles which were also evaluated.

For Q6, it can be noticed that no limitations were reported in 60% of the evaluated articles, and for Q4 the high percentage of non-validated articles is explained by conclusions exclusively based on subjective qualitative analyses in 40% of articles. Other reasons include not reporting product and/or attribute presentation orders (Q3, 10%) and not reporting contributions to the field (Q7, 10%). These results and those reported in previous sections demonstrate a lack of standards (or their application) for reporting and reviewing research related to multi-attribute temporal sensory evaluation methods and probably other sensory measurements as well. This issue is not specific to food science (Sizo et al., 2019), but questions the perception of the quality in research (see Akdag, 2019 for a discussion).

Some basic advice can be drawn from these results. When using temporal evaluation methods for research purposes, it is recommended to justify the choice of methods supported by adequate references with regards to the hypotheses and the objective of the research (these hypotheses/objectives should be clearly stated in the introduction, see Thomas & Hodges, 2010 for recommendations). If new methods or variants are introduced, extensive bibliographic research should be done to avoid “reinventing” an existing method. An appropriate use of both exploratory and confirmatory analysis should be made to draw more robust conclusions (see Fife & Rodgers, 2021 for a substantiated argument). Every research has limitations, and reporting them is a guarantee of quality and rigor in research, ensuring readers do not overemphasize or minimize findings (Ross & Bibler Zaidi, 2019). As noticed in Section 3.3.1, when sensory evaluation is a secondary measurement and the author does not have expertise in the field of sensory science or data analysis, appropriate co-authors should be associated to the research work. To help researchers to report their works and reviewers to evaluate the completion of submitted articles, we propose a checklist derived from the criteria used to evaluate the articles in this review (Visalli & Galmarini, 2022). This checklist (Visalli & Galmarini, 2023) is versioned and can be downloaded on a public repository from this URL: <https://doi.org/10.57745/JUJRTJ>.

3.6.2 Compliance to ethics in research and open science

About 40% were published in open access, with a clear increasing trend. More than 50% have been published in open access over the last two years, against less than 30% 10 years ago. This percentage is rather fair compared to other disciplines (Demeter et al., 2021).

Over the last two years, 90% of published articles included funding sources and declarations of interest, 70% authors' contributions, 60% subjects' informed consent, and 45% a review board approval (or exemption). However, less than 1% of research data are available on public repository and meet the principles of findability, accessibility, interoperability, and reusability (FAIR). This low percentage can be explained by a lack of familiarity with the FAIR principles (Brock, 2019) or by reluctance to share data on commercial products. As a second step towards FAIRification of data (Visalli et al., 2023), we propose a template for sharing data and meta-data related to sensory evaluation measurements. This template (Visalli & Galmarini,

2023) is versioned and can be downloaded on a public repository from this URL: <https://doi.org/10.57745/B35XCS>.

3.6.3 Reuse of works

Figure 20 shows the mean Field-Weighted Citation Impact (FWCI) by area of knowledge. The median FWCI computed over all the articles is 1.1, meaning that overall articles including results collected from temporal methods are 10% more cited than other articles in Food Science. The less cited articles are those dedicated to statistics. This can be explained either by an absence of need of new statistical techniques with regards to research objectives, by articles too complex to follow for scientists in sensory and food science but not statisticians, or by a lack of support in the most used software for the proposed data analyses. In any case, this result suggests that additional efforts should be made to make new statistical analyses more accessible to the sensory science community.

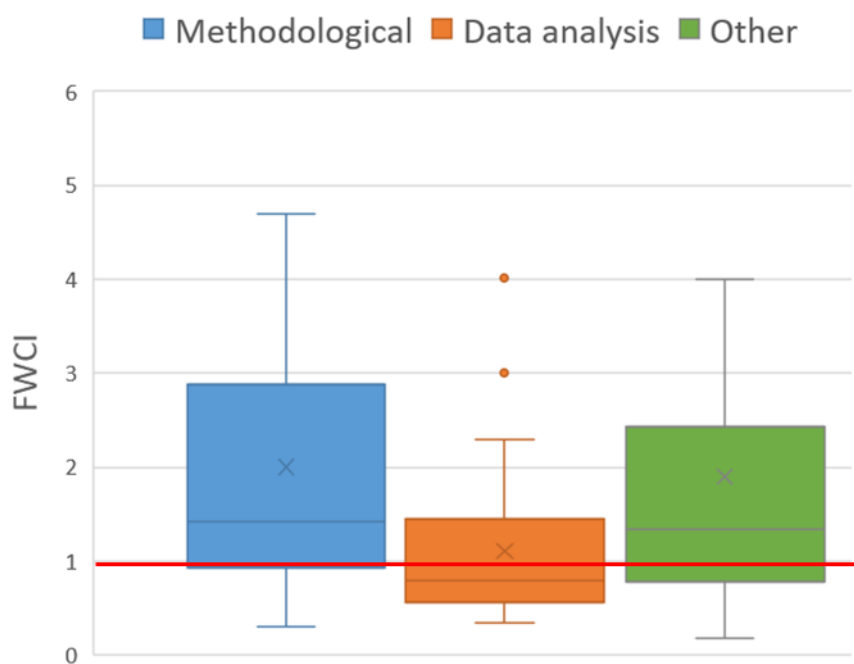


Figure 20. Field Weighted Citation Impact (FWCI) of articles by area of knowledge.

3.7 Limitations

This article is based only on published research works, which make the conclusions not necessarily representative of practices outside of the academic context.

Despite all the care and double-checking, compiling data requires some subjective decisions. Thus, the reported percentages were rounded to the nearest five and should be considered as general indications.

4. Conclusion

This review describes 30 years of research involving multi-attribute temporal methods by mapping the scientific literature in an exhaustive way (363 articles from 1991 to 2022). It presents how methods were developed, refined, disseminated, and informs about past and current trends in their implementations. The review enabled to identify some research gaps related to temporal sensory evaluation methods. The need for research on validity and reliability of the methods has been highlighted, as well as the need for recommendations about their implementation (choice of the method, type and training of the panel, number of subjects, use of replicates) and the analysis of temporal data (which analysis for which purpose).

Most temporal methods record perception closer to the moment of perception, and could also be useful tools to better understand physiological mechanisms. However, some results presented in this review suggest that knowledge was built on a fragile foundation due to a lack of guidelines in the way that studies involving temporal sensory evaluation methods have been implemented, their data analysed and their results reported. This could have resulted in a misuse of the methods, or in an overinterpretation of the results due to too much expectation about the validity and reliability of temporal data.

The objective of this review was not to point out bad practices but rather to suggest avenues for improvement that could help to increase the quality of the research. The sensory science community is invited to try the checklist and the template proposed with this article and to suggest improvements. We hope that FAIRification of data will in a near future make it possible to have a broader and more neutral body of information and opening up the prospect of meta-analyses that would allow more substantiated recommendations.

Supplementary material

Table 15. Checklist for reporting research involving (temporal) sensory evaluation methods.

Introduction

- 1. I have justified why I used a temporal descriptive approach and this method in particular.
- 2. I have cited the reference articles which present the method.
- 3. I have a research question (allowing readers to determine whether the object of study is related to products, topics, methods, etc.) and substantiated its relevance.
- 4. I have stated hypotheses, if possible.
- 5. I have reported specific objectives.

Materials and methods - Participants

- 6. I have justified the number of subjects that I chose to evaluate the products.
- 7. I have described the recruitment and selection modalities and the retribution made to participants.
- 8. I have described the demographics of the panel (age, gender, frequency of consumption, etc.)
- 9. I have explicitly described the settings of the experiment (country, lab/home, etc.)
- 10. I have explicitly described panel training in a reproducible manner (activities, number of hours, number of sessions).
- 11. I have explicitly reported if and how I have tested repeatability of the panel (for trained/semi-trained panels).
- 12. I have explicitly reported out the design (within or between subjects) when I have divided the participants into several groups.

Materials and methods - Products

- 13. I have reported relevant information about the products (commercial/model, composition, etc.) that could help to interpret sensory evaluation results.
- 14. I have reported serving conditions in a reproducible manner (portion size, temperature, type of light, blind/informed conditions, number of products per session, warm-up, etc.)
- 15. I have clearly reported the order of presentation of products (fixed, balanced, randomized, etc.)

Materials and methods - Attributes

- 16. I have reported the basis for choosing the sensory attributes (by the panel, by the experimenter, in the literature, etc.)
- 17. I have reported the attributes presented to participants in a way that sensory modalities are clear (e.g. aroma vs. flavour)

- 18. I have explicitly mentioned whether definitions and references were given to participants or not. If yes, I have reported these definitions and/or references.
- 19. I have reported the order of presentation of attributes (fixed, balanced, randomized, etc.)

Materials and methods - Data collection

- 20. I have reported all the necessary details to make tasting reproducible (free/imposed duration, start/stop buttons, breaks, rinsing, etc.)
- 21. I have reported the number of replicates and intakes, even if working with consumers.
- 22. I have reported the instructions given to the participants for the task.
- 23. I have reported the software used for collecting data.

Materials and methods - Data analysis

- 24. I have reported the alpha risk of confirmatory analyses in the data analysis section.
- 25. I have explicitly described the variables (citation rates, durations, etc.), data transformations (standardization, etc.) and models (fixed/random effect ANOVAs, etc.) used.
- 26. I have cited appropriate references for the statistical methods used.
- 27. I have reported the software used for analysing data.

Results

- 28. I have reported subject attrition (if applicable).
- 29. I have justified why I have chosen to remove or not to present certain data (if applicable).
- 30. I have presented interpretations consistent with Figures and Tables.

Discussion

- 31. I have provided a statement of support or non-support of my hypotheses.
- 32. I have compared my findings to the literature (whether they are congruent or not).
- 33. I have discussed the implications of my findings.
- 34. I have reported limitations in my experimental design and/or findings
- 35. I have discussed the generalizability of my findings.

Conclusion

- 36. I have explicitly answered my research question.
- 37. I have discussed on the contribution of the research to the topic.
- 38. I have identified paths for future research.

Meta-data

- 39. I have reported useful keywords that were not already present in the article title.

40. I have reported the contributions of each co-author (CRediT statement).

41. I have explicitly mentioned the funding of the research.

42. I have reported any potential conflict of interest.

43. I have reported appropriate ethic statements.

Details can be found in Visalli, M., & Galmarini, M. V. (2022). Multi-attribute temporal descriptive methods in sensory analysis applied in food science: Protocol for a scoping review. PLoS ONE, 17. <https://doi.org/10.1371/journal.pone.0270969>

Table 16. Template for sharing sensory data.

sheet name	column name	description	type of data	format
dataset	dataset_doi	Digital object identifier of the dataset	Text	DOI
dataset	dataset_publication_date	Publication date of the dataset	Date	yyyy-mm-dd
dataset	dataset_url	Direct URL to dataset	Text	URL
dataset	dataset_version	Version number of the dataset	Number	NA
dataset	dataset_language	Language of the dataset (English recommended)	Text	ISO 639-1, see https://www.iso.org/fr/standard/22109.html
dataset	dataset_title	Title of the dataset	Text	NA
dataset	dataset_contributors	Contributor(s) to the dataset	Text	CRediT format, see https://beta.elsevier.com/researcher/author/policies-and-guidelines/credit-author-statement?trial=true
dataset	dataset_institutions	Institution(s) of contributor(s)	Text	semicolon separated list
dataset	dataset_countries	Country(ies) of data collection	Text	ISO 639-1, semicolon separated list
dataset	dataset_data_collection_period	Begin and end date for data collection	Date	yyyy-mm-dd, semicolon separated list
dataset	dataset_description	Description of the dataset, including original purpose for data collection	Text	NA
dataset	dataset_keywords	Keywords associated with the dataset	Text	semicolon separated list
dataset	dataset_subject_area	Subject area	Text	semicolon separated list
dataset	dataset_license	License of the dataset	Text	see https://creativecommons.org/about/cclicenses/
dataset	dataset_funding	Origin of funding	Text	NA
dataset	dataset_contact_names	Contact name(s)	Text	semicolon separated list
dataset	dataset_contact_mails	Contact mail(s)	Text	semicolon separated list
dataset	dataset_contact_orcid	Contact ORCID	Text	ORCID, semicolon separated list
dataset	ethics_statement	Ethics statements	Text	NA
dataset	declaration_interest	Declaration of competing interest	Text	NA
dataset	protocol_doi	Digital object identifier of the protocol	Text	DOI

dataset	related_articles_doi	Digital object identifier(s) of the related article(s)	Text	DOI, semicolon separated list
methods	method_id	Unique identifier of the method of sensory evaluation in the dataset	Text	NA
methods	panel_type	Type of panel	Text	trained/semi-trained/expert/consumer/other (precise)
methods	recruitment_modalities	Recruitment modalities	Text	company staff/students/database of participants/social networks/flyers/other (precise), semicolon separated list
methods	selection_criteria	Selection criteria	Text	semicolon separated list
methods	compensation	Compensation	Text	none/financial/product/other (precise)
methods	method_name	Sensory evaluation method	Text	NA
methods	method_doi	Digital object identifier(s) of the reference article(s)	Text	DOI, semicolon separated list
methods	measured_concept	Measured concept	Text	intensity/applicability/dominance /other (precise)
methods	temporal_resolution	Temporal resolution	Text	global/periods/discrete time/continuous time/other (precise)
methods	measurement_moment	Moment of evaluation	Text	before tasting/during tasting/after tasting/other (precise)
methods	measurement_scale	Type of scale	Text	continuous scale/category-ratio scale/discrete scale/checkbox/button/open-ended question/other (precise)
methods	training_activities	Training activity(ies)	Text	semicolon separated list
methods	training_duration	Total duration of training (hours)	Number	NA
methods	training_sessions	Number of sessions of training	Number	NA
methods	evaluation_sessions	Number of sessions of evaluation	Number	NA
methods	evaluation_samples_per_session	Number of samples presented by session of evaluation	Number	NA
methods	evaluation_session_duration	Duration of a session of evaluation (hours)	Number	NA
methods	attribute_selection	Criteria for selection of attributes	Number	by the same panel/by another panel/by the experimenter/literature/other (precise), semicolon separated list
methods	product_presentation_order	Product experimental design	Text	balanced/randomized/fixed/NA/other (precise)

methods	attribute_experimental_order	Attribute experimental design	Text	balanced/randomized/fixed/NA/other (precise)
methods	data_collection_settings	Data collection settings	Text	lab/home/restaurant/other (precise)
methods	data_collection_context	Data collection context	Text	natural/real/immersive/virtual/other (precise)
methods	data_collection_instructions	Instructions given to subjects	Text	NA
methods	data_collection_software	Data collection software (version)	Text	NA
subjects	subject_id	Unique identifier of the subject in the dataset (anonymized)	Text	NA
subjects	age	Age of the subject (years)	Number	NA
subjects	gender	Gender of the subject	Text	man/woman/non-binary
subjects	...	Other subject related variables (complete this sheet)	-	-
products	product_id	Unique identifier of the product in the dataset (anonymized)	Text	NA
products	product_type	Type of product	Text	commercial/non-commercial/model
products	product_category	Category of product	Text	FoodEx2 category, see https://data.food.gov.uk/codes/foodtype/hierarchy/main/report/_A0B6Z
products	product_physical_state	Physical state of the product	Text	solid/semi-solid/liquid/jelly
products	product_cooking_method	Cooking method of the product	Text	FoodEx2 category, see http://data.food.gov.uk/codes/foodtype/hierarchy/facet/cookMeth
products	product_nutrition_facts	Nutritional composition (g/100g)	Text	energy:X1; fat:X2; saturates:X3; carbohydrate:X4; sugars:X5; protein:X6; salt:X7
products	product_ingredients	Ingredients (%)	Text	ingredient1:X1; ingredient2:X2; ingredient3:X3
products	portion_size	Portion size (g)	Number	NA
products	serving_temperature	Serving temperature (°C)	Number	NA
products	light	Type of light	Text	natural light/artificial white lighting/artificial red lighting/other (report)

products	information	Information given to subjects (brand, price, label, etc.)	Text	NA
subjects	...	Other product related variables (complete this sheet).	-	-
attributes	attribute_id	Unique identifier of the attribute in the dataset	Text	NA
attributes	attribute_label	Label presented to subjects	Text	NA
attributes	attribute_category	Category of attribute	Text	objective/subjective
attributes	attribute_modality	Sensory modality	Text	taste/aroma/texture/mouthfeel/trigemina l/NA/other (report) for objective attributes, hedonic/emotion/other (report) for subjective attributes
attributes	attribute_definition	Definition presented to subjects	Text	NA
attributes	attribute_reference	Reference presented to subjects	Text	NA
attributes	attribute_scale_limits	Limits of the scales (values)	Number	semicolon separated list
attributes	attribute_scale_anchors	Limits of the scales (anchors)	Text	semicolon separated list
data	method_id	Unique identifier of the method of sensory evaluation in the dataset (same as in methods sheet)	Text	NA
data	subject_id	Unique identifier of the subject in the dataset (same as in subjects sheet)	Text	NA
data	product_id	Unique identifier of the product in the dataset (same as in products sheet)	Text	NA
data	attribute_id	Unique identifier of the attribute in the dataset (same as in attributes sheet)	Text	NA
data	replicate	Replicate number	Number	NA
data	intake	Intake number	Number	NA
data	time	Time or period of evaluation	Text or Number	NA
data	value	Value given by the subject	Text or Number	NA

I.4. The challenge of measuring temporal perception of consumers

I.4.1. Temporal methods, from analytical to rapid temporal measurements

Section I.1 has shown the importance of studying sensory perception (including preferences and sensory characterization of food products) to support the development of more sustainable diets. Indeed, since perception cannot be measured by instruments, the use of sensory evaluation methods is central to answering most research questions related to food perception (Forde, 2016).

Section I.2 has illustrated the evolution of consumer sensory science, shaped by the needs of industry, but urged by the need to increase the ecological validity of sensory measurements. Thus, even if some authors (Stone, 2018) still consider that carrying out analytical tests with untrained assessors threatens the validity of the results, the assumption that consumers are able to evaluate the sensory characteristics of products using rapid methods has become increasingly accepted (Ares & Varela, 2017).

Section I.3 has shown that over the years, many temporal sensory evaluation methods were developed order to obtain the “movie” of perception. These methods attempt to dynamically measure perception during the tasting of food products to get as close as possible to the moment when the sensations are perceived. The first methods were based on the same premise as static descriptive methods: quantifying perceived attributes over time using intensity scales. As also happened with static descriptive methods, temporal methods evolved towards qualitative measurements, which gave rise to rapid temporal descriptive methods that can be applied with consumers, the most widely used being TDS and TCATA.

Rapid temporal methods are flexible and can be used early in the lifecycle of product development or research projects (Delarue et al., 2015). They can be used with “large” panels, outside the laboratory, in ecological settings (Dinnella et al., 2022), making the data collected more representative of the perception of end users (Meiselman et al., 2022). They can be employed to identify differences in perceptions due to physiological or individual characteristics, e.g. saliva (Goza et al., 2022), mastication (Jeltema et al., 2020), expertise (Weerawarna et al., 2021), age, gender or ethnicity (Aguayo-Mendoza et al., 2020), etc. Temporal sensory descriptions and hedonic scores can be collected from the same subjects, over single or successive intakes to get more representative insights on consumer experiences

(Ares et al., 2017; Corrêa Simioni et al., 2018; Greis et al., 2020; Paglarini et al., 2020; Thomas et al., 2018; Thomas et al., 2015; van Bommel et al., 2019; Wakihira et al., 2020). Thus, rapid temporal methods could help to understand product quality perceptions and consumer preferences more precisely. It is therefore possible that they help to better explain consumer behaviour (Castura, 2018).

However, the conclusions of the scoping review (Section I.3.3) suggested that knowledge related to temporal measurements was built on fragile foundations. Indeed a lack of guidelines and/or rigor was noted in the implementation of studies involving temporal sensory evaluation methods. Furthermore, even though TDS and TCATA have been considered rapid methods for years (Delarue et al., 2015; Reyes et al., 2017), they are still mainly used with trained or semi-trained panels (Figure 21, source: data from the scoping review).

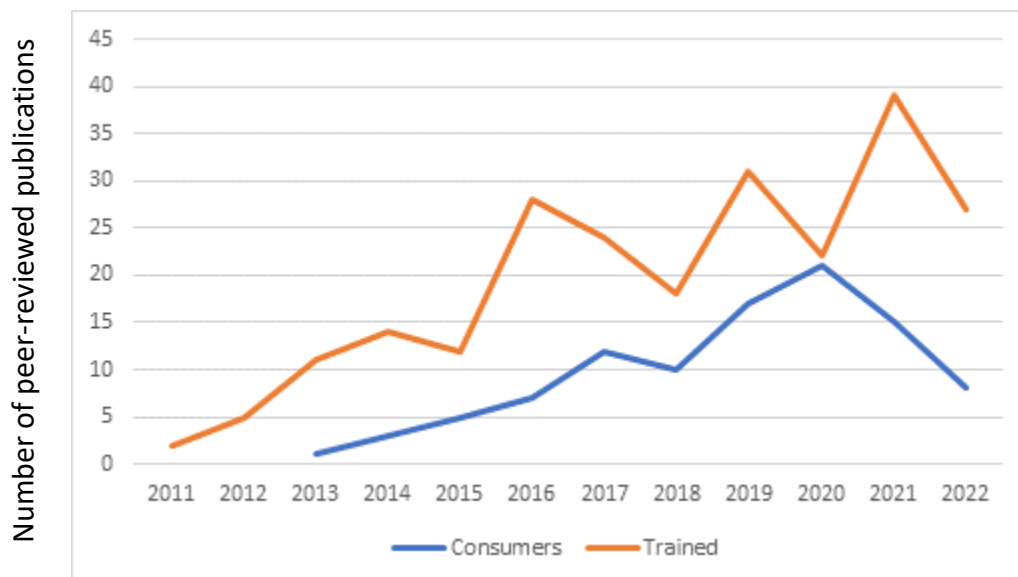


Figure 21. Evolution of the use of multi-attribute temporal sensory evaluation methods by trained and consumer panels over years.

This can be explained by the absence of clear recommendation on the application of the temporal methods (Section I.3.3, part 3.6.1) with consumers (choice of attributes, instructions for the task, number of subjects, conditions of tasting, etc.). Indeed, TDS and TCATA were first developed and used with trained panels, validated after comparison to other analytical methods, and only then transposed to consumer panels (Section I.3.3, part 3.3.3). It can be assumed that this change was probably guided more by practical aspects than by scientific considerations.

Providing data that meaningfully reflects a complex real-world experience using a task that consumers are able to perform successfully is a challenge (Castura, 2018). New variants of TDS and TCATA and new qualitative methods are continually being proposed (Section I.3.3, part 3.2.2), suggesting that some needs are still not met by existing temporal methods. However, the development of temporal methods was mainly supported by a few teams (Section I.3.3, part 3.3.6). In a context of academic competition (and sometimes software promotion), a lack of neutrality could have led to pushing the use of temporal methods “out of bounds”. This could have resulted in a misuse of the methods or an overinterpretation of the data obtained. Indeed, it is unclear to what extent the details of implementation may impact the results. It should be remembered that rapid methods compromise between costs, ecological validity and relevance of data collection (Delarue et al., 2015), and that there might be too many expectations about what can be obtained from temporal measurements collected from consumers.

Sensory data are not only complex to collect but also to analyse. Considering the temporal dimension adds new information to sensory data, but also a new source of uncontrolled variability that can cause additional noise in data. The most used criterion to assess the relevance of the results was the ability of the temporal method to discriminate between the products (Section I.3.3, part 3.5). The nature of the observed differences was rarely examined, and the reliability of temporal measurements was almost never considered (Section I.3.3, part 3.3.3). However, to draw robust conclusions from temporal data, it is important to document the performances of temporal methods beyond face validity.

Many data analysis tools were used to get insights into temporal data (Section I.3.3, part 3.4.2). However, most of the conclusions reported in the literature relied on interpretations based on exploratory data analyses (Section I.3, part 3.5). The most frequently used analyses in the literature were the TDS curves (Section I.3, part 3.4.2). Although this representation is not intended for this purpose, it has often been used to compare products.

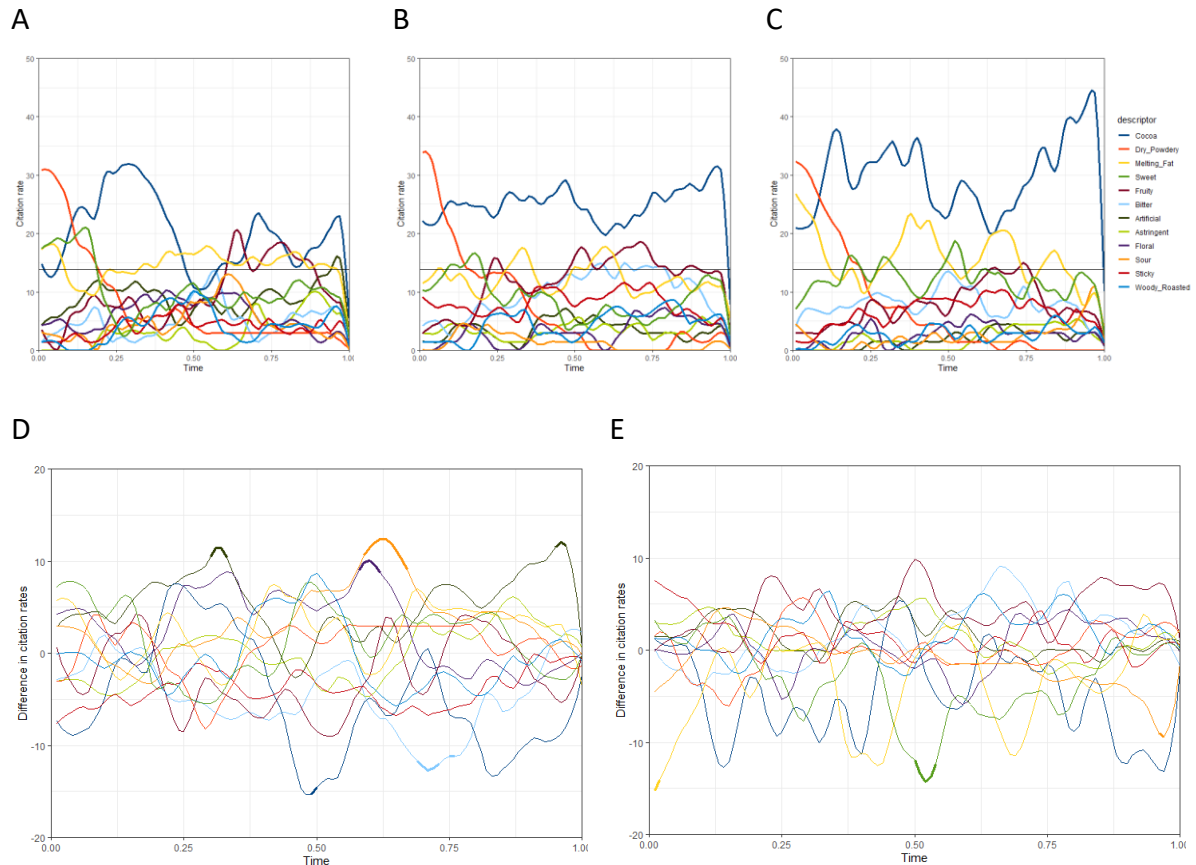


Figure 22. Standardized TDS curves and difference curves of CH3, CH4 and CH4_rep. A-C: Standardized TDS curves of CH3, CH4 and CH4_rep (replication of CH4). D-E: Standardized difference curves CH3-CH4 and CH4-CH4_rep (bold: significant differences, $\alpha=0.05$). Source: Visalli, Cordelle, et al., 2022.

To illustrate the risk of such a comparison, a survey was administrated to members of the European Sensory Network, asking them to compare the chocolates CH3 to CH4 and CH4 to CH4_rep (Figure 22.A vs Figure 22.B and Figure 22.B vs Figure 22.C). They had three possible answers: "I would say that the two samples are the same", "I would say that the two samples are different", "I can't tell".

Fifty-seven people participated in the survey (all reported experience with TDS). Only 13% answered they could not tell because they did not have statistical criterion (such as significant differences in citation rates with TDS difference curves) to conclude on the differences. CH3/CH4 and CH4/CH4_rep were reported as different by 69% and 67% of participants, respectively (Figure 23.A).

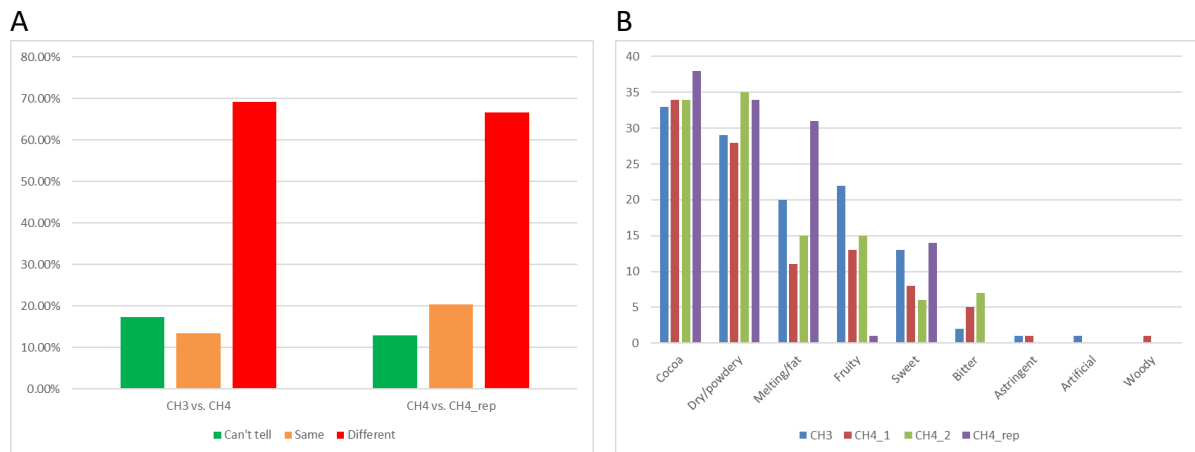


Figure 23. Results of the ESN survey. A: Percentage of answers for the comparison of samples. B: Descriptors used to describe the samples.

The TDS differences curves (not presented to ESN members) show the significant differences in citation rates between CH3 and CH4 (Figure 22.D) and CH4 and CH4_rep (Figure 22.E). Except for *Sour* (an attribute that never reached significance in CH3 and CH4), no difference last for more than 5% of the total duration of the tasting. The additional noise brought by the temporal dimension (e.g. the time offsets in subjects' perception of a same attribute) are not considered in this analysis. It is therefore possible that in fact there are no differences in the temporal perception of the three samples, except those due to measurement errors.

Participants also had to describe each sample using free comment, CH4 being described twice (CH4_1 and CH4_2). Figure 23.B shows there was no large agreement on the reported attributes, with *Fruity*, *Sweet* and *Bitter* being interpreted as dominant by half of the participants compared to *Cocoa*, *Dry/Powdery* and *Melting/Fat*. Figure 23.B also shows that the panel was not fully repeatable in the descriptions of CH4_1 and CH4_2.

Finally, in the free comments, seven participants interpreted the citation rates as intensities, and six concluded that there was no temporality in the TDS signal (including one participant who thought the curves were randomly drawn).

These preliminary results demonstrate how subjective conclusions about temporal sensory evaluation can be. More generally, although TDS and TCATA data are acquired with fine temporal resolution (in continuous time), the capacity of these methods to detect small temporal differences (within and between products) based on temporal aspects is not precisely documented. Thus, it is not guaranteed that temporal measurements systematically

have an added value compared to static measurements (Velázquez et al., 2020b; Vidal et al., 2019).

Table 17 summarizes the strengths, weaknesses, opportunities and threats related to rapid temporal sensory evaluation methods. The topics in italic will be addressed in this manuscript.

Table 17. SWOT analysis of rapid temporal sensory evaluation methods

Strength	Weakness
<ul style="list-style-type: none"> ⊕ Cost efficiency ⊕ Easy implementation ⊕ Flexibility ⊕ <i>Additional information compared to static measurements</i> 	<ul style="list-style-type: none"> ⊖ <i>Many methods available but no clear guidelines for implementing data collection</i> ⊖ <i>Concurrent data collection potentially demanding for consumers</i> ⊖ <i>Temporal data complex to analyse, with many available tools but no clear guidelines for choosing the most appropriate one and interpreting the results</i> ⊖ <i>Temporal resolution not documented</i> ⊖ <i>Validity difficult to establish</i>
Opportunities	Threats
<ul style="list-style-type: none"> ⊕ Larger panels ⊕ <i>Measurements in ecological settings</i> ⊕ <i>Study of individual differences in temporal perception of products</i> ⊕ Investigation of temporal determinants of preferences 	<ul style="list-style-type: none"> ⊖ <i>New biases and additional noise in data introduced by the temporal dimension but reliability not assessed</i> ⊖ <i>Too many expectations about temporal resolution of data</i> ⊖ <i>Misuse of methods</i> ⊖ <i>Subjective interpretation of results</i>

I.4.2. Aims and objectives of this work

Whether for researchers interested in better understanding temporal perception, or food manufacturers interested in the description of the temporal evolution of the perception of their products, the weaknesses and threats identified in Table 17 can pose a problem. Indeed, it can be difficult to choose the rapid method of temporal data collection the most appropriate, to implement it effectively with consumers, then analyse the temporal data and interpret the results in a way to draw robust conclusions.

The aim of this research was to critically question the information provided by consumers using temporal rapid sensory evaluation methods to evaluate food products. The general research question was "how appropriate are rapid sensory evaluation methods for measuring the temporal perception of consumers?".

The first objective was to make an exhaustive inventory of research practices using temporal sensory evaluation methods to evaluate food products (with all types of panels). To this end, a systematic scoping review was carried out (1.3.3), with particular emphasis given to methodological aspects.

The second objective was to investigate the granularity of the temporal data collected from consumers. To do this, different temporal methods were tested, by varying the qualitative concept measured (dominance or applicability), the attributes (lists of attributes or free comment), the moment of evaluation (concurrent measurements in continuous times or discrete times; retrospective measurements in periods). The conclusions obtained with the concurrent temporal methods (data collected in continuous time during the tasting, high temporal resolution expected for capturing within and between product differences) were compared to those obtained with retrospective temporal methods (data collected after tasting as with the static methods, inherently low temporal resolution for capturing within and between product differences). To this end, a new retrospective temporal sensory evaluation method was introduced and implemented with different variants: Attack-Evolution-Finish (AEF). AEF-Dominance (AEF-D) was compared to TDS, and AEF-Applicability (AEF-A) was compared to discrete time TCATA. Free Comment AEF-A (FC-AEF-A) was used with different panels (expert vs. consumers) and the results obtained with each type of panel compared.

The third objective was to document the performances of TDS, TCATA and AEF when used with consumers. The conclusions obtained with the three methods were compared on the basis of data collected to describe controlled temporal stimuli (delivered using a gustometer) and real food products. Confirmatory data analyses were used to assess the validity, the reliability, the temporal resolution and the discrimination ability of the methods.

The ultimate objective was to formulate recommendations on the choice of the most appropriate temporal method (depending on the objective of the sensory evaluation), on the implementation of the method with consumers, on the analysis of temporal data and on their interpretation.

I.4.3. Plan of the manuscript

Chapter I introduces this research. The first section presents an overview of the mechanisms of sensory perception and the importance of conducting research on sensory evaluation methods. The second section summarizes the history of sensory-consumer science and the development of sensory evaluation methods. The third section details through a published protocol the methodology used to perform a systematic scoping review on temporal sensory evaluation methods, then presents an article including the results of the systematic scoping review. The fourth section formulates the problematic of this research.

Chapter II details the experimental protocols implemented to meet the objectives of this research work and describes the dataset collected. Each section corresponds to a published data paper referenced by one (or more) research articles presented in Chapter III.

Chapter III is organized into two sections focusing on the second and third objectives defined in Section I.4.2. The first section includes three research articles introducing several AEF variants as new rapid temporal sensory evaluation methods based on retrospective evaluations. The second section presents three research articles related to the performances of rapid temporal sensory evaluation methods (validity, reliability, temporal resolution and discrimination ability). A synthesis of the context is presented at the beginning of each section, and the main findings are summarized at the end of each section.

Chapter IV discusses the results presented in Chapter III. The first section reports the limits of this work. The second section summarizes the key findings and their implications. The third section provides recommendations for the use of temporal sensory evaluation methods with consumers.

Chapter V concludes this work and outlines perspectives for future work. Finally, our opinion established at the end of this work on temporal methods used with consumers is presented in a position paper.

Chapter II.

Methods

During this research work, data were collected using different temporal evaluation methods (TDS, TCATA, AEF-D, AEF-A, FC-AEF-D, FC-AEF-A, AEF-RATA), with different types of panels (consumers, experts), in different countries (France, Spain, Japan), in different settings (laboratory, home). Several categories of commercial products and model solutions varying in physical states (solid, semi-solid, liquid) and complexity were evaluated. The five datasets collected during this research work are summarized in Table 18.

Table 18. Summary of datasets collected.

Data paper	Type of product	Samples	Attributes	Settings	Number of consumers					
					TDS	TCATA	AEF-A	AEF-D	FC-AEF-D	FC-AEF-A
Visalli, Mahieu, & Schlich, 2023	Dark chocolates	5 (0)	12	Lab, France	129 ¹	-	-	129 ¹	-	-
Visalli, Wakihira, et al., 2022	Lemon-flavoured carbonated alcoholic drinks	4 (0)	8	Home, Japan	-	93 ^{2,3}	97 ³	-	-	-
Visalli, Dubois, Schlich, Ric, et al., 2023	Bordeaux and Rioja red wines	4 (0)	NA	Lab, France	-	-	-	-	-	47 ⁴
				Home, France	-	-	-	-	-	106
				Home, Spain	-	-	-	-	-	98
Béno et al., 2023	Model solutions delivered by gustometer	4 (2)	8	Lab, France	50	50	49 ⁵	-	-	-
Visalli, Cordelle, et al., 2022	Guacamoles	4 (1)	12	Lab, France	70	73	74	75	72	72
	Flavoured ice teas	4 (1)	8	Lab, France	70	73	74	75	72	72
	Crisps	4 (1)	9	Home, France	70	72	74	75	66	70
	Dark chocolates	4 (1)	12	Home, France	68	72	71	75	68	70

Samples: number of samples (number of replicated samples); Attributes: number of attributes; ¹Within-subject design; ²Discrete-time TCATA; ³Multi-sip; ⁴Expert panel; ⁵AEF-RATA.

Notes: The protocols will be reminded in the research articles, making the reading of this Chapter optional. The data template proposed with the scoping review was published after the data papers were written, so they did not use it.

II.1. A dataset on concurrent and immediate retrospective measures of sensory perception and preferences of dark chocolates

This article was *published in Data in brief*:
<https://www.doi.org/10.1016/j.dib.2023.109271>

B. Mahieu: Investigation, Writing - review & editing; P. Schlich: Writing - review & editing, Supervision; M. Visalli: Data curation, Writing - original draft, Visualization.

This article describes data related to the research paper entitled “Concurrent vs. retrospective temporal data collection: Attack-evolution-finish as a simplification of Temporal Dominance of Sensations?” (Visalli et al., 2020). Temporal sensory perception data of five dark chocolates that vary in cocoa content were collected from 129 consumers who evaluated the samples in two sessions, using a different sensory evaluation method in each session. A within-subject design was set-up to compare the two data collection methods: consumers in Panel 1 (36 men and 32 women aged 19 to 63 years old) started with the Temporal Dominance of Sensations (TDS) method, and consumers in Panel 2 (35 men and 26 women aged 19 to 61 years old) started with the Attack-Evolution-Finish dominance (AEF-D) method. For each chocolate, consumers had to report the sensations they perceived either concurrently (TDS) or retrospectively (AEF-D) to the tasting. After the descriptive task, consumers were asked to rate their liking for chocolates on a 9-point discrete scale. Finally, consumers had to answer questions related to the difficulty of the descriptive task. The dataset includes information on consumers’ gender, age and frequency of consumption of dark chocolates. The dataset can be reused by sensometricians to compare methods or develop new statistical models for data analysis. It can also be reused to compare at the individual level declarative sensory measures collected either concurrently or retrospectively to tasting. Thus, the impact of cognition (due to memorization, stress or complexity of measurements) on sensory description and liking can be investigated. More specifically, this dataset can be help understand how the dynamics of perception of texture, mouthfeel and flavour attributes are integrated when using static measures.

1. Specifications table

Table 19. Specifications table.

Subject	Food science
Specific subject area	Temporal sensory evaluation
Type of data	Table Figure
How the data were acquired	<p>Two consumer panels (129 consumers in total) evaluated the samples in the sensory booths of the ChemoSens platform, using the TimeSens© software, version 2.0 (INRAE, 2022).</p> <p>Two temporal sensory evaluation methods were used by consumers to describe their temporal perception: Temporal Dominance of Sensation (TDS, Pineau et al., 2009) and Attack-Evolution-Finish Dominance (AEF-D, (Visalli et al., 2020). Liking scores were rated using 11-point discrete scales (between 0 and 10). Items related to the perception of the tasks were evaluated using 5-point Likert scales (Likert, 1932).</p>
Data format	Raw
Description of data collection	<p>In two sessions spaced 48 hours apart, five samples of dark chocolates were evaluated by all consumers separated in two panels (same samples in each session, within-subject design). The consumers had first to qualitatively describe their temporal perception of each sample by selecting in a predefined list of attributes (<i>Astringent, Bitter, Cocoa, Dry, Fat, Floral, Fruity, Melting, Sour, Sticky, Sweet, Woody</i>) those they perceived as dominant. Half started with TDS in session 1, half by AEF-D, and the order was reversed in session 2. When using TDS, the evaluation task was concurrent to the tasting, the times of citations of each attribute and the duration of the tasting were recorded. When using AEF-D, the evaluation task was retrospective to the tasting and only the duration of tasting was recorded. After the descriptive task, without re-tasting the samples, the consumers had to rate their liking for the sample, and the time it took was recorded. After having evaluated the five samples, the consumers filled a questionnaire about the difficulty of the descriptive task. At the end of session 2, they reported their opinion on the relative difficulty of each descriptive task.</p>
Data source location	<p>Institution: INRAE City/Town/Region: Dijon Country: France</p>
Data accessibility	<p>Repository name: Mendeley data Data identification number: 10.17632/9c9g3rh8rd.1 (Visalli, Mahieu, & Schlich, 2023) Direct URL to data: https://data.mendeley.com/datasets/9c9g3rh8rd/1</p>
Related research article	<p>M. Visalli, B. Mahieu, A. Thomas, P. Schlich, Concurrent vs. retrospective temporal data collection: Attack-evolution-finish as a simplification of Temporal Dominance of Sensations? <i>Food Quality and Preference</i>, 85, (2020), 103956, ISSN 0950-3293. https://doi.org/10.1016/j.foodqual.2020.103956.</p>

2. Value of the data

- These data are useful because they allow a comparison at the panel and individual levels of concurrent and retrospective descriptive and hedonic measures collected with sensory evaluation methods.
- The food science and sensometrics community can benefit from these data to gain insight on sensory perception of dark chocolates or test new statistical models, respectively.
- Researchers in cognitive science can also benefit from these data for testing if Kahneman's theory (O'Brien, 2012) (fast vs. slow thinking) applies on sensory perception.
- These data can be reused to investigate individual differences in perception due to cognitive biases (complexity of the instructions, memorization, stress) and related to retrospective and concurrent measures, and test if all subjects are affected by these biases equally or differently (depending on individual characteristics).
- They can also be reused for studying differences between what is perceived (concurrent measures) and what is recalled and integrated (retrospective measures) according to sensory modalities (differences between texture, mouthfeel, basic taste and aroma attributes), and how these differences affect hedonic perception.

3. Objective

As sensory perception is a dynamic process, sensory evaluation methods such as TDS have been developed to collect data simultaneously with tasting. However, if the dynamic evaluation task is natural, it remains demanding for consumers as it attaches great importance to the moment at which the sensations are cited. To ensure valid measures and limit heterogeneity in consumers' behaviours with TDS, a familiarization step with the method may be required, which is not always possible outside of laboratory settings. The published article introduced AEF-D as an alternative, simplified method for temporal measurements. It focused on methods comparison, practical aspects of data collection and statistical analysis related to discrimination of products at the panel level. However, the data collected can also be used for addressing questions related to the subjects and the descriptors. More precisely, it is possible to study if individual differences can be observed about retrospective integration of dynamic perception, and if this integration varies according to the sensory modalities and affects the

liking. These questions not covered in the published research article are of primary importance because they are linked to the ecological validity of the measures.

4. Data description

The dataset is provided as an Excel file (data.xlsx) including seven sheets.

Subjects provides information about the consumers.

- “Panel”: name of the panel to which the participant has been assigned.
- “Subject”: unique code of the participant.
- “Gender”: gender of the participant (M: male or F: female).
- “Age”: age of the participant.

Products provides information about the dark chocolates evaluated by the participants.

- “Product”: code of the product. The 2-digit number corresponds to the percentage of cocoa.

TDS and AEF-D provide data collected with the corresponding temporal evaluation methods.

- “Subject”, “Product”: see above.
- “Attribute”: dominant attribute clicked by the participant (*Astringent, Bitter, Cocoa, Dry, Fat, Floral, Fruity, Melting, Sour, Sticky, Sweet, Woody* + *START* and *STOP* in TDS).
- “Time”: time (in seconds) of click on the attribute (TDS)
- “Period”: period (A for Attack, E for Evolution, F for finish) during which the attribute was retrospectively declared applicable (AEF-D).

AEF-D tasting duration provide data about the duration of the perception before the AEF-D evaluation.

- “Subject”, “Product”: see above.
- “Duration”: duration of perception (in seconds).

Liking provide data about the liking scores given after the descriptive task.

- “Subject”, “Product”: see above.
- “AfterTDS-Liking” and “AfterAEF-Liking”: value rated on a 11-point discrete scale (between 0 and 10) after the TDS and AEF task, respectively.
- “AfterTDS-TimeToReportLiking” and “AfterAEF-TimeToReportLiking”: time (in seconds) required to report the liking score.

Questionnaire provide answers of the participants to task-related questions. Items measured on a Likert scale used the following labels: “strongly agree” (5), “agree” (4), “neither agree nor disagree” (3), “disagree” (2), and “strongly disagree” (1).

- “Subject”: see above.
- “ExperienceWithTDS”: previous experience of the participant with the TDS method (Yes, No or Don’t know)
- “AfterTDS-Q1”: Answer to the question “the oral explanations were useful” (Likert scale, TDS only).

- “AfterTDS-Q2”: Answer to the question “The explanations displayed on the screen about how to evaluate chocolates were useful” (Likert scale, TDS only).
- “AfterTDS-Q3” and “AfterAEF-Q3”: Answers to the question “I understood how to evaluate the chocolates” (Likert scale).
- “AfterTDS-Q4” and “AfterAEF-Q4”: Answers to the question “The list of sensations was exhaustive” (Likert scale).
- “AfterTDS-Q5” and “AfterAEF-Q5”: Answers to the question “The sensations were sufficiently explanatory” (Likert scale).
- “AfterTDS-Q7A”: Answer to the question “It was easy to identify the sensations that caught my attention during the tasting” (Likert scale, TDS only).
- “AfterTDS-Q8”: Answer to the question “It was easy to quickly click on a sensation when it caught my attention” (Likert scale, TDS only).
- “AfterTDS-Q9”: Answer to the question “It was easy to identify when to click STOP” (Likert scale, TDS only).
- “AfterTDS-Q10” and “AfterAEF-Q10”: Answers to the question “The order in which I listed the sensations was important” (Likert scale).
- “AfterTDS-Q11” and “AfterAEF-Q11”: Answers to the questions “I could list the same feeling several times” (Likert scale).
- “AfterTDS-Q12” and “AfterAEF-Q12”: Answers to the questions “The questionnaire interface was easy to use” (Likert scale).
- “AfterTDS-Q13” and “AfterAEF-Q13”: Answers to the questions “The task that was asked of me was easy” (Likert scale).
- “AfterAEF-Q1”: Answer to the question “I wish I had oral explanations” (Likert scale, AEF-D only).
- “AfterAEF-Q2”: Answer to the question “The explanations displayed on the screen about how to evaluate chocolates were sufficient” (Likert scale, AEF-D only).
- “AfterAEF-Q6”: Answer to the question “I wished I could select more than 3 sensations” (Likert scale, AEF-D only).
- “AfterAEF-Q7B”: Answer to the question “It was easy to identify the sensations perceived at the beginning of the tasting” (Likert scale, AEF-D only).
- “AfterAEF-Q7C”: Answer to the question “It was easy to identify the sensations perceived at the middle of the tasting” (Likert scale, AEF-D only).
- “AfterAEF-Q7D”: Answer to the question “It was easy to identify the sensations perceived at the end of the tasting” (Likert scale, AEF-D only).
- “RelativeDifficulty”: Answer to the question “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?’”. Coded between -2 (TDS much easier) and 2 (AEF much easier).
- “Comment”: open-ended question about the overall opinion of the participant about the experiment (in French).

Figure 24 describes the procedure of data collection.

5. Experimental design, materials and methods

5.1. Participants

Participants were preselected from a population registered in the ChemoSens Platform's PanelSens database (declared to the relevant authority, Commission Nationale Informatique et Libertés – CNIL, authorization number 1148039). The inclusion conditions for participating in this study were as follows: being between 18 and 65 years old; not suffering from food or non-food allergies; not being pregnant or breastfeeding and not following a restrictive diet incompatible with the consumption of sugar; being a regular consumer of dark chocolate (at least once every two weeks). The purpose of the study was explained via an information sheet sent by email. The participants have to accept the conditions and fill out a written informed consent form before to be included.

One hundred and forty consumers were selected and randomly assigned to one of two panels, with a constraint of balance in gender and age between panels. Due to attrition, a total of 129 consumers (71 men and 58 women, between 19 and 63 years old) finally participated in this study (68 in panel 1, 61 in panel 2). They were compensated for their participation in the study (vouchers worth 20€).

5.2. Samples

Five dark chocolates (provided by Barry Callebaut, cocoa product manufacturer), varying in cocoa content (54%, 65%, 68%, 70% and 73%) and in origin of cocoa bean, were selected for this study. The samples were given to the consumers in transparent plastic containers containing four callets of chocolates of 0.5 g each and labelled with 3-digit codes.

5.3. Data collection

The consumers participated in sessions of approximately one hour in the sensory lab of ChemoSens at the Centre for Taste and Feeding Behaviour, Dijon, France.

In session 1, consumers in panel 1 firsts attended a collective briefing (groups of 16 persons) aiming at explaining the procedure and demonstrating how to report their perception using the TDS method. The concept of dominance was presented as “the sensation that catches the attention at any time”, and the panel leader shown how to interact with a TDS measurement screen while simulating a tasting. To ensure their understanding of the task, the participants were invited to ask any question.

Then, they were installed in individual booths equipped with computers running TimeSens© V2 software, the data acquisition program. The instructions were reminded on the screen: “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”.

The sequence presented in green on Figure 1 was repeated for each chocolate, the samples being presented under white light, at ambient temperature, in a sequential monadic order according to a Williams Latin square. First, the sensory attribute list (chosen based on previous experiments) was displayed on the screen. This list included the following attributes: *astringent, bitter, cocoa, dry, fat, floral, fruity, melting, sour, sticky, sweet, woody*. The consumers were reminded to familiarize themselves with the location of the attributes on the screen before clicking on the button START (which triggered the chronometer) while at the same time consuming the four callets in a single intake. Once they have clicked on the button STOP (which stopped the chronometer), the consumers had to report their liking on a 11-point discrete scale labelled from 0 (“I did not like at all”) to 10 (“I liked it very much”). Then, a 30-second pause was imposed, during which consumers were asked to rinse their mouths with mineral water.

After having evaluated the five samples, the consumers had to fill a questionnaire related to the TDS evaluation task. The questionnaire included 12 items (see Section “data description>questionnaire”) to evaluate using a 5-point Likert (labels: “strongly agree”, “agree”, “neither agree nor disagree”, “disagree”, and “strongly disagree”). That ended the session, and consumers were invited to come back 48 hours later at the same time for session 2.

In session 2, consumers in panel 1 were directly installed in individual booths. The instructions for the AEF-D task were presented on the first screen of the software: “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”. The second screen presented examples of description: “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”.

The sequence presented in green on Figure 1 was repeated for each chocolate (with the same experimental design as TDS). The consumers were instructed to consume in a single intake the four callets while clicking on the START button at the same time. Then, they were invited to focus and memorize the sensations they perceived. When they did not perceive any sensation, they had to click on the STOP button (which was enabled after 10 s). It was only then that the AEF-D measurement screen allowing them to report the perceived sensations appeared. The instructions were: “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...”. After the descriptive task, the consumers had to report their liking (same as with TDS) and a 30-second pause was imposed to rinse their mouths.

After having evaluated the five samples, the consumers had to fill a questionnaire related to the AEF-D evaluation task. The questionnaire included 13 items (see Section “data description>questionnaire”). Finally, a last question invited the consumers to compare the relative difficulty of TDS and AEF-D: “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?’”. They also had the possibility to report anything related to the task using an open-ended question.

For consumers in panel 2, the order of tasks was reversed: AEF-D in session 1, and TDS in session 2.

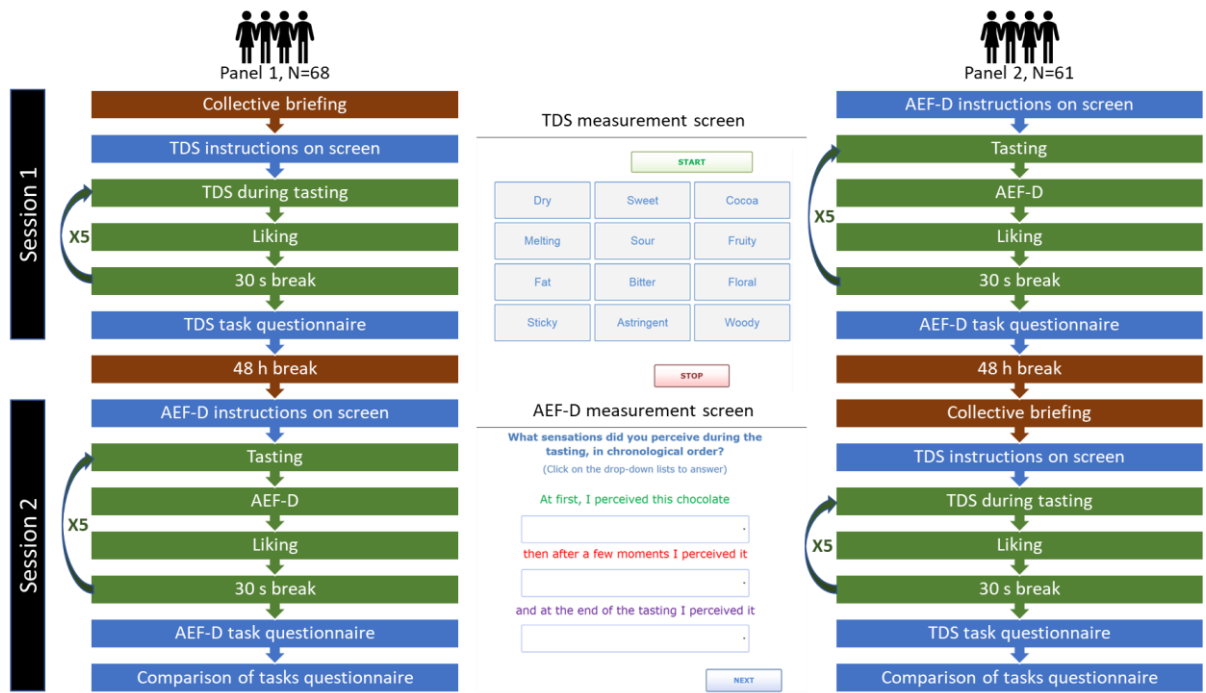


Figure 24. Procedure of data collection.

II.2. A dataset on concurrent and retrospective sensory perception and preferences for lemon-flavoured carbonated alcoholic drinks

This article was *published in Data in brief*:

<https://www.doi.org/10.1016/j.dib.2022.108346>

P. Schlich: Conceptualization, Methodology, Software, Supervision, Funding acquisition, Writing - Review & Editing; M. Visalli: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization. T. Wakihiro: Conceptualization, Methodology, Validation, Data Curation, Resources, Project administration, Investigation, Writing - Review & Editing.

This article describes a dataset providing temporal sensory descriptions and preferences for four lemon-flavoured carbonated alcoholic drinks. The recruited Japanese consumers (97 men, 96 women) corresponded to the target for this kind of drink: aged between 20 and 40 and regular consumers of flavoured alcoholic drinks. They had to consume a whole can of each drink at home, each on a different day. For sips 1, 4 and 7, they had to check from a check-all-that-apply (CATA) list of eight attributes (alcohol, bitter, carbonated, lemon, refreshing, sour, sweet aroma and sweet taste) that were applicable during three periods of perception – “in mouth before swallowing”, “immediately after swallowing” and “aftertaste”. They were separated into two panels: the consumers in panel SIM (96 consumers) had to do the task simultaneously with the tasting, while the consumers in panel RET (97 consumers) had to do it retrospectively. They also had to rate their liking and report the number of crackers they consumed during the tasting. Once the can had been fully consumed, they had to score their satisfaction level and optionally report comments about the products and the task. The data were used to compare retrospective and concurrent temporal evaluations in a methodologically oriented article entitled “Concurrent vs. immediate retrospective temporal sensory data collection: A case study on lemon-flavoured carbonated alcoholic drinks.” The data could also be reused by researchers interested in understanding interactions between alcohol, carbonation, sour, sweet and bitter or to relate temporal perception and preferences for improving product formulation.

1. Specifications table

Table 20. Specifications table

Subject	Food science
Specific subject area	Sensory analysis > Temporal perception > Alcoholic drinks
Type of data	Table Figure
How the data were acquired	Sensory data were acquired by 2 panels of consumers (193 consumers) at home, using the web application TimeSens version 2 (INRAE, 2022).
Data format	Tables in raw format (XLSX file)
Description of data collection	Japanese consumers (97 men, 96 women) corresponding to the target of this kind of drink (aged between 20 and 40 and regular consumers of flavoured alcoholic drink) were recruited. They had to consume a whole can of each drink at home, each on a different day. For sip 1, 4 and 7, they had to check from a CATA list of eight attributes (alcohol, bitter, carbonated, lemon, refreshing, sour, sweet aroma and sweet taste) which were applicable during three periods of perception – “in mouth before swallowing”, “immediately after swallowing” and “aftertaste”. They were separated into two panels, and the consumers in panel SIM (96 consumers) had to do the task during the tasting, while the consumers in panel RET (97 consumers) had to do it retrospectively. They also had to rate their liking (on a 0-10 visual analogic scale) and report the number of crackers they consumed during the tasting. Once the can had been fully consumed, they had to score their satisfaction level (on a 0-10 visual analogue scale), and optionally report comments about the products and the task (free text).
Data source location	City/Town/Region: All over Japan Country: Japan
Data accessibility	Repository name: Mendeley data Data identification number: 10.17632/729tsts8ng.1 Direct URL to data: https://data.mendeley.com/datasets/729tsts8ng/1
Related research article	M. Visalli, T. Wakihira, P. Schlich, Concurrent vs. immediate retrospective temporal sensory data collection. A case study on lemon-flavoured carbonated alcoholic drinks. Food Quality and Preference.

2. Value of the data

- These data are useful because they provide information about temporal sensory perception and preferences for lemon-flavoured carbonated alcoholic drinks by consumers over several sips of the full consumption of a can.
- Researchers interested in comparing temporal descriptions of products obtained concurrently and retrospectively to the tasting can benefit from these data.
- Researchers can also use these data to study interactions between flavoured alcoholic drinks key descriptors (alcohol, carbonation, sour, sweet, bitter).
- Product developers can relate temporal perception and preferences to investigate drivers of liking to improve their product formulation.
- Product developers can reproduce the original protocol that allowed us to collect temporal descriptions from consumers at home in ecological conditions.
- Sensometricians can reuse data to develop or test new statistical methods for the analysis of discrete time temporal sensory data.

3. Data description

The dataset is provided as an Excel file (.xlsx) including 5 sheets.

The sheet “Consumers” provides information about the recruited consumers. “Panel” is the panel to which the consumers have been randomly assigned (RET: retrospective temporal description, SIM: simultaneous temporal description). “Consumer” is the 4-character code of the consumer. The “Panel” and “Consumer” columns are reported in each sheet following this one. “Gender” is the gender of the consumer (male or female). “Age” is the age of the consumer (numeric). “Frequency” is the frequency of drinking of flavoured alcoholic drinks of any type (approximately two or three days a week, approximately four or five days a week or almost every day). “Brand” is the brand that they most frequently drank in the past month (P2, P3, P4, P5 or Other).

The sheet “Drinking mode” contains in column “Answer” the answer to the question [Q1]: “Please select your drinking mode in the list” (Directly from the can, From a glass without ice, From a glass with ice).

The sheet “Description” contains the sensory descriptions of each consumer for each of the 3 measured sips (1, 4, 7) of the 5 products. “Sip” is the sip position (1, 4 or 7). “Product” is the

2-character code of the lemon-flavoured carbonated alcoholic drink. "Product" is reported in each sheet following this one. "Attribute" is the code of the evaluated descriptor (Carbonated, SweetF for "sweet aroma", Refreshing, Bitter, Lemon, SweetT for "sweet taste", Sour, Alcohol). "Period" is the period of the evaluation (T1 for "in mouth before swallowing", T2 for "immediately after swallowing" and T3 for "aftertaste"). "Score" is 1 if "Attribute" has been checked and thus considered applicable for "Product" by "Consumer" during "Period", 0 otherwise.

The sheet "Hedonic" contains the liking and satisfaction scores of each consumer for the 5 products. "Sip" is the sip position for liking, (1, 4, or 7). "Attribute" is the code of the evaluated hedonic descriptor (Liking for question [Q2] "The sip I just drank: How delicious did you think it was?", Satisfaction for question [Q3] "After drinking the full can, how is your overall satisfaction?"). "Score" takes a numeric value between 0 and 10 (precision=0.01).

The sheet "Crackers" contains in column "Answer" the answer to the question [Q5] "How many crackers did you eat?" (numeric value between 0 and 12) for each consumer for the 5 products.

The answers to the free text questions Q4 and Q6 have not been included in the dataset because of the number of missing answers.

Table 21 displays information about the products. Alcohol content and lemon juice content were obtained from the product package. Brix, acidity, pH, limonene and citral were analysed in the Asahi Group R&D Center. ABV=alcohol by volume, ppm=parts per million.

Figure 25 describes the process of the tasting of the can over seven sips (common for RET and SIM).

Figure 26 describes the procedure of the description of sips 1, 4, and 7 for panel RET.

Figure 27 describes the procedure of the description of sips 1, 4, and 7 for panel SIM.

4. Experimental design, materials and methods

4.1. Samples

The five products (P1, P2, P3, P4, and P5) were 350 ml cans of commercial lemon-flavoured carbonated drinks. P1 was a non-alcoholic drink that served as a warm-up. P2 to P5 were white liquor-based (Japanese Shochu or Vodka) alcoholic drinks, referred to as “Chu-hai” in Japanese.

Table 21. Information about products

Sample	Alcohol content (% ABV)	Lemon juice content (%)	Brix (%)	Acidity (g/100 ml)	pH	Limonene (ppm)	Citral (ppm)
P2	6	3	6.07	0.41	3.43	2.3	0.05
P3	4	14	4.82	0.53	3.44	43.6	1.36
P4	5	10	6.33	0.52	2.88	28.4	0.08
P5	5	1.6	2.63	0.49	3.65	3.7	0.13

The cans were purchased at stores and sent from a research agency to the consumers' houses. The participants were instructed to put the products in the refrigerator for five or more hours after they had received them and, to the greatest extent possible, to taste the products at the same temperature. To accompany drinks, they also received plain, unsalted crackers. The cans were blinded by white-coloured masking films and coded using three-digit labels, and they were presented according to a Williams Latin square, but P1 was served first to every consumer.

4.2. Consumers

Two hundred consumers aged 20 to 39, who were regular consumers of lemon-flavoured carbonated alcoholic beverages, were recruited through an online questionnaire from a panel of consumers belonging to a research agency in Japan. “Regular” referred to consumers drinking flavoured carbonated alcoholic beverages with a frequency of at least twice a week and lemon-flavoured carbonated alcoholic beverages at least once a month. The design of the test was explained to consumers in the online questionnaire. They were informed they would do the test on their smartphones, and they had to sign a consent form to participate in the study. They were financially compensated for their participation. The consumers were separated into two panels that were balanced in their composition (gender, age, frequency of

consumption). The first panel (RET) had to evaluate the products retrospectively to the tasting just after they declared that they no longer perceived anything, while the second (SIM) evaluated them concurrently to the tasting. One hundred and ninety-three consumers finally participated, 97 consumers in panel RET, and 96 in panel SIM.

4.3. Descriptors

The descriptors were chosen according to the expertise of Asahi. The same list of descriptors was provided for both the RET and SIM panels: Alcohol, Bitter, Carbonated, Lemon, Refreshing, Sour, Sweet aroma and Sweet taste. (In Japan, sweet aroma is used when it is not possible to describe detailed quality of sweetness in terms of aroma. Sweet aroma can include different types of aromas, such as fruity, floral, caramel, vanilla, honey, etc.) The descriptors were presented as a check-all-that-apply (CATA) list in a random order on the screen, but this order was constant for each consumer across evaluations. No definitions of the descriptors were given to the consumers.

4.4. Data collection

Consumers received an email containing an individualized URL to invite them to connect to the session using their smartphones. All instructions (here translated from Japanese) were presented in the web application. The first screen reminded the consumers they had to fully consume each can, drink one a day, and to use their smartphones for the test: “In this survey, you are going to drink one can a day. On the first day, a non-alcoholic lemon carbonated drink will be tasted to familiarize ourselves with the tasting method. Alcoholic lemon carbonated drinks will be tasted from Day 2 to Day 5. Please complete this survey on your smartphone”. Another screen asked the consumers about their drinking mode and invited them to keep their drinking mode constant for all the products: “During the tasting period of five days, please taste the drinks in the same way. Those who usually drink directly from the can also drink directly from the can during the tasting period. Those who usually pour into a glass and drink pour into a glass even during the tasting period. [Q1] Please select your drinking mode in the list.” Finally, an overview of the tasting of each can over multiple sips was presented: “While drinking, follow the instructions on your smartphone and answer in order. First, you will be asked your perceived sensations and liking for the first sip. Then, you will be asked to take a second and third sip (no question). Then, you will be asked your perceived sensations

and liking for the fourth sip. Then, you will be asked to take a fifth and sixth sip (no question). Then, you will be asked your perceived sensations and liking for the seventh sip. After the 7th sip, you will be asked to finish the can (no question) according to your own timing.”.

Procedure for tasting, repeated on each day

Before each tasting of a new product (on each day), the tasting procedure was reminded, depending on the panel.

For panel RET: “First, you will be asked to focus and remember the perceived sensation(s) following the mouthing (before swallowing). The screen will change after eight seconds. Second, when "Please swallow" will be displayed on the screen, you will have to swallow the sip. Third, you will have to focus and remember the perceived sensation(s) after swallowing. The screen will change after 18 seconds. Fourth, you will have to remember the taste and aroma remaining. When you no longer sense anything, you will have to press the Stop button. Finally, you will have to check the attribute(s) corresponding to the sensation(s) that applies during the three preceding periods: before swallowing, immediately after swallowing, and taste and aroma remaining afterwards.”.

For panel SIM: “First, you will have to check the perceived sensation(s) following the mouthing (before swallowing). The screen will change after eight seconds. Second, when "Please swallow" will be displayed on the screen, you will have to swallow the sip. Third, you will have to check the perceived sensation(s) after swallowing. The screen will change after eight seconds. Fourth, you will have to wait for ten seconds. Fifth, you will have to check the remaining perceived sensation(s). When you no longer sense anything, you will have to press the Next button.

Then, a training exercise presenting the list of attributes (CATA list) was proposed: “Please select all the sensations that apply from the eight items (by clicking on the checkbox). The order of the eight items will be the same for each evaluation. Remember the position of each item. Do not taste yet. Practise clicking on the checkbox from the list below on this screen. If you make a mistake, click it again to uncheck.”

Then, the procedures for the full consumption of the can and for the evaluation of each sip were displayed on the screen as charts.



Figure 25. Description of the tasting of the drink over seven sips (common for RET and SIM). Only sips 1, 4 and 7 were evaluated.

Tasting: sips 1, 4 and 7

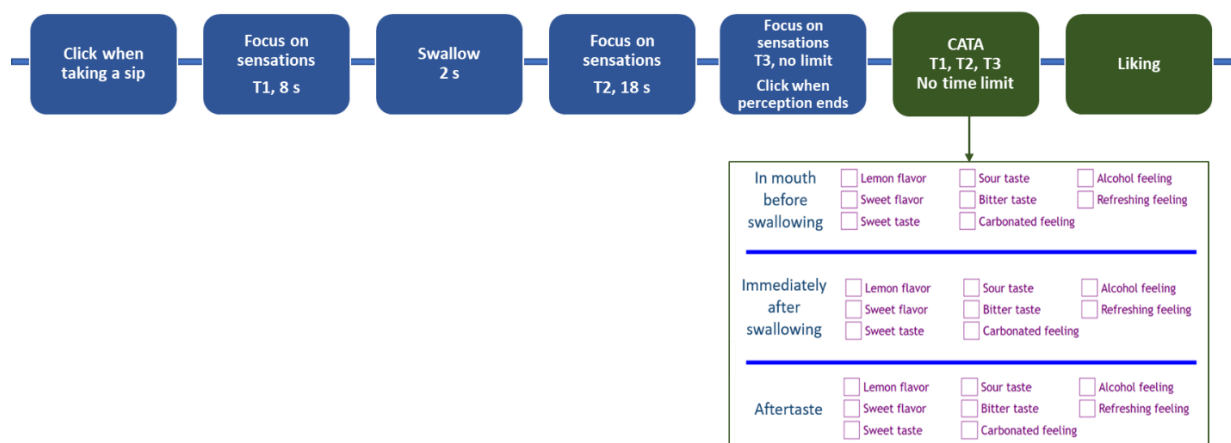


Figure 26. Evaluation of sips 1, 4, 7 for panel RET using retrospective CATA. Only one CATA measurement screen was displayed, after the tasting. “In mouth before swallowing”, “Immediately after swallowing” and “Aftertaste” correspond to periods T1, T2 and T3.

Figure 26 describes the tasting of sips 1, 4, 7 for panel RET. During the tasting, several screens were displayed to guide the consumer. A first screen invited the consumer to take a sip without swallowing. Then, a screen was displayed during eight seconds (T1, “In mouth before swallowing”) instructing the consumer not to swallow while focusing on perceived sensations and memorizing them. Then, a screen was displayed during two seconds, inviting the consumer to swallow. Then, a screen was displayed during eighteen seconds (T2, “Immediately after swallowing”) instructing the consumer to focus on perceived sensations and memorize them. Then, a screen instructing the consumer to focus on perceived sensations and memorize them was displayed until the consumer declared he no longer perceived anything (T3, “Aftertaste”). Then, the CATA screen was displayed with no time limit, with the instruction: “What did you perceived during the different steps of the tasting? Please select all the sensations that apply during each period”.

Finally, the liking score (“[Q2] The sip I just drank: How delicious did you think it was? Touch on the line and answer intuitively”) was asked to the consumer (on a 0-10 unstructured scale, 0 being labelled “not at all” and 10 “very much”).

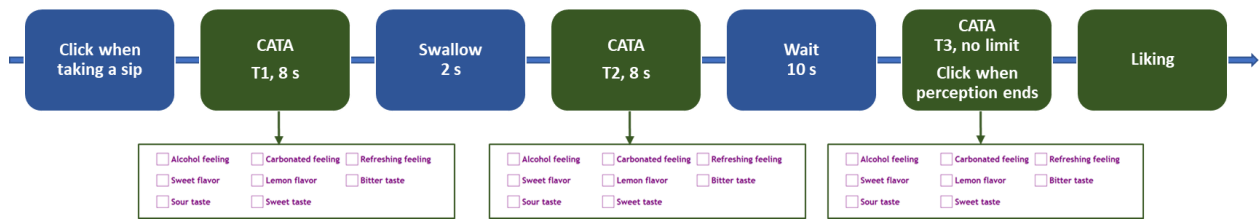


Figure 27. Evaluation of sips 1, 4, 7 for panel SIM using simultaneous CATA. The CATA measurement screen was displayed during the tasting within periods “In mouth before swallowing” (T1), “Immediately after swallowing” (T2) and “Aftertaste” (T3).

Figure 27 describes the tasting of sips 1, 4, 7 for panel SIM. During the tasting, several screens were displayed to guide the consumer. A first screen invited the consumer to take a sip without swallowing. Then, a CATA screen was displayed during eight seconds (T1, “In mouth before swallowing”), with the instruction: “What do you perceive now? Please select all the sensations that apply”. Then, a screen was displayed during two seconds, inviting the consumer to swallow. Then, a screen was displayed during eight seconds (T2, “Immediately after swallowing”) instructing the consumer to select all the sensations that apply (same instruction as in T1). Then, a screen inviting the consumer to wait was display during ten seconds. Then, the last CATA screen was displayed until the consumer declared he no longer perceived anything (T3, “Aftertaste”), instructing the consumer to select all the sensations that apply (same instruction as in T1). Finally, the liking score (“[Q2] The sip I just drank: How delicious did you think it was? Touch on the line and answer intuitively.”) was asked to the consumer (on a 0-10 unstructured scale, 0 being labelled “not at all” and 10 “very much”).

Tasting: sips 2, 3, 5 and 6

On sips 2, 3, 5 and 6, neither the CATA nor the hedonic measurement screen was displayed, but the following message was shown: “Please relax and wait until you want to take another sip. If you want to eat crackers, you can. If you want to take another sip, press Next before taking it.

Tasting: drinking of the rest of the can

After the 7th sip, the consumers were invited to finish the can: “Please continue to drink at your own pace until you finish drinking the whole can. If you want to eat crackers, you can. (two bags = 12 pieces). When you finish drinking, press Next.”

Then, they had to rate their overall satisfaction using an unstructured scale (0 being labelled “Not satisfied at all” and 10 “Strongly satisfied”): “[Q3] After having finished the full can, how is your overall satisfaction? Touch on the line and answer intuitively.”

Then, a free answer question was displayed: “[Q4] From the first sip to the last one, have you perceived any change in your sensations? If yes, how did your perception change? (feel free to enter anything). If no, enter “No change”.”

Then, the consumers were requested to select in a drop-down list the number of consumed crackers: “[Q5] How many crackers did you eat? (One bag contains six pieces of crackers).” This ended the session for the 4 first products: “Today's tasting is over. Open the same link again at a later date.”

After the 5th and last product, a free answer question was displayed: “All tastings are finished. [Q6] Please tell us what you felt about this tasting survey (if the tasting method was difficult, easy, etc.). Please be as specific as possible.” Finally, the last screen thanked the consumers for their participation and invited them to exit the session.

II.3. A dataset on the sensory and affective perception of Bordeaux and Rioja red wines collected from French and Spanish consumers at home and international wine students in the lab

This article was *published in Data in brief*:
<https://www.doi.org/10.1016/j.dib.2022.108873>

J.M. Cardebat: Funding acquisition, Review & Editing; M. Dubois: Conceptualization, Data Curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - Review & Editing; N. Georgantzis: Review & Editing; F. Ric: Review & Editing; P. Schlich: Review & Editing; M. Visalli: Conceptualization, Data Curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing - Original Draft.

This article describes a dataset providing temporal sensory descriptions and affective answers for red wines: two Bordeaux and two Riojas. The wines were tasted at home by French (FR, n=106) and Spanish (SP, n=98) consumers and in the lab by wine students (WC, n=47). Standardized information was displayed on the samples (country and region of origin, name, producer, vintage, alcohol content). The FR and SP panels were split into three groups, the first having no rating information, the second having expert rating information (based on Wine Advocate ratings), and the third having consumer rating information (based on online Vivino reviews). The participants first rated their expected liking for the four wines. Then, for each wine sample, they had (in order) to taste the sample while being video recorded, rate their liking, temporally describe the sequence of sensations they perceived using Free Comment Attack-Evolution-Finish, answer several questions about familiarity and quality perception, and declare their willingness to pay (reserve price). Then, they had to rank the four wines according to their quality. General questions about wine involvement, subjective wine knowledge, valuation behaviour, purchasing, and consumption patterns were asked. Finally, an auction was resolved: participants declaring a reserve price greater than the drawn price won a bottle. The data were used to assess the influence of culture and expertise on temporal sensory evaluations in an article entitled “Using Free Comment to investigate expertise and cultural differences in wine sensory description”. The data can be reused by researchers interested in studying the impact of external information on preferences and choices or investigating the sensory drivers of liking.

1. Specifications table

Table 22. Specifications table

Subject	Food science
Specific subject area	Wines
Type of data	Tables Questionnaire Figures
How the data were acquired	Sensory data were acquired by recruiting two panels of consumers (98 Spanish, 106 French) at home through Qualtrics and one panel of wine students (47 international students) using a web application.
Data format	Tables in raw format (XLSX file)
Description of data collection	Three panels tasted two Bordeaux and two Rioja red wines. Standardized information was displayed on the samples (origin, designation, producer, vintage, alcohol content). The French and Spanish panels were split into three groups, the first having no rating information, the second and third having additional external information, respectively expert rating and consumer rating for the evaluated sample. The participants first rated their expected liking for the four wines on a 7-point scale. Then, for each wine sample, they had (in order) to taste the sample while being video recorded; rate their liking on a 7-point scale; temporally describe the sensations they perceived using Free Comment Attack-Evolution-Finish; answer several questions (about their familiarity, perception of quality, perception of others' liking); and declare their willingness to pay (reserve price). Then, they had to rank the four wines according to their quality. General questions about their behaviour towards wines were asked (interest, choice, knowledge, purchases, consumption).
Data source location	City/Town/Region: Dijon (students), everywhere in France or Spain (consumers) Country: France, Spain
Data accessibility	Repository name: Mendeley data Data identification number: 10.17632/f9wtj7s9b8.1 Direct URL to data: https://data.mendeley.com/datasets/f9wtj7s9b8/1
Related research article	M. Visalli, M. Dubois, P. Schlich, F. Ric, J.M. Cardebat, N. Georgantzis. Using Free Comment to investigate expertise and cultural differences in wine sensory description. Food Quality and Preference.

2. Value of the data

- These data are useful because they provide information about wines' perceived quality, stated and revealed preferences collected from consumers and experts from different countries under different information conditions.
- Researchers or product developers can reuse these data to test the impact of external information (peer and expert ratings) on sensory perception, preferences or willingness to pay for wines. They can also study whether this impact varies according to the expertise and culture of the participants.
- They can compare the information obtained with a sensory evaluation with that available in online wine reviews. They can also benefit from these data to investigate the drivers of liking for red wines.

3. Objective

This dataset has been generated in order to compare the expectations, temporal sensory perception, willingness to pay, and preferences for two Bordeaux and two Rioja red wines tasted in two settings (at home and in the lab) by three panels varying in culture and expertise and having different information about the wines. The research article associated with this data paper only reports results on temporal sensory perception.

4. Data description

The dataset is provided as an Excel file (.xlsx) including five sheets:

Participants provides information about the participants collected during the screening and in the questionnaire.

- "Panel" is the panel to which the participant has been assigned (WC: wine connoisseurs, FR: French consumers, SP: Spanish consumers).
- "Language" is the language of the participant. "Participant" is the unique anonymized identifier of the participant.
- "Age" is the age range of the participant (18-29, 30-40, 41-50, 51-60, 61-70, 71 and older).
- "Gender" is the gender of the participant (Male, Female, or Other).
- "PCS" is the socio-professional category of the participant (Employed, Unemployed, Student, Retired, Other).
- "Diploma" is the higher level of diploma obtained by the participant (None, High school, Vocational school, Bachelor's degree, Master's degree, PhD, Other).

- “Income” is the income-related quality of life estimated by the participant (Living comfortably on present income, Coping on present income, Finding it difficult on present income, Finding it very difficult on present income, I prefer not to answer).
- “Group” is the group of the participant (No rating information, Consumer rating information, Expert ratings information).
- “QuestionnaireBeginDate” is the date (YYYY-MM:DD hh:mm:ss) when the participant started the study (first connection).
- “QuestionnaireEndDate” is the date (YYYY-MM:DD hh:mm:ss) when the participant ended the study.
- “PCI1”, “PCI2” and “PCI3” are related to product category involvement (Bruwer et al., 2014).
- “PC1” is the answer to the question “Wine interests me a lot” rated on a 5-point Likert scale (1: strongly disagree, 2: disagree, 3: neither agree nor disagree, 4: agree, 5: strongly agree).
- “PC2” is the answer to the question “I often discuss wine with other people” rated on a 5-point Likert scale (same values as PCI1).
- “PC3” is the answer to the question “It gives me pleasure to shop for wine” rated on a 5-point Likert scale (same values as PCI1).
- “SWK1”, “SWK2”, “SWK3” and “SWK4” are related to subjective wine knowledge (Mueller Loose et al., 2008).
- “SWK1” is the answer to the question “I feel confident in my ability to choose wine” (Kapferer & Laurent, 1993) rated on a 5-point Likert scale (same values as PCI1).
- “SWK2” is the answer to the question “I know more about wine than many other people” rated on a 5-point Likert scale (same values as PCI1).
- “SWK3” is the answer to the question “I would describe myself as being very knowledgeable about wine” rated on a 5-point Likert scale (same values as PCI1).
- “SWK4” is the answer to the question “Did you already follow a wine education course?” (1: no, 2: yes, without certification, 3: yes, with certification).
- “CP1” and “CP2” are related to consumption patterns (Bruwer & Buller, 2013).
- “CP1” is the answer to the question “How often do you consume wine at home?” (1: daily, 2, at least once a week, 3: at least once a month, 4: less than once a month, 5: never).
- “CP2” is the answer to the question “How often do you consume wine outside of home (restaurant, bar, club, etc.)?” (same values as CP1).
- “AC1” is the answer to the question “If you read this sentence correctly, please answer ‘strongly disagree’” (Lunardo et al., 2021) rated on a 5-point Likert scale (same values as PCI1, expected answer=1).
- “VB1”, “VB2”, “VB3” and “VB4” are related to valuation behaviour (Thach & Chang, 2015).
- “VB1” is the answer to the question “I use wine apps to help me decide which wine to buy” (1: never, 2: once in a while, 3: often, 4: always).

- “VB2” is the answer to the question “I use wine professional expert ratings (wine reviews, point scores, medals, and awards) to help me decide which wine to buy” (same values as VB1).
- “VB3” is the answer to the question “I often seek advice from other people before purchasing a wine” rated on a 5-point Likert scale (same values as PCI1).
- “VB4_1” to “VB4_8” are the answers to the question “Whose advice do you trust most when selecting a wine?” (VB4_1: Friends, VB4_2: Family members, VB4_3: Colleagues, VB4_4: Sommelier, VB4_5: Professional Wine Expert, VB4_6: Wine blogger or influencer, VB4_7: Wine Guide or Magazine, VB4_8: Only my own). Answers are 1 if the option was checked and 0 otherwise.
- “PP1”, “PP2”, “PP3”, and “PP4” are related to purchasing patterns (Thach & Olsen, 2015).
- “PP1_1” to “PP1_7” are the answers to the question “Where do you buy your wine?” (PP1_1: I do not buy wine, PP1_2: Supermarket, PP1_3: Wine store, PP1_4: Online, PP1_5: Directly from the winemaker, PP1_6: Restaurants and bars, PP1_7: Other). Answers are 1 if the option was checked and 0 otherwise.
- “PP2” is the answer to the question “For a 75-cl bottle of red wine, you spend on average - for informal drinking” (1: I do not buy wine, 2: less than 5€, 3: 5-10€, 4: 11-20€, 5: 21-30€, 6:30€ and more).
- “PP3” is the answer to the question “For a 75-cl bottle of red wine, you spend on average - for a formal occasion or a gift” (same values as PP2).
- “PP4” is the answer to the question “How much does your household spend on wine monthly?” (1: 0€, 2: 50€ or less, 3: 51-100€, 4: 101-150€, 5: 151€ or more).

Qualities provides information about the subjective qualities evaluated for each wine by the participants.

- “Panel” and “Participant” are the same as in the “Participants” tab.
- “Wine” is the code of the evaluated wine.
- “EL1” is the answer to the question “How much do you think to like this wine?” (Danner et al., 2020), rated on a 7-point hedonic scale (1: dislike extremely, 2: dislike moderately, 3: dislike slightly, 4: neither like nor dislike, 5: like slightly, 6: like moderately, 7: like extremely).
- “SL1” is the answer to the question “How much did you like this wine?” (Danner et al., 2020), rated on a 7-point hedonic scale (same values as EL1).
- “QE1” is the answer to the question “I think this wine is high quality”, rated on a 5-point Likert scale (same values as PCI1).
- “QE2” is the answer to the question “Most people would like this wine”, rated on a 5-point Likert scale (same values as PCI1).
- “F1” is the answer to the question “This wine profile is familiar to me”, rated on a 5-point Likert scale (same values as PCI1).
- “RQ1” is the answer to the question “Rank the 4 wines by clicking on ‘click to choose a wine’ and then giving each wine a rank. #1=most qualitative wine, #4=least qualitative wine”.

- “BDM1” is the answer (reserve price) to the question “What is the maximum price (in euros) you are willing to pay for a 75cl bottle of the wine you just tasted?” (Lohéac & Issanchou, 2007).

TemporalPerception provides information about temporal sensory perception evaluated for each wine by the participants.

- “Panel” and “Participant” are the same as in the “Participants” tab.
- “Wine” is the code of the evaluated wine.
- “Period” is the code of the period defined in the question (AEF1: “At first, I perceived this wine”, AEF2: “Then, after a few moments, I perceived it”, AEF3: “At the end of the tasting, I perceived it”).
- “Description” is the free comment description of “Product” at “Period” depending on the period, as entered by “Participant” (in English, French or Spanish).
- “Keywords” contains the lemmas related to sensory attributes (canonical form, masculine, singular), separated by commas, translated into English (if required).

AuctionResolution provides information about the results of the auction.

- “Panel” and “Participant” are the same as in the “Participants” tab.
- “RandomProduct” is the code of the randomly drawn wine (between W1, W2 or W4, W3 being not available at the end of the study).
- “RandomPrice” is the price randomly drawn in the distribution of the prices of red wines of Bordeaux and Rioja (extracted from Vivino).
- “Result” was the result of the auction (win if RandomPrice \geq reserve price for RandomProduct, lose otherwise).

Questionnaire 1 includes commented screenshots of the online questionnaire used to collect data, translated from French and Spanish to English.

Table 23 provides objective information about wines.

Table 24 provides subjective information about wines collected on the Robert Parker Wine Advocate and Vivino websites in May 2022.

Table 25 describes the individual characteristics of the participants in the three panels by group.

Figure 28 shows the standardized labels displayed on the wine samples.

Figure 29 is the flowchart of the participants (recruited, participated, completed).

5. Experimental design, materials and methods

5.1. Samples

The four products (W1, W2, W3, and W4) were nonorganic red wines produced in France or in Spain.

Table 23: Objective information about wines.

Code	Appellation	Winery	Variety	Vintage	Alcohol Content	Price in euros
W1	AOC Bordeaux Supérieur	Chateau Féret Lambert	Merlot 90%, Cabernet 10%	2018	14.5	15.5
W2	AOC Pessac Léognan	La Louvière	Merlot 40%, Cabernet Sauvignon 60%	2018	13.5	15.6
W3	DOP Rioja	Bhilar	Tempranillo 85%, Grenache 10%, Viura 5%	2018	13.5	16
W4	DOP Rioja	Miguel Merino	Tempranillo 100%	2018	14.0	14.5

AOC: Appellation d'Origine Contrôlée (protected designation of origin). DOP: Denominación de Origen Protegida (protected designation of origin).

The 750 ml bottles were purchased from an online store. The wines chosen came from two regions to study the cultural impact and have noncongruent Vivino and Parker scores to study the impact of information (expert or consumer ratings from reviews). The year of production was the same, and the prices were very close.

The reviews in Wine Advocate were conducted by highly experienced tasters, specially trained to understand and recognize wine quality in a glass. They used 100-point quality scales to evaluate the wines. The Vivino taste profiles of wines were based on user reviews. The values reported in Table 24 were measured on the bipolar scales ("Light/strong", "Supple/Tannic", "Dry/Liquorous", "Sweet/Acid") displayed on the website and converted between 0 and 10. The percentages associated with the flavours correspond to the percentages of consumers who cited the flavour according to Vinino's counts.

Table 24. Subjective information about wines.

Code	Parker score	Vivino score	Parker description	Vivino description
W1	85/100	3.76/5 (383 notes)	“Deep garnet-purple coloured, the 2018 Feret-Lambert leaps from the glass with crème de cassis, boysenberries and black raspberries followed by plum pudding and cloves nuances. Full-bodied, it coats the mouth with dried berries and exotic spice flavours, framed by chewy tannins and just enough freshness, finishing earthy”. (Lisa Perrotti-Brown, 23rd Apr 2019)	Light/Strong: 8.5/10; Supple/Tannic: 6.5/10; Dry/Liquorous:0.5/10; Sweet/Acid: 7/10; Prune, blackberry, black fruit (79%); Oak, vanilla, tobacco (46%); Earthy, leather, smoky (41%); Cherry, raspberry, red berries (32%); Pepper, liquorice, anise (14%)
W2	94/100	4.07/5 (84 notes)	“Medium to deep garnet-purple in colour, the 2018 la Louviere leaps from the glass with notions of redcurrant jelly, fresh blackberries and warm black plums, plus nuances of dried mint, cedar chest and ground cloves. The medium-bodied palate is refreshing and savoury in the mouth, featuring a light touch of finely grained tannins and bold freshness, finishing with a compelling red berry lift. It's an elegant, lively expression of this vintage and one that really works! “(Lisa Perrotti-Brown, 31st Mar 2021)	Light/Strong: 8/10; Supple/Tannic: 7.5/10; Dry/Liquorous:0.5/10; Sweet/Acid: 8.5/10; Oak, vanilla, tobacco (37%); Somey, leather, cocoa (21%); Blackberry, black fruit, blackcurrant (19%); Cherry, strawberry, sour cherry (14%); Licorice, pepper, anise (13%)
W3	88/100	3.21/5 (16 notes)	“It has a strong note of cider on the nose, with a volatile touch and a nutty touch and very low alcohol.” (Luis Gutiérrez, 28th Jun 2019)	Light/Strong: 4.5/10; Supple/Tannic: 6/10; Dry/Liquorous:3.5/10; Sweet/Acid: 6.5/10; Blackcurrant (50%)
W4	90/100	4.10/5 (135 notes)	“The wine is juicy, very drinkable but serious, with fantastic balance and very clean aromas and flavours. It's quite fruit driven but has the complexity of a more serious wine produced with attention to detail. It already has Rioja character.” (Luis Gutiérrez, 30th Oct 2020)	Light/Strong: 7/10; Supple/Tannic: 6.5/10; Dry/Liquorous:2/10; Sweet/Acid: 6.5/10; Vanilla, oak, chocolate (41%); Prune, black cherry, black fruit (27%); Cocoa, leather, earthy (21%); Cherry, strawberry (14%); Pepper, liquorice, anise (7%)

Vinovae (*Vinovae*) used a patented process to repack the bottles in 20 ml polyethylene terephthalate (PET) screw bottles in an inert atmosphere devoid of oxygen. The process was designed to avoid the risk of oxygenation and ensure the preservation of the organoleptic qualities of the wines. The screw bottles were labelled with original information (origin, designation, producer, vintage, alcohol content) and displayed in a standardized way. The samples were sent to the consumers' houses by postal mail.

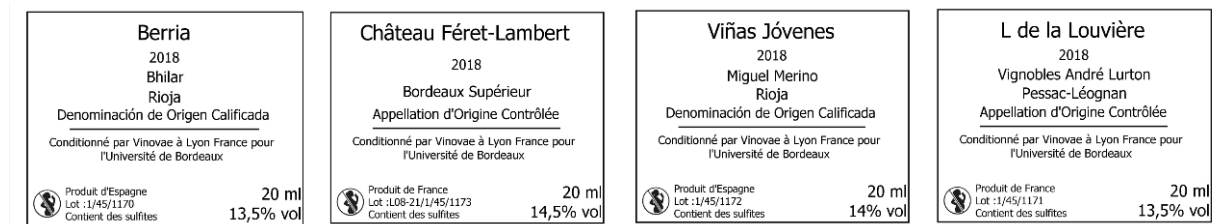


Figure 28. Standardized labels of the samples.

5.2. Participants

“Wine connoisseur” panel (WC): Fifty students of the School of Wine and Spirits Business of the Burgundy School of Business in Dijon were recruited through a mailing. They were natives from different countries and selected based on their knowledge about wines (they were at least holders of WSET certificate (Wine and Spirits Education Trust) Level 2 Award in Wines (*Wine & Spirit Education Trust qualification*)).

French (FR) and Spanish (SP) consumer panels: A total of 150 French consumers and 150 Spanish consumers were recruited from a panel recruitment agency database through online questionnaires (Qualtrics). The selection criteria included (i) being available to participate in a 20-minute online study involving the at-home tasting of four wines; (ii) having consumed red wine at home within the past month; (iii) possessing at least one wine glass at home; (iv) owning a computer with a webcam and a good internet connection (Chrome, Firefox or Edge browser); (v) agreeing to be video recorded during the tasting; and (vi) agreeing to provide a postal address for the shipment of samples. Quotas on individual characteristics (age, gender, employment status, education level, family income) were also established to balance the two consumers panels. The appropriate sample size was determined based on the literature (Asioli et al., 2021).

All participants (WC, FR, SP) signed an informed consent form. They were informed that they would receive compensation worth 20 euros (that may include a bottle of wine depending on

the resolution of an auction), and they could refuse to participate or stop participating in the study at any time without providing a reason; however, in that case, they would not receive any compensation. They were also informed that the study was an academic research project without any commercial interests and that the information collected would be used exclusively for research purposes.

Consumers in FR and SP panels were both divided into three groups of 50: the first (control group) had no other information than the label (“no rating information”), the second received information about expert ratings (“expert rating information”), the third received information about consumer ratings (“consumer rating information”).

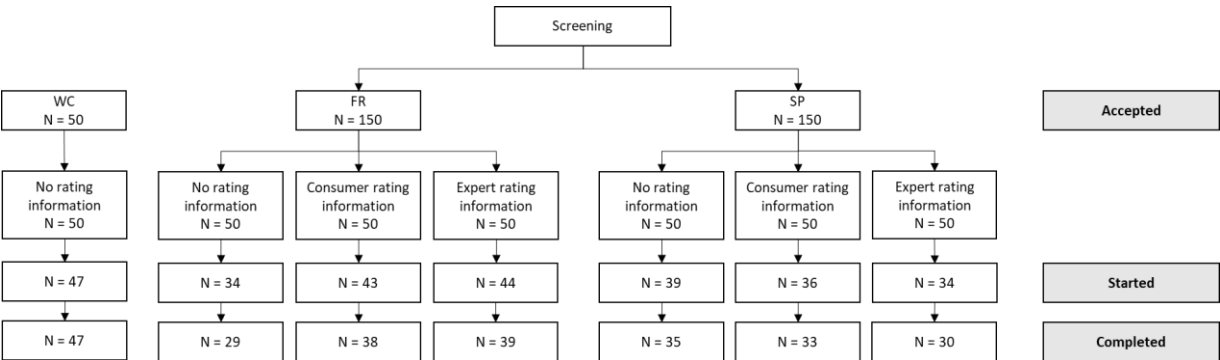


Figure 29. Participant flowchart.

Forty-seven (94%) students finally completed the study in the WC panel, 106 consumers (70%) in the FR panel, and 98 (65%) in the SP panel.

Table 25. Participant characteristics.

	WC	FR no rating	FR consumer rating	FR expert rating	SP no rating	SP consumer rating	SP expert rating
Age 18-29	57%	7%	11%	13%	6%	10%	6%
Age 30-40	21%	17%	18%	15%	37%	30%	27%
Age 41-50	15%	21%	13%	26%	23%	30%	27%
Age 51-60	4%	17%	16%	13%	23%	13%	18%
Age 61-70	2%	31%	32%	21%	11%	17%	18%
Age 70+	-	7%	11%	10%	-	-	3%
Female	62%	59%	48%	46%	46%	60%	52%
Male	36%	41%	53%	52%	54%	40%	48%
Other	2%	-	-	3%	-	-	-
Employed	11%	55%	55%	59%	77%	73%	79%
Other	-	14%	-	3%	3%	3%	3%
Retired	-	28%	37%	26%	9%	3%	8%
Student	89%	3%	5%	5%	-	-	6%
Unemployed	-	-	3%	8%	11%	20%	3%
PhD	4%	-	-	8%	3%	7%	6%
High school	-	34%	24%	23%	6%	13%	12%
Bachelor's degree	51%	17%	17%	13%	49%	33%	48%
Master degree	43%	17%	24%	21%	20%	30%	21%
None	-	-	-	5%	-	-	-
Other	2%	-	5%	-	-	3%	-
Vocational school	-	31%	32%	28%	23%	13%	12%
Living comfortably	34%	10%	24%	23%	45%	40%	45%
Finding it difficult	9%	21%	13%	15%	12%	3%	12%
No answer	23%	-	3%	3%	-	-	-
Coping on income	34%	69%	61%	59%	42%	57%	42%

5.3. Data collection

The consumers received an email containing an individualized URL to invite them to connect to the TimeSens version 2 web application (INRAE, 2022) using a web browser (Chrome, Firefox, or Edge were recommended to ensure maximum compatibility with the web app) on their computer. The FR and SP panels completed the experiment at home, and the WC panel completed the experiment in the sensory lab of the Burgundy School of Business (Dijon, France, 32 available individual boxes) during three sessions.

The experimental procedure followed the steps of the questionnaire described below.

Screen 1: reading and acceptance of the conditions of the study.

Screen 2: reminder that four wines had to be evaluated and that the tasting part would be video recorded.

Screen 3: rating of expected liking of the four wines using a 7-point hedonic scale (question EL1).

Screen 4: instructions to prepare by having a glass of water and an empty wine glass available for the tasting.

Screen 5: instructions for webcam calibration included facing the webcam; having the face and forefront visible (no glasses); adapting the light to be homogeneous; having the face occupy 25 to 30% of the screen; avoiding white clothing, direct lighting, a dark environment, and anything that masks the face; and turning off the phone during the study.

Screen 6: instruction checklist displayed on screen 5.

Screen 7: displaying the video flux of the webcam to adjust the calibration of the webcam.

Screen 8: instructions for water tasting (warm-up) included pouring some water in the wine glass; looking at the water, swirling and sniffing it; and taking a small mouthful and at the same time clicking on the button to start the video recording.

Screen 9: displaying the video flux from the webcam during the water tasting (ten seconds).

Screen 10: instructions for preparing a new sample included emptying the wine glass; pouring the appropriate wine sample in the glass (the order of presentation of the samples was balanced over participants based on a Williams' Latin square); looking at the wine, swirling and sniffing it; and taking a small mouthful and at the same time clicking on the button to start the video recording. The label corresponding to the wine sample they had to taste was displayed on the screen (see Figure 28). The participants in the "expert rating information" group received the following supplementary information to the right of the sample label: "This wine was scored $x/100$ by the Wine Advocate – Robert Parker". The participants in the "consumer rating information" group also received supplementary information: "This wine was rated $x/5$ by consumers – Vivino website".

Screen 11: displaying the video flux from the webcam during the wine tasting (ten seconds). This was the only moment when a wine sample was tasted.

Screen 12: rating of liking on a 7-point hedonic scale (question SL1).

Screen 13: explanation of the wine description task, a Free Comment Attack-Evolution-Finish (FC-AEF, Mahieu et al., 2020b). The participants were informed that they had to retrospectively describe the sensations they perceived in mouth during the tasting in chronological order. Three periods were defined to summarize the tasting: "at first", "after a

few moments” and “at the end of the tasting”. For each period, they had to describe their sensations (tastes, aromas) using their own words. The same words could be used in different periods. A fictive example with chocolate was shown to help the participants understand the task.

Screen 14: FC-AEF task (questions AEF1, AEF2 and AEF3), as explained on screen 13.

Screen 15: rating of question F1 (“The wine I just tasted is similar to the wines I normally select”).

Screen 16: rating of questions QE1 (“I think this wine is high quality”) and QE2 (“Most people would like this wine”).

Screen 17: explanation of the Becker–DeGroot–Marschak (BDM) method (Becker et al., 1964). The participants were instructed they would have to propose a price corresponding to the maximum price they would pay for a 75-cl bottle of the tasted wine. They were informed that they could indicate 0 if they did not like the product and did not wish to buy it. They were told that an auction would happen only for one of the four wines randomly drawn at the end of the survey. For this wine, if the price drawn was higher than the indicated price, the participant would lose the auction and would not receive the wine. If the price drawn was lower than the indicated price, the participant would win the auction and receive the bottle in the following few days.

Screen 18: scoring of their maximum willingness to pay for a 75-cl bottle (question BDM1), as explained on screen 17.

Screen 19: instructions for glass rinsing.

Screen 20 to 42: the procedure described on screens 10 to 19 was repeated for the three other wines (explanation screens 13 and 17 were displayed once).

Screen 43: ranking of the four wines (no ex-aequo allowed, question RQ1).

Screen 44: rating of questions PC1 (“Wine interests me a lot”), PC12 (“I often discuss wine with other people”) and PC13 (“It gives me pleasure to shop for wine”).

Screen 45: rating of questions SWK1 (“I feel confident in my ability to choose wine”), SWK2 (“I know more about wine than many other people”), SWK3 (“I would describe myself as being

very knowledgeable about wine”) and SWK4 (“Have you already taken a wine education course?”).

Screen 46: answering questions VB1 (“I use wine apps to help me decide which wine to buy”) and VB2 (“I use wine professional expert ratings (wine reviews, point scores, medals and awards) to help me decide which wine to buy”).

Screen 47: rating of question VB3 (“I often seek advice from other people before purchasing a wine”).

Screen 48: answering question VB4 (“Whose advice do you trust most when selecting a wine?”, multiple answers authorized).

Screen 49: answering question PP1 (“Where do you buy your wine?”, multiple answers authorized).

Screen 50: answering questions PP2 (“For a 75cl bottle of red wine you spend on average - for informal drinking”), PP3 (“For a 75cl bottle of red wine you spend on average - for a formal occasion or for a gift”) and PP4 (“How much does your household spend on wine monthly?”).

Screen 51: answering questions CP1 (“How often do you consume wine at home?”); CP2 (“How often do you consume wine outside from home (restaurant, bar, club, etc.)?”) and AC1 (“If you read this sentence correctly, please answer ‘strongly disagree’”).

Screen 52: drawing of the wine and the random price and resolution of the auction.

Screen 53: study debriefing.

Screen 54: end screen.

II.4. A dataset of sensory perception of chocolates, guacamoles, iced teas and crisps collected with consumers using six temporal methods

This article was *published in Data in brief*:
<https://www.doi.org/10.1016/j.dib.2022.108708>

S. Cordelle: Methodology, Validation, Investigation, Resources, Project administration, Writing - Review & Editing; M. Coutière: Methodology, Investigation, Resources, Data Curation; B. Hoffarth: Resources; B. Mahieu: Methodology, Writing - Review & Editing; C. Pedron: Resources; M. Praudel: Methodology, Resources, Data Curation; P. Schlich: Conceptualization, Methodology, Supervision, Funding acquisition, Writing - Review & Editing; M. Visalli: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization.

This article describes a dataset providing temporal sensory perception data of four dark chocolates, four guacamoles, four crisps and four ice teas collected from 436 consumers divided in six groups. Each group of consumers has tested all products using only one sensory evaluation method among: Temporal Dominance of Sensations (TDS, n=70), Temporal Check-All-That-Apply (TCATA, n=73), Attack-Evolution-Finish (AEF) dominance (n=74), AEF applicability (n=75), Free Comment Attack-Evolution-Finish (FC-AEF) dominance (n=72) and FC-AEF applicability (n=72). Each consumer evaluated all the products: guacamoles and ice tea were evaluated in the lab in one session; chocolates and crisps were evaluated at home in two separate sessions. Within each product category, one sample has been replicated. The consumers started with product descriptions, then they gave a hedonic score, and after having tasted all the products related to a same category, they answered questions about product complexity and difficulty of the task. Consumer information included in the dataset is sex, age and frequency of consumption of each product category. This dataset is unique as it addresses several temporal methods applied on four product categories with different textures and levels of complexity. Thus, it could be very useful for the sensometric community to compare the different methods and their parameters: dominance vs. applicability, periods vs. continuous time, simultaneous vs. retrospective measures, list of terms vs. free comment.

1. Specifications table

Subject	Food science
Specific subject area	Sensory analysis. Temporal methods.
Type of data	Tables Figures Questionnaire
How the data were acquired	Sensory data were acquired from six panels of consumers (436 consumers in total) at lab and at home using TimeSens© software 2.0 (INRAE, Dijon, France).
Data format	Raw Analysed
Description of data collection	<p>The consumers were randomly assigned (with the constraint of balance in gender and age between groups) to one of the six panels, each panel using a different method to describe the temporal perception of the products: Temporal Dominance of Sensations – TDS, Temporal Check All That apply – TCATA, Attack-Evolution Finish dominance – AEF-D, AEF applicability – AEF-A, Free Comment AEF dominance – FC-AEF-D, FC-AEF-applicability – FC-AEF-A</p> <p>The four evaluated product categories (including various commercial brands) varied in composition, texture and sensory complexity. Crisps (solid) varied in fat and salt content; ice teas (liquid) varied in sugar/sweetener content, in flavour and variety of tea; guacamoles (semi-solid) varied in avocado and fat contents; chocolates (solid) varied in cocoa origin and content. Guacamoles and ice teas were evaluated at lab, crisps and chocolates at home. Hedonic data have been rated using 9-points discrete scales. The difficulty of the task has been scored on a 0-10 VAS scale. The items of sensory complexity of each product category have been scored on 0-10 VAS scales.</p>
Data source location	City/Town/Region: Dijon and vicinity Country: France
Data accessibility	The raw data, provided as a Microsoft Excel Worksheet, are available on the Mendeley data open-access research data repository. Repository name: Mendeley data Data identification number: 10.17632/fshtbhffth.1 Direct URL to data: https://data.mendeley.com/datasets/fshtbhffth/1
Related research article	

2. Value of the data

- These data are useful because they enable the comparison of six temporal methods applied on four product categories varying in sensory complexity, textures and compositions.
- The sensometric community can benefit from these data to compare different concepts: dominance vs. applicability, periods vs. continuous time, simultaneous vs. retrospective measures, list of terms vs. free comment.
- These data can be reused to compare and document the sensory methods performances, to develop new statistical analyses or to study the links between perception, liking and complexity.

3. Objective

This dataset has been generated in order to compare the six sensory methods on different product spaces varying in complexity and sensory differences. It will serve as “material and methods” for several articles to come that will study temporal resolution, performances and drivers of liking. The ultimate objective is to conclude on the most appropriate method to use with consumers based on the type of product, location and expected level of performances (discrimination, repeatability, reproducibility).

4. Data description

The dataset is provided as an Excel file (.xlsx) including 11 sheets.

The sheet “Consumer” provides information about the recruited consumers.

- “Panel” is the panel to which the consumer has been randomly assigned (TDS, TCATA, AEF_D, AEF_A, FC_AEF_D, FC_AEF_A).
- “Consumer” is the 3-character code of the consumer.
- “Gender” is the gender reported by the consumer (M for male or F for female).
- “Age” is the age range reported by the consumer (18_30: from 18 to 30 years old, 31_45: from 31 to 45 years old, 46_64: from 46 to 64 years old).
- “Consumption_IceTea”, “Consumption_Guacamole”, “Consumption_Chocolate”, “Consumption_Crisp” are the frequencies of consumptions of each product category (ice teas, guacamoles, dark chocolates, crisps) reported by the consumers (never, less than once a month, at least once a month, at least once a week).
- “Panel”, “Consumer” and “ProductCategory” columns are reported in each sheet following this one.

The sheets “TDS” and “TCATA” contain the temporal descriptions reported by the consumers of the panels of the same name.

- “Product” is the identifier of the product (character).
- “Time” is the time of each click on the attribute in seconds (numeric).
- “Attribute” is the code of the attribute (character).
- “Score” is 1 if “Attribute” has been considered dominant (TDS) or applicable (TCATA) by “Consumer” for “Product” during “Period” (numeric). For TCATA, an attribute remains applicable until the end or until deselected, in this case a new entry with score=0 is recorded.

The sheets “AEF_D” and “AEF_A” contain the temporal descriptions reported by the consumers of the panels of the same name.

- “Product” is the identifier of the product (character).
- “Period” is the identifier of the period (A: attack, E: evolution, F: Finish).
- “Attribute” is the code of the attribute (character).
- “Score” is 1 if “Attribute” has been considered dominant (AEF_D) or applicable (AEF_A) by “Consumer” for “Product” during “Period”, 0 otherwise (numeric).

The sheets “FC_AEF_D” and “FC_AEF_A” contain the temporal descriptions reported by the consumers of the panels of the same name.

- “Product” is the identifier of the product (character).
- “Period” is the identifier of the period (A: attack, E: evolution, F: Finish).
- “FrenchRawDescription” is the free comment reported by the consumer (free text, in French).
- “EnglishRawDescription” is the English translation of “FrenchRawDescription” made using deepL translator (<https://www.deepl.com/translator>) and checked by the authors of this article.

The sheet “Duration” contains the durations of tasting of each “Product” by each “Consumer” from each “Panel”.

- “Duration” is the duration from the click on the start button to the click on the stop button, in seconds (numeric).

The sheet “Liking” contains the liking scores reported for each “Product” by each “Consumer” from each “Panel”.

- “Liking” is the value rated on a discrete scale (numeric, between 1 and 9).

The sheet “Complexity” contains the scores of the different items of the complexity questionnaire reported for each “ProductCategory” by each “Consumer” of each “Panel”.

- “Attribute” is the code of the item (IntensityOfDifferences, Familiarity, NumberOfSensations, EaseOfIdentification, Harmony, Balance, Persistence, Power, Complexity).
- “Score” is the score on the structured scale (numeric, between 0 and 10, precision of 0.01).

The sheet “Difficulty” contains the scores of difficulty of the evaluation task reported for each context by each “Consumer” from each “Panel”.

- “Context” is the location of the measure (lab or home).
- “Score” is the score on the structured scale (numeric, between 0 and 10, precision of 0.01).

Table 26 summarizes the main characteristics of the six temporal methods.

Table 27 describes the product categories, codes and composition.

Table 28 reports the averaged scores of the items of the complexity questionnaire, by product category.

Table 29 summarizes the attributes used with the TDS, TCATA, AEF-A and AEF-D methods, by product category.

Table 30 reports the number of consumers having evaluated each product category, by panel.

Table 31 summarizes the individual characteristics of the consumers, by panel.

Figure 30 is the Principal Component Analysis of the averaged scores of the items of the complexity questionnaire by product category.

Figure 31 is the experimental procedure chart.

Questionnaire includes commented screenshots of the online questionnaire (TimeSens V2 web app). It has been translated from French to English.

5. Experimental design, materials and methods

5.1. Temporal methods

Six temporal methods were compared: Temporal Dominance of Sensations - TDS (Pineau et al., 2009); Temporal Check-All-That-Apply - TCATA (Castura et al., 2016); Attack-Evolution-Finish - AEF (Visalli et al., 2020) called here AEF-D (D for dominance); AEF-A (adaptation of AEF-D, the reported attributes being the applicable ones instead of the dominant ones); Free Comment Attack-Evolution-Finish dominance - FC-AEF-D (an adaptation of FC-AEF, Mahieu et al., 2020b), the reported attributes being the dominant ones instead of the applicable ones); FC-AEF, called here FC-AEF-A (A for Applicability).

For AEF-D and FC-AEF-D, no definition was given for dominance, but the forced choice was supposed to imply the dominance concept.

Table 26. Summary of the main characteristics of the six temporal methods.

Method	Moment of measure	Choice of attributes	Reported attributes	Temporal resolution
TDS	During tasting	Predefined list	Dominant ones, one at a time	Continuous
TCATA	During tasting	Predefined list	Applicable ones, zero, one or several at a time	Continuous
AEF-D	After tasting	Predefined list	Dominant ones, one at a time	Periods
AEF-A	After tasting	Predefined list	Applicable ones, zero, one or several at a time	Periods
FC-AEF-D	After tasting	Free comment	Dominant ones, one at a time	Periods
FC-AEF-A	After tasting	Free comment	Applicable ones, zero, one or several at a time	Periods

5.2. Product categories and samples

In order to get more generalizable conclusions, the methods were compared on different product families varying in composition, sensory modalities and sensory complexity. The chosen products were commercial products accepted by a majority of consumers and easy to prepare for the experimenters. Bibliographical research on the products used in previous studies implementing these methods and pre-tests led to the selection of four product categories, each containing four variants. Crisps (solid, portion size: 3 g) varied in fat and salt, ice teas (liquid, portion size: 20 ml) varied in sugar/sweetener, flavour and variety of tea, guacamoles (semi-solid, portion size: 7 g) varied in avocado content and fat, and dark chocolates (solid, portion size: 10 g) varied in cocoa origin and content. To assess for individual and panel repeatability, one variant was replicated inside each product category.

The ice teas and guacamoles were stored in the refrigerator at 4°C until tasting. They were taken out of the refrigerator a few minutes before the test so as not to be too fresh when tasted.

Table 27. Product categories, codes and composition (as reported on the packaging).

Product category	Codes	Composition
Crisp	C1, C1_rep (replicate of C1)	34 g fat, 1.3 g salt
Crisp	C2	23.9 g fat, 1.52 g salt
Crisp	C3	29 g fat, 1 g salt (sea salt)
Crisp	C4	34 g fat, 0.10 g salt
Guacamole	G1, G1_rep (replicate of G1)	92 % avocado, 16 g fat
Guacamole	G2	13 % avocado, 9.5 g fat
Guacamole	G3	90 % avocado, 14.6 g fat
Guacamole	G4	95 % avocado, 18 g fat
IceTea	IT1	4.7 g sugar, black tea, white peach
IceTea	IT2	4.5 g sugar, white tea, peach and rosemary
IceTea	IT3, IT3_rep (replicate of IT3)	4.3 g sugar, sweeteners, black tea, peach
IceTea	IT4	0 g sugar, sweeteners, black tea, peach
Chocolate	CH1	85 % cocoa, origin Madagascar
Chocolate	CH2	80 % cocoa, origin Equator
Chocolate	CH3	70 % cocoa, origin Peru
Chocolate	CH4, CH4_rep (replicate of CH4)	74 % cocoa, origin Côte d'Ivoire

The samples were evaluated according to a between-subjects design (the treatment being the temporal method), with products evaluated in a fixed order, which is quite unconventional in sensory analysis. Indeed, this study was purely methodological, and one of the objectives was to compare temporal methods (not products) regarding to individual differences. In this specific case, it is common to fix the order of testing to minimize the variance introduced by different orders of presentation across the subjects (Bell, 2013; Humes et al., 2013). By doing so, it must be remembered before drawing any product dependent conclusions that an order effect may have affected the product comparison. Since there was no indication of which order would minimize this potential product order effect, the presentation rank of each product was determined randomly, with the exception of the second evaluation of the repeated sample which was always presented at the fifth rank. The position of the product in the presentation design was indicated in the code (example: G1 was served at position 1). The samples were labelled using random 3-digits codes.

Table 28 summarizes the results of the complexity questionnaire.

Table 28. Averaged scores of the items of the complexity questionnaire by product category.

	Balance	Complexity	Ease of identification	Familiarity	Harmony	Intensity of differences	Number of sensations	Persistence	Power
Chocolate	5.58	5.86	5.48	6.15	5.77	5.81	5.23	6.70	6.82
Crisp	5.38	4.47	4.79	6.66	5.82	6.38	4.87	5.58	5.60
Guacamole	4.92	6.01	5.23	5.37	5.69	7.56	6.77	6.99	7.18
IceTea	4.96	5.37	5.19	5.84	5.39	7.36	5.62	6.00	6.51

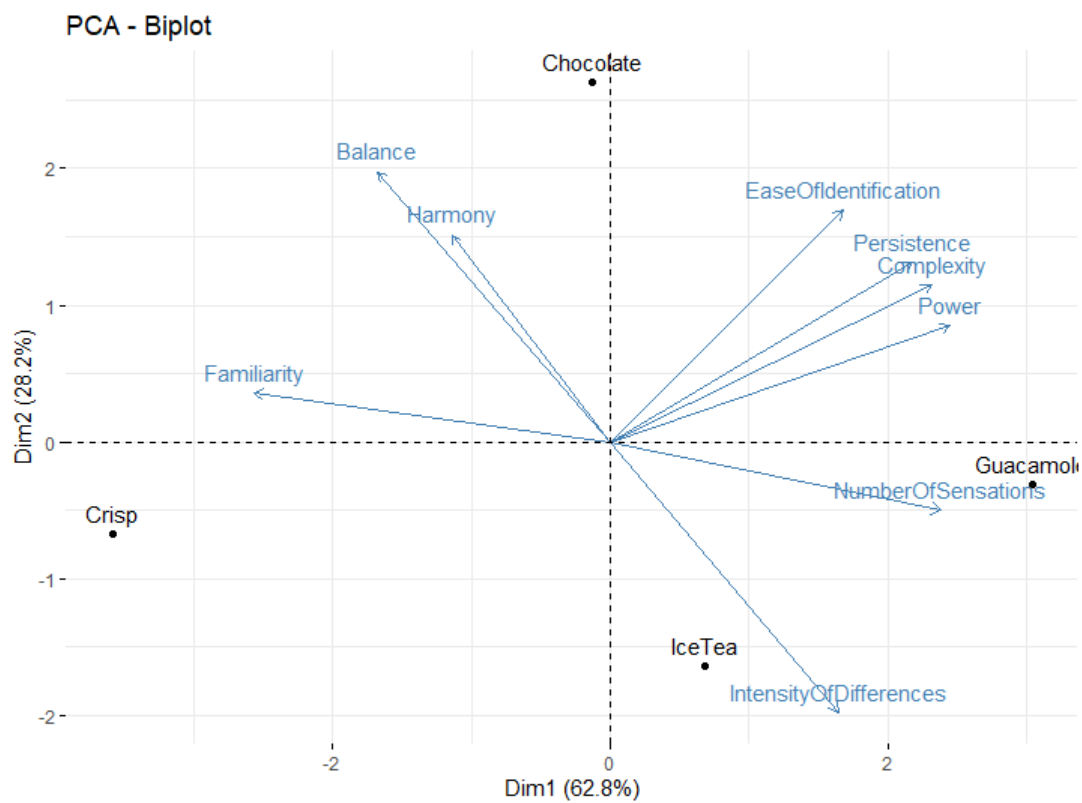


Figure 30. PCA of the averaged scores of the items of the complexity questionnaire by product category.

Figure 30 was plotted using R 4.0.1 (R Core team, 2017) and package FactoMineR (Lê et al., 2008). It shows the correlations between the different dimensions of the complexity questionnaire.

5.3. Attributes

A pre-test carried out by an internal panel of six persons familiar with temporal methods was carried out to select the descriptors for the TDS, TCATA, AEF-D and AEF-A methods.

Table 29. Attributes used with the TDS, TCATA, AEF-D and AEF-A methods, by product category.

Product category	Sensory modality	French word	English translation
IceTea, Guacamole, Chocolate	Basic taste	Acide	Sour
IceTea, Guacamole, Chocolate	Basic taste	Sucré	Sweet
IceTea, Guacamole, Chocolate	Basic taste	Amer	Bitter
IceTea	Mouthfeel	Astringent/Râpeux	Astringent/Rough
IceTea	Flavour	Pêche	Peach
IceTea, Chocolate	Flavour	Chimique	Artificial
IceTea	Flavour	Thé	Tea
IceTea	Texture	Aqueux/Dilué	Watery/Diluted
Guacamole	Texture	Onctueux/Gras	Smooth/Fat
Guacamole	Texture	Fluide	Fluid
Guacamole	Texture	Épais	Thick
Guacamole	Flavour	Citronné	Lemon
Guacamole	Flavour	Avocat	Avocado
Guacamole	Flavour	Pimenté	Spicy
Guacamole	Flavour	Oignon/Échalote	Onion/Shallot
Guacamole	Flavour	Tomate	Tomato
Guacamole	Flavour	Herbes aromatiques	Aromatic/Herb
Crisp	Texture	Craquant/Dur	Crackly/Hard
Crisp	Texture	Croustillant	Crispy
Crisp	Texture	Collant/Pâteux	Sticky/Pasty
Crisp	Texture	Fondant	Melting
Crisp	Basic taste	Salé	Salty
Crisp	Flavour/Mouthfeel	Gras	Fat
Crisp	Flavour	Pomme de terre	Potato
Crisp	Flavour	Grillé	Roasted
Crisp	Flavour	Fade	Bland
Chocolate	Texture	Sec/Poudreux	Dry/Powdery
Chocolate	Texture	Fondant/Gras	Melting/Fat
Chocolate	Texture	Collant	Sticky
Chocolate	Mouthfeel	Astringent	Astringent
Chocolate	Flavour	Cacao	Cocoa
Chocolate	Flavour	Fruité	Fruity
Chocolate	Flavour	Floral	Floral
Chocolate	Flavour	Boisé/Torréfié	Woody/Roasted

The descriptors were presented in a random order on the screen but this order was constant for each consumer across the evaluations of a same product category. Due to a technical problem, only nine attributes over twelve have been presented to the AEF-A panel on the Chocolate product category.

5.4. Consumers

At the moment of the recruitment, the objective was to have a minimum size of 64 consumers for each panel, or 384 subjects in total. Taking into consideration the possible withdrawals, 504 consumers were preselected 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—n° d'autorisation 1148039). The selection criteria included gender, age and frequency of consumption of the product categories. Restrictions related to the health context related to COVID-19 were added to the usual restrictions. People with an allergy, people on a restrictive diet, people considered vulnerable and pregnant women were therefore excluded. A total of 436 people participated in this study. They were compensated for their participation with vouchers worth 15 euros.

The consumers were randomly assigned to one of the six panels with a constraint of balance in gender and age between panels. Each panel used a different temporal method to describe the temporal evolution of the samples from each product category over three sessions, one at lab and two at home, on different days.

Table 30. Number of consumers having evaluated each product category, by panel.

Context	Product category	Panel TDS	Panel TCATA	Panel AEF-D	Panel AEF-A	Panel FC-AEF-D	Panel FC-AEF-A
Lab, 1 session	IceTea	70	73	74	75	72	72
	Guacamole						
Home, 2 days	Crisp	70	72	74	75	66	70
	Chocolate	68	72	71	75	68	70

Table 31. Individual characteristics of the consumers, by panel. Cells indicate percentages.

Individual characteristics	Panel TDS	Panel TCATA	Panel AEF-D	Panel AEF-A	Panel FC-AEF-D	Panel FC-AEF-A
Age: 18-30	20.00	21.92	21.62	14.67	20.83	18.06
Age: 31-45	35.71	31.51	36.49	42.67	41.67	37.50
Age: 46-64	44.29	46.58	41.89	42.67	37.50	44.44
Gender: male	68.57	72.60	70.27	70.67	66.67	68.06
Gender: female	31.43	27.40	29.73	29.33	33.33	31.94
IceTea: at least once a month	42.86	26.03	37.84	41.33	45.83	38.89
IceTea: at least once a week	25.71	36.99	28.38	21.33	18.06	22.22
IceTea: less than once a month	27.14	36.99	31.08	36.00	30.56	34.72
IceTea: never	4.29	0.00	2.70	1.33	5.56	4.17
Guacamole: at least once a month	57.14	47.95	43.24	41.33	47.22	38.89
Guacamole: at least once a week	14.29	10.96	9.46	13.33	9.72	12.50
Guacamole: less than once a month	28.57	36.99	43.24	44.00	41.67	47.22
Guacamole: never	0.00	4.11	4.05	1.33	1.39	1.39
Chocolate: at least once a month	21.43	12.33	24.32	16.00	13.89	12.50
Chocolate: at least once a week	71.43	79.45	59.46	70.67	70.83	75.00
Chocolate: less than once a month	2.86	6.85	12.16	10.67	11.11	4.17
Chocolate: never	1.43	0.00	0.00	2.67	1.39	2.78
Crisp: at least once a month	32.86	38.36	40.54	54.67	40.28	44.44
Crisp: at least once a week	48.57	45.21	36.49	32.00	40.28	26.39
Crisp: less than once a month	14.29	15.07	22.97	12.00	16.67	20.83
Crisp: never	0.00	0.00	0.00	1.33	0.00	0.00

5.5. Experimental procedure

Collective briefing (session 1)

To start, all consumers were invited to a collective briefing at lab in groups of 16. The temporal method was explained and the attributes were presented. A demonstration of the task was done by the panel leader by video projecting the screens (see attached questionnaire, TDS: screens 1-27, TCATA: screens 28-54, AEF-D/A: screens 55-83, FC-AEF-D/A: screens 84-112) of the software (TimeSens V2, INRAE, 2022). Then, the consumers were encouraged to ask questions about the task before taking place in individual booths.

Instructions for temporal description

The instructions, depending on the temporal method, were reminded on the screen. The screen was displayed during a minimum time of 30 seconds to ensure the consumers read the instructions. The instructions were quite similar for all product categories, below are instructions for IceTea.

TDS instructions (screen 3): “For each ice tea, you will proceed as follows. You will take a sip, while simultaneously pressing the “Mouthing” button. A list of buttons will be displayed on the screen. Throughout the tasting of the sip, as soon as you perceive a dominant sensation, you will have to press the button corresponding to this sensation. Some sensations may never be selected, others may be selected multiple times during the tasting of the sip. You will continue to indicate the sensations perceived after swallowing the sip. When you no longer perceive anything, you will click on “I don’t perceive anything anymore” button. You should only taste one sip of each sample of ice tea. Familiarize yourself with the sensations available and their location on the screen before putting the product in your mouth.”

TCATA instructions (screen 30): “For each ice tea, you will proceed as follows. You will take a sip, while simultaneously pressing the “Mouthing” button. A list of checkboxes will be displayed on the screen. Throughout the tasting of the sip, as soon as you perceive a sensation, you will have to check the checkbox corresponding to this sensation. You will need to uncheck the checkbox as soon as you no longer perceive this sensation. Some sensations may never be selected, others may be selected multiple times during the tasting of the sip. You will continue to check the attributes perceived and uncheck the attributes no longer perceived after swallowing the sip. When you no longer perceive anything, you will click on the “I don’t perceive anything anymore” button. You should taste only one sip of each sample of ice tea. Familiarize yourself with the sensations available and their location on the screen before putting the product in your mouth.”

AEF (D and A) instructions (screen 57): “For each ice tea, you will proceed as follows. You will take a sip, while simultaneously pressing the “Mouthing” button. When you no longer perceive anything, you will press the “Next” button. At this moment, we will ask you to describe your perception by choosing, from a list of terms, a sensation [AEF-D] / one or several sensations [AEF-A] for each period of your perception: beginning, middle and end. Some sensations may never be selected, others may be selected in several periods. An example is given to you on the following page. You should taste only one sip of each sample of ice tea.

FC-AEF (D and A) instructions (screen 86): “For each ice tea, you will proceed as follows. You will take a sip, while simultaneously pressing the “Mouthing” button. When you no longer perceive anything, you will press the “Next” button. At this moment, we will ask you to describe, using your own words, the sensation [FC-AEF-D] / sensation(s) [FC-AEF-A] you

experienced for each period of your perception: beginning, middle and end. An example is given to you on the following page. Use only words, don't make sentences. Compound words and expressions are allowed. Example: "long in the mouth". You should taste only one sip of each sample of ice tea."

Tasting, descriptive and hedonic evaluation of the five ice tea samples

The consumers had to evaluate the products of category IceTea under red light. They first had to check the sample code (TDS: screen 4, TCATA: screen 31, AEF-D/A: screen 59, FC-AEF-D/A: screen 88). Consumers of AEF (screen 60) and FC-AEF (screen 89) panels started with a screen with no attribute, allowing to record the duration of the tasting. The measurement screens were displayed just after, depending on the temporal method (TDS: screen 5, TCATA: screen 32, AEF-D/A: screen 61, FC-AEF-D/A: screen 90). After having described their perception, whatever the panel, the consumers had to rate their preference for the tasted sample on a 9-point discrete scale (TDS: screen 6, TCATA: screen 33, AEF-D/A: screen 62, FC-AEF-D/A: screen 91). After that, a 30-second forced break (TDS: screen 7, TCATA: screen 34, AEF-D/A: screen 63, FC-AEF-D/A: screen 92) was imposed, inviting the consumers to rinse their mouth with mineral water. The procedure was repeated for the three other samples and the replicated one.

Evaluation of the complexity of the ice tea product category

After the evaluation of the five samples, the consumers had to evaluate the complexity of the product category, using an adaptation of the questionnaire of (P. Schlich et al., 2015). Nine items were evaluated over four consecutive screens (TDS: screens 9-12, TCATA: screens 36-39, AEF-D/A: screens 65-68, FC-AEF-D/A: screens 94-97) using structured scales. Items included, in this order: "Intensity of differences", "Familiarity", "Number of perceived sensations" "Easiness of identification of sensations", "Harmony between sensations", "Balance", "Persistence", "Power" and "Overall complexity".

Evaluation of the five guacamole samples

After a 5-minute forced break (TDS: screen 13, TCATA: screen 40, AEF-D/A: screen 69, FC-AEF-D/A: screen 98), the five guacamole samples were evaluated following the procedure described in sections 5.3 and 5.4.

Participants characteristics and frequency of consumption

Participants characteristics and frequency of consumption were asked (TDS: screen 16, TCATA: screen 43, AEF-D/A: screen 72, FC-AEF-D/A: screen 101) in order to check consistency with answers collected during the recruitment phase.

Evaluation of the difficulty of the task at lab

To end the lab session, the consumers had to evaluate the difficulty of the temporal description task using a structured scale (TDS: screen 17, TCATA: screen 44, AEF-D/A: screen 73, FC-AEF-D/A: screen 102).

End of lab session and preparation of home session

The duration of the session was approximately 45 minutes. After that, each consumer had to take home samples from the two other product categories: Chocolates, and Crisps. The samples of crisps were stored in individual disposable plastic cups with lids. The samples of chocolate were packed into aluminium foils. All the samples were put into a bag and the consumers were asked to keep the bag at room temperature.

The consumers were informed that they would receive two separated e-mails inviting them to connect to the internet session using a browser (TDS: screen 18, TCATA: screen 45, AEF-D/A: screen 74, FC-AEF-D/A: screen 103). There was one mail for each product category, and the sessions were still designed using TimeSens V2. The consumers were instructed to evaluate all the samples of a product category on a same day, but the two categories on different days.

Evaluation of the five samples of crisps, at home (session 2)

The consumers received the first mail the same day of the tasting in lab. They were instructed to do the session 2 in the same day, or if unable to do so, the next morning. In session 2, they had to evaluate the samples of crisps following the procedure described in sections 5.3, 5.4, and 5.6.

Evaluation of the five samples of chocolate, at home (session 3)

A second email was sent to the consumers the day after they completed session 2. In session 3, they had to evaluate the samples of chocolate following the procedure described in sections 5.3, 5.4, and 5.6. Then, a final question was asked about the usefulness of having done the

task at lab before doing the test at home (TDS: screen 27, TCATA: screen 54, AEF-D/A: screen 83, FC-AEF-D/A: screen 112). They had to answer on a structured scale.

The whole procedure was summarized in Figure 31 below.

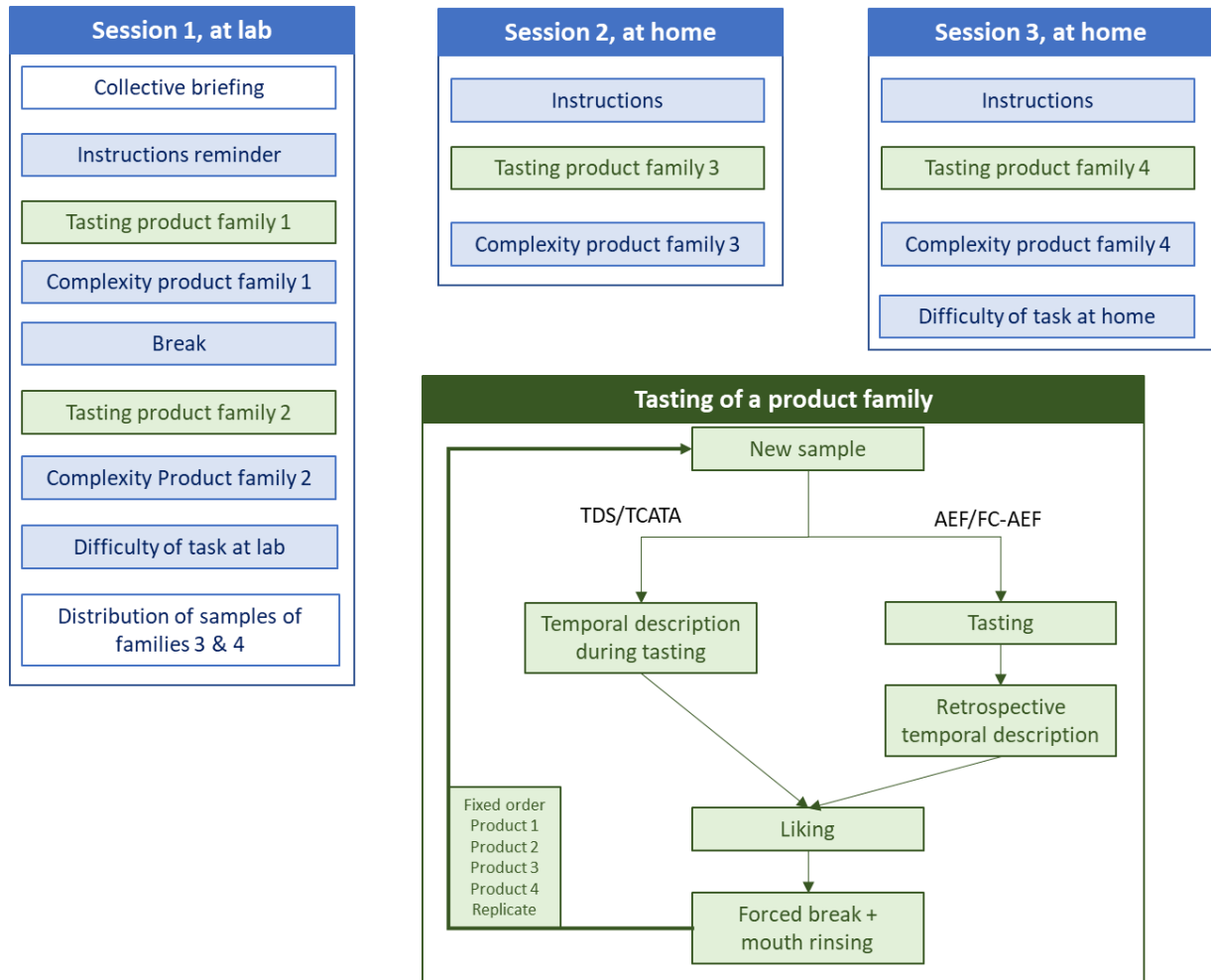


Figure 31. Experimental procedure chart. Steps in blue and green boxes were related to the TimeSens session. Steps in green boxes were repeated for each sample of each product category.

II.5. A dataset on sensory evaluation of gustometer controlled stimuli with TDS, TCATA and AEF

This article was *published in Data in brief*:
<https://www.doi.org/10.1016/j.dib.2023.109271>

N. Béno: Conceptualization, Data Curation, Investigation, Methodology, Project administration, Resources, Validation, Writing - Original Draft. L. Nicolle: Data Curation, Investigation, Methodology, Resources, Writing - Review & Editing. M. Visalli: Conceptualization, Funding acquisition, Methodology, Software, Supervision, Visualization, Writing - Review & Editing.

This paper describes data on the consumer sensory perception of liquid mixtures including sapid and aromatic compounds. A total of 149 consumers participated in this study. They were randomly assigned to one of three panels. Each panel used a different temporal sensory evaluation method among Temporal Dominance of Sensation (TDS, n=50), Temporal Check-All-That-Apply (TCATA, n=50) and Attack-Evolution-Finish Rate-All-That-Apply (AEF-RATA, n=49) to evaluate solutions delivered by a gustometer (Burghart GU002). First, four simple solutions (composed of a single compound) were delivered to the consumers to evaluate their recognition ability using free comment. Second, eighteen complex solutions (composed of two to five compounds varying in their sequence, intensity and duration of stimulation) were delivered to the consumers to evaluate their ability to use the three temporal evaluation methods. The compounds included sodium chloride (“salty”), saccharose (“sweet”), citric acid (“acid”), citral (“lemon”) and basil hydrosol (“basil”). The data were used to assess the validity and reliability of the temporal sensory methods in an article entitled “Assessment of the validity and reliability of temporal sensory evaluation methods used with consumers on controlled stimuli delivered by a gustometer”. The data could be reused by researchers interested in studying the effect of interactions between sapid and aromatic compounds on perception.

1. Specifications table

Table 32. Specifications table

Subject	Food science
Specific subject area	Temporal sensory evaluation
Type of data	Table Figure
How the data were acquired	<p>Three consumer panels evaluated liquid solutions using each a different temporal sensory evaluation method among Temporal Dominance of Sensation (TDS (N Pineau et al., 2009), n=50), Temporal Check-All-That-Apply (TCATA (John C Castura et al., 2016), n=50) and Attack-Evolution-Finish Rate-All-That-Apply (AEF-RATA, n=49), an adaptation of AEF-Applicability (Visalli, Wakihira, et al., 2022b).</p> <p>Experiences were carried out individually in the human olfaction/taste laboratory of the ChemoSens platform.</p> <p>The solutions were delivered by a gustometer Burghart GU002 at a constant flow rate of 350 μL/s.</p> <p>The data were recorded using the TimeSens© (INRAE, 2022) software, version 2.0.</p>
Data format	<p>Tables in raw format (XLSX file)</p> <p>Videos (MP4)</p> <p>Questionnaire (PDF)</p>
Description of data collection	<p>Recognition task</p> <p>Four single iso-intense compound solutions were delivered to the consumers: two sapid compounds over three (sodium chloride, saccharose, citric acid), one aromatic compound over two (citral, basil hydrosol), plus one replicate. The solutions were presented according to a William's Latin square design balanced at the panel level. The solutions were delivered during eight seconds. The durations of perception were recorded, and the consumers had to self-report the sensation(s) they perceived using free comment.</p> <p>Temporal perception task</p> <p>Eighteen (14 different and four replicated) multi-compound solutions were delivered to the consumers in the same order. They were composed of two to five compounds (same compounds as for the recognition task), varying in their sequence (with or without overlap), intensity (three iso intense concentration levels: weak, medium, and strong) and duration of stimulation. Each sequence lasted 30 seconds. The same list of eight attributes was used in TDS, TCATA and AEF-RATA. This list includes five attributes corresponding to the compounds: sweet, salty, acid, lemon and basil; and three distractors: bitter, liquorice, and mint. For TDS and TCATA, the record of the perception started when the consumers clicked on an attribute. The times and durations of perception of dominance (TDS) and applicability (TCATA) were thus recorded. For AEF-RATA, the consumers had to retrospectively rate the intensity they perceived (weak, medium, strong) for each applicable attribute and for each of three periods: "at the beginning", "after a few seconds", "at the end".</p>
Data source location	<p>Institution: INRAE</p> <p>City/Town/Region: Dijon</p> <p>Country: France</p>
Data accessibility	<p>Repository name: Mendeley data</p> <p>Data identification number: 10.17632/3j6h7mrxnf.1</p> <p>Direct URL to data: https://data.mendeley.com/datasets/3j6h7mrxnf/1</p>
Related research article	

2. Value of the data

- These data are useful because they allow a better understanding of what is actually measured with temporal evaluation methods by making it possible to compare the declarative results relating the sensory perception of consumers with the chemical reality of the controlled stimuli delivered using a gustometer.
- The food science and sensometric community can benefit from these data to assess the construct and criterion validity of concepts related to sensory evaluation methods (e.g. dominance vs. applicability) and their derived measurements (e.g. intensities vs. durations vs. citation rates, periods vs. continuous time, etc.)
- These data can also be reused to compare and document the performances of temporal sensory evaluation methods (e.g. repeatability, discrimination) or to develop new statistical analyses better suited to the specific nature of temporal measurements.
- These data can also be reused by researchers interested in studying the effects of congruent and incongruent interactions between sapid and aromatic compounds on the sensory perception of consumers (e.g. when acid and lemon compounds are delivered at the same moment).

3. Objective

These last years, numerous sensory evaluation methods have been developed to capture the dynamic of perception during the tasting of food products. However, “sensory reality” remains unknown, thus it is difficult for sensory scientists to determine which method is more valid. The data were collected to investigate how the perception of controlled temporal stimuli delivered using a gustometer are transcribed by consumers using three different sensory evaluation methods: Temporal Dominance of Sensations (TDS), Temporal Check-All-That-Apply (TCATA), and Attack-Evolution-Finish Rate-All-That-Apply (AEF-RATA). The ultimate objective was to confront the results of the methods to the chemical reality of the stimuli to document the validity and reliability of the methods.

4. Data description

The dataset is provided as an Excel file (data.xlsx) including six sheets.

Consumer provides information about the participants.

- "Panel": name of the panel to which the participant has been assigned. The name of the panel corresponds to the name of the method used for the temporal sensory evaluation: TDS, TCATA, and AEF-RATA.
- "Consumer": unique code of the participant.
- "Gender": gender of the participant (Male or Female).
- "Age": age of the participant.

Stimuli provides information about the sequences delivered in the recognition task and in the temporal perception task.

- "Stimulus": code of the stimulus. Stimuli with the following codes have been delivered during the recognition task: Acid, Sweet, Salty, Lemon and Basil. Other codes correspond to stimuli delivered during the temporal perception task.
- "Attribute" corresponds to the attribute related to the delivered compound (salty for sodium chloride; sweet for saccharose; acid for citric acid; lemon for citral, and basil for basil hydrosol).
- "Time": beginning time of delivery of "Attribute" expressed in seconds.
- "Quantity": quantity delivered at the different concentrations of the intensity levels (10% for weak; 50% for medium, and 90% for strong).

RecognitionTask provides data related to the recognition task.

- "Panel", "Consumer", "Stimulus": see above.
- "FrenchDescription": free comment (in French) reported by the participants to describe the stimulus.
- "Attributes": sensory attributes extracted from free comment and translated to English by the experimenters.
- "Result": categorization of the result of the recognition task performed by the experimenters in five levels: not identified (NotIdentified), approximately identified with the need for several attributes (Identified_Approximate_NotUnique), approximately identified with the need for a single attribute (Identified_Approximate_Unique), identified with the need for several attributes (Identified_Exact_NotUnique) and identified with the need for a single attribute (Identified_Exact_Unique).
- "Duration": duration of perception of the stimulus, in seconds.

TDS, *TCATA* and *AEF-RATA* provide data collected with the corresponding temporal evaluation methods during the temporal perception task.

- "Panel", "Consumer", "Stimulus": see above.
- "Time": time (in seconds) of click on the attribute (TDS and TCATA)

- “Period”: period (A for Attack, E for Evolution, F for finish) during which the attribute was retrospectively declared applicable (AEF-RATA).
- “Attribute”: attribute clicked by the participant (acid, basil, bitter, lemon, liquorice, mint, salty or sweet) + end of perception (“stop”, TDS and TCATA).
- “Score”: for TDS: dominant attributes (always 1), for TCATA: applicable attributes (1: start of applicability, 0: end of applicability), for AEF-RATA: perceived intensity for applicable attributes (1: weak, 2: medium, 3: strong).

Figure 32 shows a picture of the gustometer.

Figure 33 shows the screens displayed to the consumers.

Table 33 describes the composition of the three consumer panels.

Table 34 describes the sapid and aromatic compounds used in the stimuli.

5. Experimental design, materials and methods

5.1. Participants

A total of 149 consumers (56 men and 93 women, between 21 and 65 years old) participated in this study. They were preselected from a population registered in the ChemoSens Platform's PanelSens database (declared to the relevant authority, Commission Nationale Informatique et Libertés – CNIL, authorization number 1148039). The inclusion conditions for participating in this study were as follows: being between 18 and 65 years old; not suffering from food or non-food allergies; not being pregnant or breastfeeding and not following a restrictive diet incompatible with the consumption of sugar or salt.

The purpose of the study was explained via an information sheet sent by email. The consumers had to fill out a written informed consent form. They were compensated for their participation (one session) with vouchers worth 10 euros.

The selected consumers were randomly assigned to one of three panels, with a constraint of balance in gender and age between panels. Each panel used a different temporal method among AEF-RATA (n=49), TCATA (n=50) and TDS (n=50).

Table 33. Composition of the consumer panels.

Panel	TDS	TCATA	AEF-RATA
Average age	46.8	46.2	44.4
Gender: male	21	16	19
Gender: female	29	34	30

5.2. Stimuli

Different stimuli were delivered to the consumers using a gustometer model Burghart GU002.



Figure 32. Gustometer model Burghart GU002.

This device consists of five syringes (2.5 ml) containing the sapid and aromatic compound solutions, and two syringes (2.5 ml) containing water. The stimuli temperature is regulated at 35°C to avoid thermal irritations. Up to five compounds can be loaded, mixed, and diluted by the gustometer.

The selected compounds are described in Table 34. Their concentrations were chosen to be as close to iso-intensity as possible, based on internal pre-tests.

Controlled by a computer, the gustometer enables to deliver liquids in the form of pulses (70 μ l every 100 ms) through a tube (350 μ l/s flow rate). Up to 39 classes could be defined for the pulses, one class defining the selected syringe(s), the dilution of the compound, and

sequences defining the duration of the class(es) delivery. The total dilution of the compounds could not exceed 100% within a class. Three classes of concentration were chosen for the compounds: weak, medium, and strong, corresponding respectively to 10, 50 and 90 % of the concentrated solution.

As swallowing was made more difficult by the continuous flow of liquid, the stimulation durations should not exceed 30 seconds to limit the volume in the mouth to 10 ml and reduce the discomfort of consumers.

Table 34. Information about the sapid and aromatic compounds used in the stimuli, and corresponding sensory attributes.

Compound	Attribute	Raw formula	CAS Number	Lot number	Expiration date	Concentration
Sodium chloride	Salty	NaCl	7647-14-5	19100184/D	10-2024	25 g/l
Saccharose	Sweet	C ₁₂ H ₂₂ O ₁₁	57-50-1	71616173	-	250 g/l
Citric acid	Acidic	C ₆ H ₈ O ₇ x H ₂ O	5949-29-1	20070058/A	06-2023	10 g/l
Citral	Lemon	C ₁₀ H ₁₆ O	5392-40-5	MKCJ7159	03-2024	10 mg/l
Basil hydrosol (Plantago)	Basil	-	-	HY2120 HY2120	08-2022 08-2023	20 g/l

5.3. Data collection

The consumers participated in individual sessions of approximately one hour in the olfactometry lab of ChemoSens. They were installed in a booth and had to read and accept the condition of the study. The consumers' responses were collected using TimeSens© V2 software.

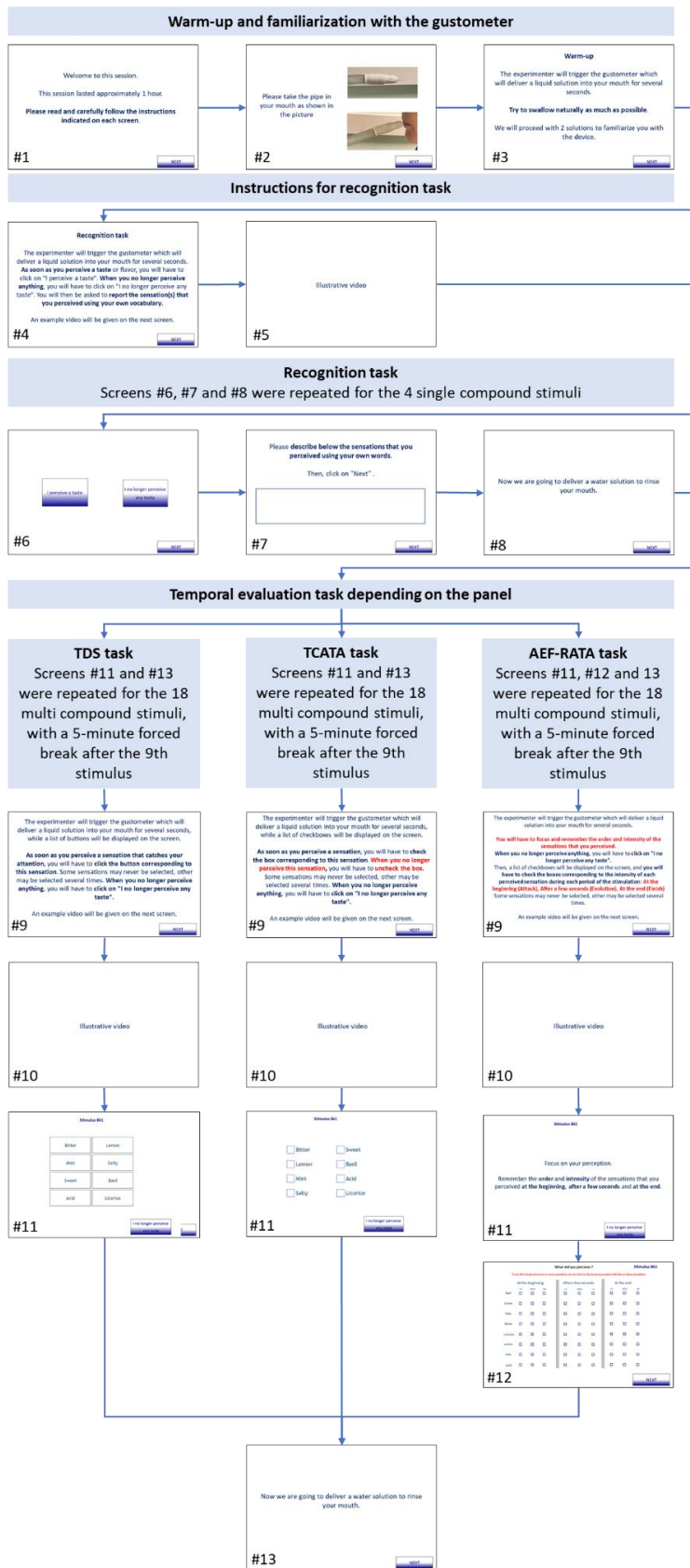


Figure 33. Screens displayed to the consumers.

Familiarization with the gustometer (screens #1-3)

The consumers were instructed about how the gustometer works and how to position their mouth (screen #2). They could ask all the questions they wanted to the experimenter who piloted the gustometer. Then, they were invited to experiment the stimulation with the gustometer (screen #3) so that they get used to the device, and in particular to swallow at the same time as they received a liquid flow under pressure (they were free to swallow whenever they wished). Two water solutions were delivered during 20 and 30 s.

Recognition task (screens #4-5)

Four medium-concentration single-compound solutions were delivered to the consumers, each for 16 seconds at a constant flow rate of 350 $\mu\text{l/s}$: water for 4 s, compound for 8 s, then water for 4 s. The order of presentation of the stimuli follows an incomplete William's Latin square design balanced at the panel level. Each consumer evaluated two sapid compounds over three (sodium chloride, saccharose, citric acid) and one aromatic compound over two (citral, basil hydrosol). The first or second solutions delivered was replicated last.

The instructions for the task were presented on the screen (screen #4), then a tutorial video was shown (screen #5). The durations of perception were recorded: the consumers had to click on a button when they started to perceive and when they no longer perceived anything (screen #6). Then, the consumers had to self-report the sensation(s) they perceived using free comment (screen #7). Between two stimulations, water was delivered by the gustometer during 10 seconds to clean the tubes and rinse the mouths of consumers (screen #8).

Temporal evaluation task (screens #9-12)

Eighteen multiple-compound solutions were delivered to the consumers (14 different and four replicated). The order of presentation was the same for all the consumers: S01, S07, S011, S04, S12, S14, S10, S02, S013, S03, S08, S09, S06, S05, S11_2, S02_2, S08_2, S14_2. The stimuli varied in number, sequence (with or without overlap), duration, and concentration of the compounds. Each sequence lasted 30 seconds and 10.5 ml were delivered at a constant flow rate of 350 $\mu\text{l/s}$.

The instructions for the task were presented on the screen (screen #9), then a tutorial video was shown (screen #10). Each panel used a different temporal sensory evaluation method. The same list of eight attributes was used in TDS, TCATA and AEF-RATA. This list includes five

attributes related to the compounds: sweet, salty, acid, lemon and basil; and three distractors: bitter, liquorice and mint. The order of presentation of the attributes was randomized between the consumers but remained the same for a consumer within the session. For TDS and TCATA (screen #11), the chronometer started when the consumers clicked on an attribute. The times and durations of perception of dominance (TDS) and applicability (TCATA) were recorded. For AEF-RATA (screen #12), the consumers had to retrospectively rate the intensity they perceived (weak, medium, strong) for each applicable attribute and for each of three periods: “at the beginning”, “after a few seconds”, “at the end”. Between two stimulations, water was delivered by the gustometer for 10 seconds to clean the tubes and the rinse mouths of consumers. A five-minute break was imposed after the 9th stimulus.

Chapter III.

Results

III.1. Concurrent vs. retrospective temporal sensory data collection

III.1.1. Context and aims

Concurrent data collection gives great importance on the times when sensory attributes are cited and the total durations of perception, but these times and durations can be very different from one subject to another. When this heterogeneity is too great, data can become complex to analyse. Thus, to simplify data analysis or interpretation, TDS and TCATA data are sometimes transformed into periods (see Section 1.3.2, part 3.4.1). One might wonder why not directly ask consumers to report their perception in this way. Empiric evidences from wine tasting suggest that when the duration of perception is not too long (which is the case for the majority of food products), it is possible to ask subjects to retrospectively report the sensations they perceived during specific time periods, immediately after the tasting. This was the basic idea that led to the development of a “new” method called Attack-Evolution-Finish (AEF), consisting in asking subjects immediately after the tasting the sensory attributes they perceived during three periods: the “attack” phase (first perception(s)), the “evolution” phase (in between attack and finish), and the “finish” phase (last perception(s)).

The lack of attention paid to temporal retrospective measurements in sensory analysis can probably be explained by the tacit assumption that as concurrent measurements are closer to immediate perception, they are more accurate. However, the fact that TDS and TCATA collect data in continuous time does not necessarily imply that the temporal information returned at the panel level is accurate. In other areas of research and in particular in psychology, the performances of retrospective and concurrent measurements were compared, without it being possible to conclude on the superiority of one or the other. Indeed, each mode of collection suffers from different biases due to the different cognitive processes involved.

Compared to TDS and TCATA, AEF was expected to present several advantages: (i) simplify data collection for consumers (the task is self-explanatory); (ii) reduce the noise inherent to concurrent continuous-time measurements (data are collected in three periods); (iii) generate data easier to analyse (there is no need to transform data prior to data analysis). AEF was designed to be versatile, and different variations of the method were used: measurements of dominance, applicability or intensity; use of list of attributes or free comment; subjective or food oral processing-related periods. Thus, AEF was interesting for studying the impact of different implementations on data collection and results obtained. AEF made it possible to collect temporal sequences of attributes, without citation times or durations. This mode of collection with inherent low temporal resolution makes AEF close to static sensory evaluation

methods, which made it useful for studying the importance of temporal resolution in drawing conclusions about the evolution of sensations during tasting.

In this chapter, it was investigated to what extent the capacity of methods to capture temporal differences (within and between products) depended on the resolution of the data collection. Three studies were carried out, aiming to understand the differences between retrospective and concurrent measurements and to discuss different aspects of the implementation of temporal data collection.

The first article was the starting point of this research work. It was initiated following personal doubts about the real ability of the consumers to report their temporal perception in an accurate, valid and reliable manner. It introduced the AEF method (now called “AEF-D”) using the concept of dominance. AEF-D was compared to TDS to question the granularity of temporal sensory data based on a study at laboratory involving French consumers (within-subject design) who tasted dark chocolates (presenting relatively large sensory differences).

The second article went further by investigating the nature of the differences between retrospective and concurrent measurements. It compared the results obtained with a new implementation of AEF (“AEF-A” for AEF with applicability measurement) with those obtained with a discrete time TCATA. The data were collected as part of a research project with an industry partner who wanted to investigate the differences between lemon-flavoured carbonated alcoholic beverages tasted by Japanese consumers (between-subjects design, multi-intakes).

The third article studied the appropriateness of FC-AEF-A, another implementation of AEF (combination of AEF-A and free comment) previously introduced by Mahieu et al. (2020). Indeed, if FC-AEF-A has theoretical advantages related to the absence of predefined list of attributes (tedious to establish, limited in number in dynamic methods, likely to bias responses, etc.), it has only been tested once with products with large sensory differences (the same chocolates as in the first article of this chapter). To test whether the methodological conclusions could be generalized, FC-AEF-A was reused in a research project with an academic partner who needed to collect sensory descriptions for red wines tasted in two contexts: in the laboratory with expert panellists from different countries, and at home with two consumer panels (French and Spanish).

III.1.2. Concurrent vs. retrospective temporal data collection: Attack-Evolution-Finish as a simplification of Temporal Dominance of Sensations?

This article was published in Food Quality and Preference:
<https://www.doi.org/10.1016/j.foodqual.2020.103956>

B. Mahieu: Conceptualization, Methodology, Resources, Validation, Writing - Review & Editing P. Schlich: Methodology, Writing - Review & Editing, Supervision, Funding acquisition; A. Thomas: Methodology, Writing - Review & Editing; M. Visalli: Conceptualization, Formal analysis, Methodology, Software, Writing - Original Draft, Writing - Review & Editing, Visualization.

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-All-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 three 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-Intensity (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 et al., 2012; Hutchings et al., 2014; Jaeger et al., 2018, 2017; Rodrigues et al., 2016; Thomas et al., 2015; Visalli et al., 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 et al., 2011; Torres-Moreno et al., 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 et al., 2011) was used to measure the temporality between several bites corresponding to a full portion. P3R consists in retrospectively picking then ranking the three 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 three 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 three phases. AEF proposes to retrospectively split the tasting in three 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 five 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 four “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 two 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.

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

Figure 34. 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 (INRAE, 2022). 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 34, left), the following instruction was displayed: "Now, put the four 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 five chocolates. After each tasting, we will ask you to choose from a list the three 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 four callets in your mouth and taste them. Focus on the chronological order of the three 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 34, 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 36) 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 three 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.

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 four 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 + error*”, 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 three 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 $(1/\text{number of descriptors})$, 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 & 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 likelihood 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 et al., 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 three points corresponding to the periods of AEF was also plotted to evaluate the within-product evolution. TimeSens software was used to produce these 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 three 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

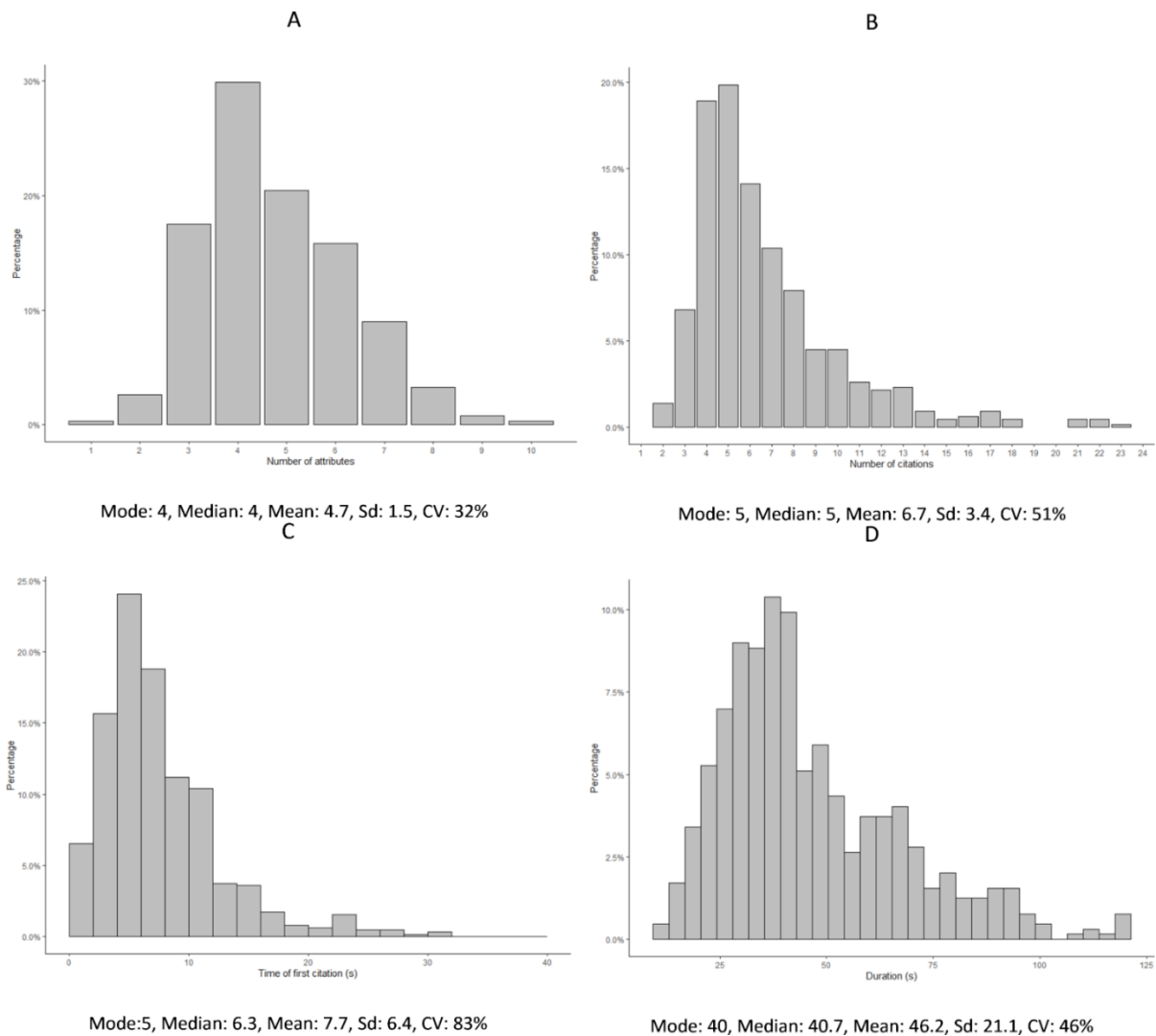


Figure 35. Histograms of consumer behaviour observed during the TDS sequences. A: Distribution of the number of descriptors used in a sequence, B: Distribution of the number of dominant sensations (number of clicks), C: Distribution of the times of first dominant sensations, D: Distribution of the total durations of the evaluation.

Figure 35.A 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 35.B 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 35.A, 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 35.C 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 35.D 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

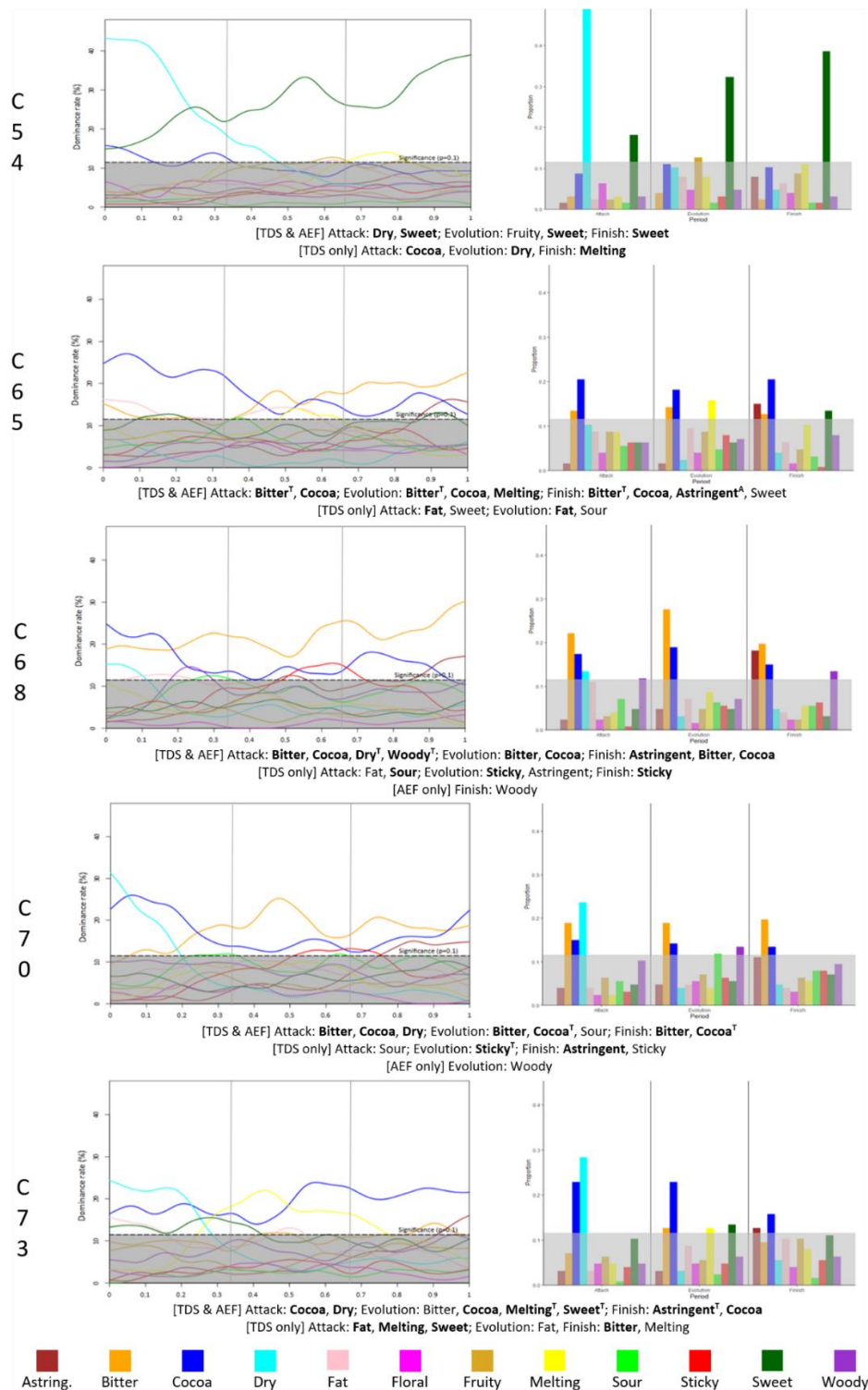


Figure 36. TDS curves (left) and AEF barplots (right) of the proportion of dominances for C54, C65, C68, C70, and C73 (from top to bottom). The grey mask corresponds to the region below significance level (as defined in TDS). The descriptors significant in sense of the binomial test are summarized below 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 36 represents the TDS curves and AEF barplots for the five 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 35 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.

Table 35. ANOVA of durations or citations by descriptor, method and period.

Attribute	Protocol	p-values				NPD			
		All periods	Attack	Evolution	Finish	All periods	Attack	Evolution	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

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).

3.4. Multidimensional analyses

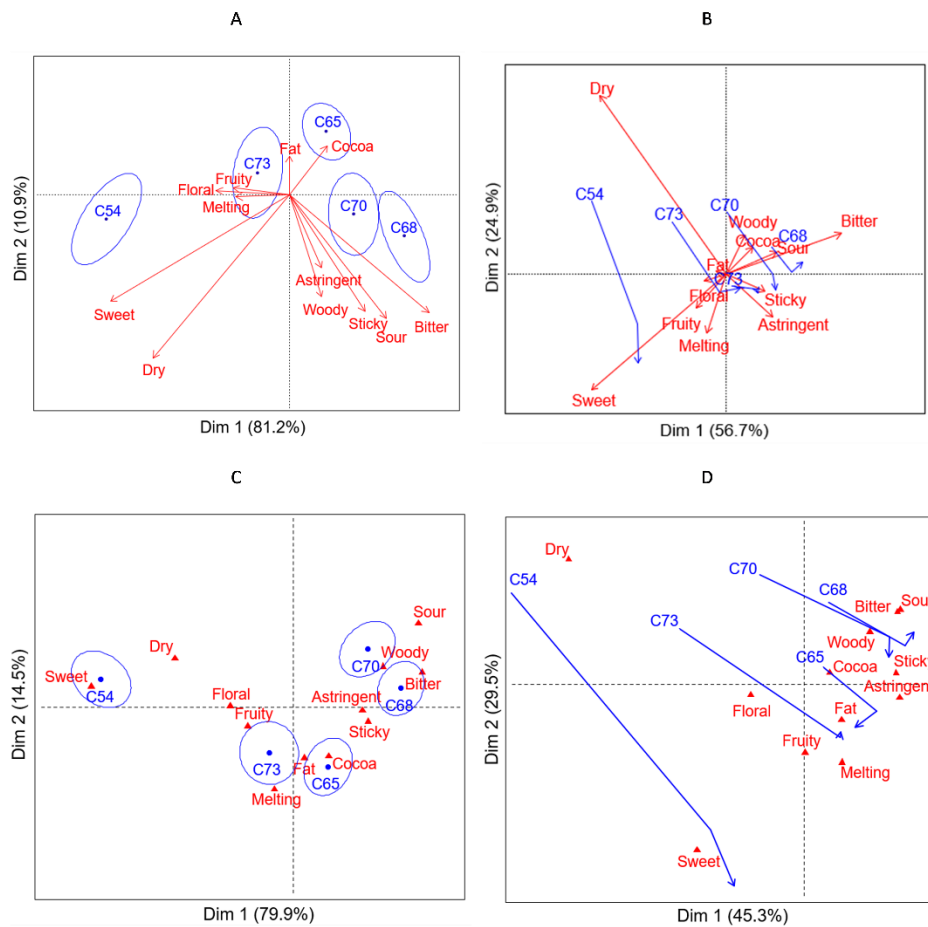


Figure 37. Multidimensional maps, axes 1 and 2. A - CVA of the TDS durations for all periods combined, with 90% confidence ellipses. B – Trajectory PCA of the TDS durations over the three periods of AEF. C – CA of the AEF sensations for all periods combined, with 90% confidence ellipses. D – Trajectory CA of the AEF sensations over the three periods of AEF.

All periods combined and considering all axes (Figure 37.A and Figure 37.C), 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 37.A), and the test for the determination of the number of significant axes (not presented here) recommends to keep three axes. The two first axes of the CA of citations (Figure 37.C) of citations explained 94.4% of variance. Whatever the map, the first axis mainly separated the products in three 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 (Figure 37.B and Figure 37.D), was very similar in TDS and AEF. The PCA of durations (Figure 37.B) explained 81.6% of the variance on the two 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 37.D) explained 74.8% of the variance and can be interpreted in the same way as the PCA.

3.5. Answers to questionnaire

Table 36 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 three 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" (six 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" (five) and that "it is easier to choose the sensations right after placing the chocolate into your mouth or during the tasting" (seven). Positive comments about AEF reported the "easier choice after the tasting, no time pressure" (six), "easier because less choices to do, quick, concise" (four). Negative comments about AEF concerned the difficulty of "picking only three sensations" (five), "differentiating the periods" (three) or "memorizing the order" (five). Negative comments about TDS were about the "lack of clarity of the explanations" (two) or "a need for more instructions, a warm-up, a training" (two). Regardless of the method, several consumers also found it "too bad for not being able to evaluate more flavours" (two), that "some descriptors are useless" (three), that "some descriptors are missing" (two), or they regretted not being able to "express a free opinion" (one).

Table 36. 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).	3.88 ± 0.17	
7B. It was easy to identify the sensations perceived at the beginning of the tasting.		3.25 ± 0.16
7C. It was easy to identify the sensations perceived at the middle of the tasting.		3.98 ± 0.14
7D. It was easy to identify the sensations perceived at the end of the tasting.		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

4. Discussion

4.1. Overall differences in conclusions obtained with TDS and AEF

AEF measured temporality in products in a minimalistic way, as each sequence was composed of three 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 two methods were able to discriminate all product pairs.

4.2. 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 (three 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 reproducible than those of TDS.

4.3. 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 that 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, Délérís, & 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 favourably by the experimenter (bias of social desirability, Cerri et al., 2019). However, the (few) answers to the open-ended question about the overall opinion of the task suggested that 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 three 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.

4.4. 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 three 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).

4.5. 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 that in this study durations did not provide additional information compared to citations.

4.6. 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.

4.7. 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 three 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.

4.8. 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 three 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).

4.9. 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 five seconds then *Bitter* two 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 three 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 three phases. AEF

was based on a concept close to dominance and allowed consumers to select three 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 three 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 three 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.

III.1.3. Concurrent vs. immediate retrospective temporal data collection. A case study on lemon-flavoured carbonated alcoholic drinks

This article was published in Food Quality and Preference:
<https://www.doi.org/10.1016/j.foodqual.2022.104629>

P. Schlich: Conceptualization, Methodology, Software, Supervision, Funding acquisition, Writing - Review & Editing; M. Visalli: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization; T. Wakihira: Conceptualization, Methodology, Validation, Data Curation, Resources, Project administration, Investigation, Writing - Review & Editing.

Most temporal sensory methods measure product perception concurrently with tasting. However, retrospective measurement could have the advantage of being easier to implement with consumers. To date, no study has explicitly compared these two modes of temporal data acquisition. The objective of this study was to compare the temporal description obtained from consumers who had drunk and evaluated a full can of four lemon-flavoured carbonated alcoholic drinks at home and on different days. The consumers were separated into two panels and asked to select from eight attributes on a check-all-that-apply list that were applicable during three periods of perception of a sip - “in mouth before swallowing”, “immediately after swallowing” and “aftertaste” - for three sips of each can. The first panel (RET, 97 consumers) did the evaluation retrospectively immediately after the tasting, while the second (SIM, 96 consumers) did it concurrently with the tasting. Data were analysed using the multiple-response correspondence analysis (mrCA) framework applied at different levels: product, period and sip. The data from RET and SIM captured the differences between products and periods, with the differences between the products being larger than the differences between the periods. In both methods, no sip effect was observed. Perception of the products were identical in sips 1, 4 and 7. The consumers of RET and SIM agreed overall on the relative differences between products, although the level of discrimination was stronger for the consumers of RET. The consumers of RET and SIM only partially agreed on the between-period differences by product. The results suggest that the retrospective temporal evaluation could better discriminate the products and that the temporality patterns may be different between the two methods; therefore, there is a need for more research about the resolution of temporal data.

1. Introduction

Temporal perception has been investigated in sensory science for the last 60 years using different methodologies (Kemp et al., 2017). The first developed methods measured the evolution of the intensity of the attributes over time. Some did it in a continuous way: time-intensity (TI, Lee & Pangborn, 1986), dual-attribute time-intensity (DATI, Duizer et al., 1997) or multi-attribute time-intensity (MATI, Kuesten et al., 2013). Others did it at fixed discrete time points: discrete time-intensity (Clark & Lawless, 1994), progressive profiling (Jack et al., 1994) or sequential profiling (Methven et al., 2010). More recent methods, which did not rely on intensities, measured the evolution of dominant attributes (Temporal Dominance of Sensations - TDS, Pineau et al., 2009) or applicable attributes (Temporal Check All That Apply - TCATA, Castura et al., 2016) in continuous time. The common point between all these methods is that the data collection is carried out simultaneously to the tasting.

Few studies have applied retrospective temporal measurements in sensory analysis. Visalli et al. (2020) proposed the attack-evolution-finish (AEF) method to study the temporal perception of chocolates. They compared the results obtained with AEF with those obtained with TDS and concluded on very similar product discrimination. Mahieu, Visalli, Thomas, et al. (2020) extended the AEF concept by using free comment, avoiding the issues inherent in the use of a predefined list of descriptors. While losing temporal resolution, the two methods present the advantage of being easy to implement with consumers, as the task does not require specific training or familiarization, contrary to TDS (Boinbaser et al., 2015; Hutchings, Foster, Grigor, et al., 2014; Kantono et al., 2018; Rodrigues, de Souza, et al., 2016; Thomas et al., 2015; Velázquez et al., 2020b) or TCATA (Ares et al., 2015; Jaeger et al., 2017; Rizo et al., 2020; Weerawarna et al., 2021).

The lack of attention to retrospective measurements in sensory analysis can probably be explained by the tacit assumption that concurrent temporal sensory data collection is more precise and less biased than retrospective data collection. Some psychologists have judged real-time data capture preferable for measuring changes over time, using concurrent assessment to ask people to report on their current experience (Stone et al., 2007). Others claimed that concurrent data collection does not eliminate other potential sources of bias in self-reports (Schwarz, 2012). Aldrovandi et al. (2015) presented a series of studies showing that retrospective evaluation of event sequences could be biased by memory and notably by

the tendency to retrieve the most readily available information that is itself related to the valence of the information. However, meta-analytic findings (Block et al., 2010) revealed that if the duration was less than 60 seconds, there was no difference observed between conclusions obtained with simultaneous and retrospective paradigm. Noncongruent results related to taste perception have been reported. Liang et al. (2018) showed that the stress of the memory load influenced sensory perception by decreasing taste sensibility, while Daniel and Katz (2018) showed that a higher accuracy in taste recognition was observed over a 30-s delay. Recently, Botha et al. (2021) compared TCATA data collection with a retrospective approach for phenotyping thermal taste. The authors judged that TCATA avoids memory effects and allows participants to fully focus on the new taste instead of relying on their working memory, but they also recognized that the cognitive aspects of TCATA have not been studied to date. Varela et al. (2018) also used an immediate retrospective think-aloud protocol to review TDS results to better understand the dominance concept and gain insight about the temporal perception.

Indeed, memory is not the only aspect to consider, as it is just one component of the cognitive load. The cognitive load is a complex and multidimensional construct consisting of causal factors related to the task and the subject and assessment factors related to mental load, mental effort, and performance (Paas & Van Merriënboer, 1994). In sensory analysis, TDS and TCATA have been reported to be potentially demanding tasks by some authors (Botha et al., 2021; Castura et al., 2016; Jaeger et al., 2017; Pineau et al., 2012). This could affect the outcomes of these methods, as Wal & Dillen (2013) suggested that an increased task load could reduce taste perception. This was partially confirmed by Velázquez et al. (2020), who showed that a TCATA task was feasible by children on a simple video stimulus, but when applied to more complex food stimuli, they used TCATA (and TDS) methods as static methods. The performances of retrospective and concurrent measurements have been compared in other research fields, still leading to controversy. Church et al. (2019) compared humans' metacognitive performances (the ability to choose between two answers with and without feedback) in prospective and retrospective paradigms, showing that the metacognitive response was used more robustly and accurately retrospectively when it was not in direct competition with the primary perceptual responses. Van Den Haak et al. (2003) showed that concurrent and retrospective think-aloud protocols revealed comparable sets of usability

problems, but in the concurrent protocols, the requirement to think aloud while working had a negative effect on the task performance, raising the question about the reactivity of concurrent think-aloud protocols in the case of high task complexity. Peute et al. (2015) showed that the concurrent think-aloud method was more efficient in assessing the usability of a data query tool but also that it did not outperform retrospective methods that additionally elucidated unique problems. Kuusela & Paul (2000) reported that the concurrent protocol generally outperformed the retrospective method in verbal protocol analysis, but they also noticed that the retrospective method provided more statements about final choices. Similarly, Whyte et al. (2010) concluded that concurrent verbal reports provided by nurses during and after administering care in a simulated task environment provided a more complete representation of the cognitions of research participants, but additional unique data were exclusively present in the retrospective reports. Ryan & Haslegrave (2007) showed that concurrent reports only contained a proportion of information about workers' thoughts, while additional information could be obtained from retrospective reports, although these appeared to be vulnerable to bias and reordering of reported information due to the rationalization of thoughts.

With regard to this context, the objective of this study was to compare the sensory descriptions of commercial products collected from consumers using either a concurrent or a retrospective data collection method.

2. Materials and methods

2.1. Simultaneous and retrospective protocols

Two data collection protocols based on Check-all-that-apply (CATA, Ares & Jaeger, 2015) questions were used in this study, the first one simultaneous to the tasting (SIM), the second one retrospective to the tasting (RET). A detailed description of the two protocols can be found in the related data paper (Visalli et al., 2022).

2.2. Samples

The five products (P1, P2, P3, P4, and P5) were 350 ml cans of commercial lemon-flavoured carbonated drinks. P1 was a non-alcoholic drink that served as a warm-up. P2 to P5 were white liquor-based (Japanese Shochu or Vodka) alcoholic drinks, referred to as "Chu-hai" in Japanese. The cans were blinded by white-coloured masking films and coded using three-digit

labels, and they were presented according to a Williams Latin square, but P1 was served first to every consumer.

2.3. Consumers

Two hundred consumers aged 20 to 39, who were regular consumers of lemon-flavoured carbonated alcoholic beverages, were recruited through an online questionnaire from a panel of consumers belonging to a research agency in Japan. “Regular” referred to consumers drinking flavoured carbonated alcoholic beverages with a frequency of at least twice a week and lemon-flavoured carbonated alcoholic beverages at least once a month. The design of the test was explained to consumers in the online questionnaire. They were informed they would do the test on their smartphones, and they had to sign a consent form to participate in the study. They were financially compensated for their participation. The consumers were separated into two panels that were balanced in their composition (gender, age, frequency of consumption). The first panel (RET) had to evaluate the products retrospectively to the tasting just after they declared that they no longer perceived anything, while the second (SIM) evaluated them concurrently to the tasting.

2.4. Descriptors

The same list of descriptors was provided for both the RET and SIM panels: *Alcohol*, *Bitter*, *Carbonated*, *Lemon*, *Refreshing*, *Sour*, *Sweet aroma* and *Sweet taste* (in Japan, sweet aroma is used when it is not possible to describe detailed quality of sweetness in terms of aroma. Sweet aroma can include different types of aromas, such as fruity, floral, caramel, vanilla, honey, etc.) The descriptors were presented as a Check-All-That-Apply (CATA) list in a random order on the screen, but this order was constant for each consumer across evaluations. No definitions of the descriptors were given to the consumers.

2.5. Experimental procedure

The consumers had to evaluate the products using the browser of their smartphones (TimeSens V2 web app, INRAE, 2022). The procedures for the full consumption of the can is summarized in Figure 38. Only three sips were evaluated to limit the duration of the task and to avoid boredom (Thomas et al., 2018).

Figure 39 describes the tasting of sips 1, 4, 7 for panel RET. During the tasting, several screens were displayed to guide the consumer. A first screen invited the consumer to take a sip without swallowing. Then, a screen was displayed during eight seconds (T1, "In mouth before swallowing") instructing the consumer not to swallow while focusing on perceived sensations and memorizing them. Then, a screen was displayed during two seconds, inviting the consumer to swallow. Then, a screen was displayed during eighteen seconds (T2, "Immediately after swallowing") instructing the consumer to focus on perceived sensations and memorize them. Then, a screen instructing the consumer to focus on perceived sensations and memorize them was displayed until the consumer declared he no longer perceived anything (T3, "Aftertaste"). Then, the CATA screen was displayed with no time limit, with the instruction: "What did you perceive during the different steps of the tasting? Please select all the sensations that apply during each period". Finally, the liking score was asked to the consumer (on a 0-10 continuous scale, 0 being labelled "not at all" and 10 "very much"). Figure 40 describes the tasting of sips 1, 4, 7 for panel SIM. During the tasting, several screens were displayed to guide the consumer. A first screen invited the consumer to take a sip without swallowing. Then, a CATA screen was displayed during eight seconds (T1, "In mouth before swallowing"), with the instruction: "What do you perceive now? Please select all the sensations that apply". Then, a screen was displayed during two seconds, inviting the consumer to swallow. Then, a screen was displayed during eight seconds (T2, "Immediately after swallowing") instructing the consumer to select all the sensations that apply (same instruction as in T1). Then, a screen inviting the consumer to wait was displayed during ten seconds. Then, the last CATA screen was displayed until the consumer declared he no longer perceived anything (T3, "Aftertaste"), instructing the consumer to select all the sensations that apply (same instruction as in T1). Finally, the liking score was asked to the consumer (on a 0-10 continuous scale, 0 being labelled "not at all" and 10 "very much").



Figure 38. Description of the tasting of the drink over seven sips (common for RET and SIM). Only sips 1, 4 and 7 were evaluated.

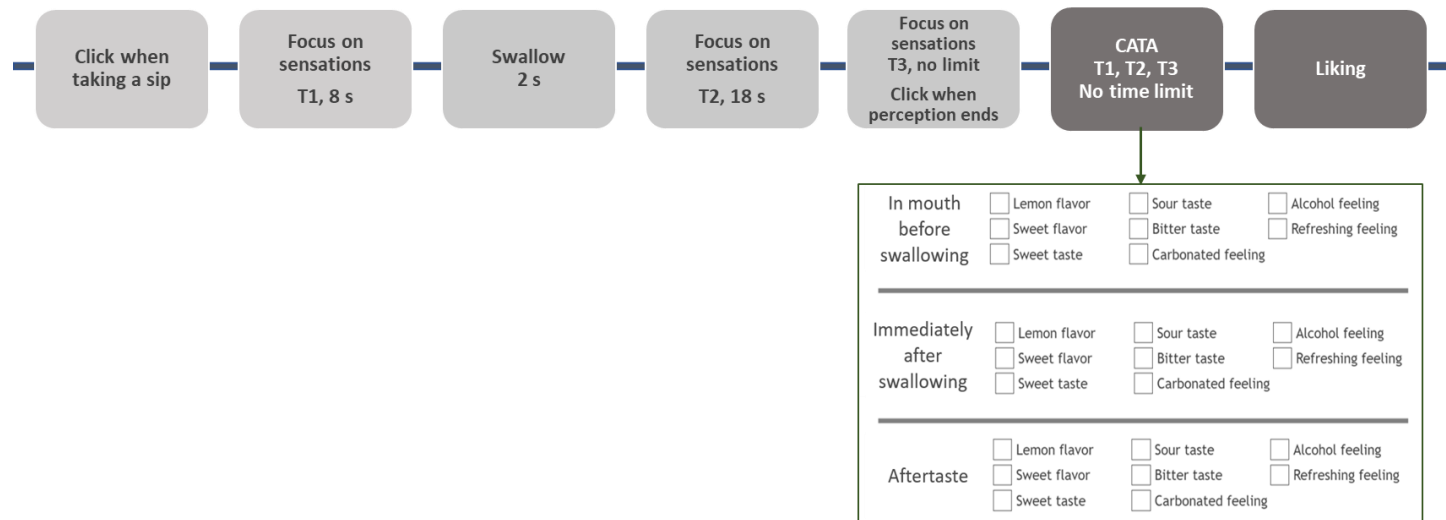


Figure 39. Evaluation of sips 1, 4, 7 for panel RET using retrospective CATA. Only one CATA measurement screen was displayed, after the tasting. “In mouth before swallowing”, “Immediately after swallowing” and “Aftertaste” correspond to periods T1, T2 and T3.

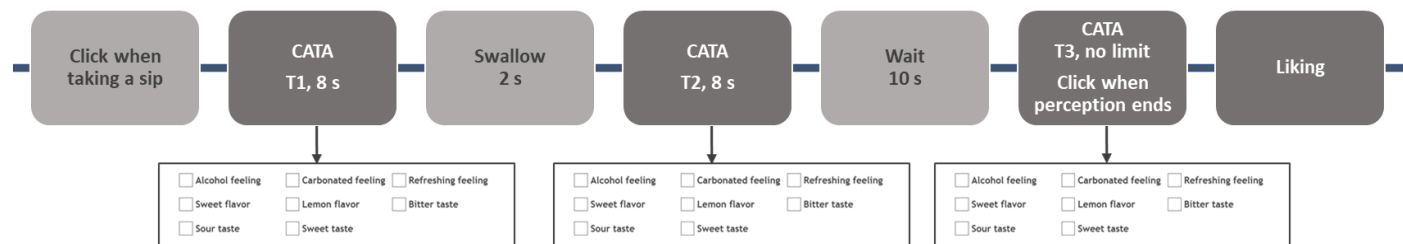


Figure 40. Evaluation of sips 1, 4, 7 for panel SIM using simultaneous CATA. The CATA measurement screen was displayed during the tasting within periods “In mouth before swallowing” (T1), “Immediately after swallowing” (T2) and “Aftertaste” (T3).

2.6. Data analysis

This article focuses on methodological aspects related to concurrent and retrospective evaluation. Conclusions on products and liking will be published elsewhere.

The data for product P1 (the warm-up) were excluded. Six consumers had incomplete data in the SIM session and one in RET because they did not complete their sessions. Their data were kept because the statistical methods used can deal with incomplete datasets.

Individual consumer data were stored as a table with seven columns: "Panel" (RET, SIM), "Consumer", "Sip number" (1, 4 or 7), "Product" (P2, P3, P4 or P5), "Attribute" (*Carbonated, SweetF, Refreshing, Bitter, Lemon, SweetT, Sour, Alcohol*), "Period" (T1, T2, T3) and "Score" (1 applicable, 0 otherwise). Each consumer contributed to 288 rows (3 sips x 4 products x 8 descriptors x 3 periods).

Statistical analyses were performed using R 4.1.0 software (R Core team, 2020) and the package MultiResponseR (Mahieu, 2021).

2.6.1. Use of descriptors

For the RET and SIM panels, mean number of citations and descriptors were obtained in various ways. Citations were averaged: (i) across sips, products and consumers within each period (T1, T2, T3); (ii) across products and consumers by sip (1, 4 or 7) and overall. Descriptors were averaged by sip (1, 4 or 7) and overall: (i) across products and consumers; (ii) across consumers. The 95% confidence intervals were computed based on Poisson log-linear models for count data (R function `glm`, family = Poisson). Two means were considered different when their confidence intervals did not intersect.

2.6.2. Product/period/sip comparisons

CATA results were aggregated into contingency tables, the column variables being the descriptors (see examples in supplementary materials). For studying product-by-sip sensory trajectories over periods, one contingency table was computed for each panel, the row variables being the combinations of "product x period x sip". For studying between-sip differences, twelve contingency tables (one for each "product x period" combination) were computed for each panel, the row variables being the sip. For studying between-product differences, four contingency tables (one for each period plus one with citations averaged

over periods, with sips being pooled and considered to be independent observations) were computed for each panel, the row variables being the products. For studying between-period differences, five contingency tables (one for each product plus one with citations averaged over products, with sips being pooled and considered to be independent observations) were computed for each panel, the row variables being the periods.

As the consumers could check multiple attributes for a product during a period, the usual χ^2 framework was not well suited to analyse these multiple-response data. To overcome this limitation, the approach described in Mahieu et al. (2021) was used. The dimensionality of the dependence between row and column variables was tested using the dimensionality test (2000 simulations) based on multiple-response χ^2 framework. Then, if and only if at least one dimension was significant ($\alpha=0.05$), the multiple-response correspondence analysis (mrCA) was computed on the contingency table. Outputs of mrCA were displayed using a standard biplot, and 95% confidence ellipses were computed with a total bootstrap procedure (Cadoret & Husson, 2013) with 2000 simulations, Procrustes rotations being performed on the significant dimensions. When the periods were analysed, the sensory trajectory of each product was represented by a two-segment arrow from period 1 to 2 and then from period 2 to 3 as in Mahieu et al. (2020). For each pair of row variables of the contingency table, a total bootstrap test was performed on the significant dimensions for assessing the significance of difference. When significant ($\alpha=0.05$), for each pair of row and column variables (cell), a multiple-response hypergeometric test (2000 simulations, $\alpha = 0.05$) was performed to test if the descriptor was cited in a proportion that was significantly greater than the overall average citation proportion. No adjustments have been made for the multiple hypergeometric tests.

2.6.3. Feedback of the consumers about the task

The answers to the free-text question about the difficulty of the task (Q6, Visalli et al., 2022) were investigated. Q6 was labelled: "Please tell us what you felt about this tasting survey (if the tasting method was difficult, easy, etc.)". Similar answers about the difficulty of the task were manually grouped by the experimenter, then counted. No statistical analysis was performed with these data.

3. Results

One hundred and ninety-three consumers participated in the study: 97 consumers in panel RET, and 96 in panel SIM.

3.1. Use of descriptors by the consumers

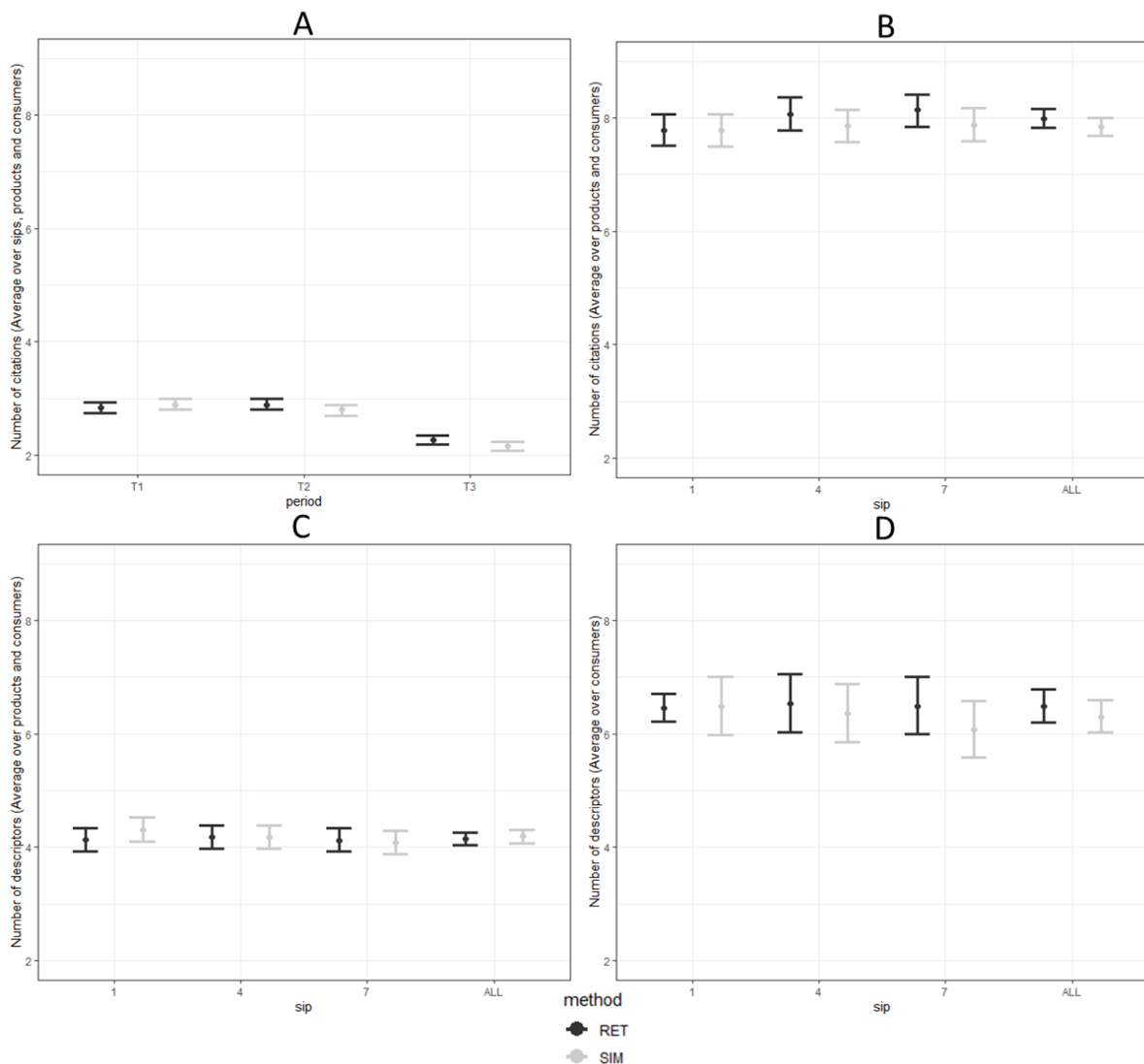


Figure 41. Use of descriptors. A – Number of citations averaged over sips, products and consumers, by period. B – Number of citations averaged over products and consumers, by sip and all sips pooled. C – Number of descriptors averaged over products and consumers, by sip and all sips pooled. D – Number of descriptors averaged over consumers, by sip and all sips pooled. For all figures, vertical bars represent 95% confidence intervals of the means based on Poisson log-linear models for count data.

Figure 41 shows how the descriptors have been used by the two panels. Figure 41.A shows that the consumers of the two panels cited almost three descriptors by product in periods T1 and T2; then, this number significantly decreased to just over two in period T3. Figure 41.B shows that the number of citations by product was approximately eight. These number of citations slightly increased over the three sips (not significantly). Figure 41.C shows that the number of descriptors used by the product was approximately four for the two panels. The number of descriptors slightly decreased over the sips in panel SIM (not significantly). Figure 41.D shows that the total number of descriptors used to characterize the four products was between six and seven (over the eight of the CATA list). The decrease observed in Figure 41.D for panel SIM was more pronounced but still not significant.

3.2. Product comparison

3.2.1. Product-by-sip sensory trajectories over periods

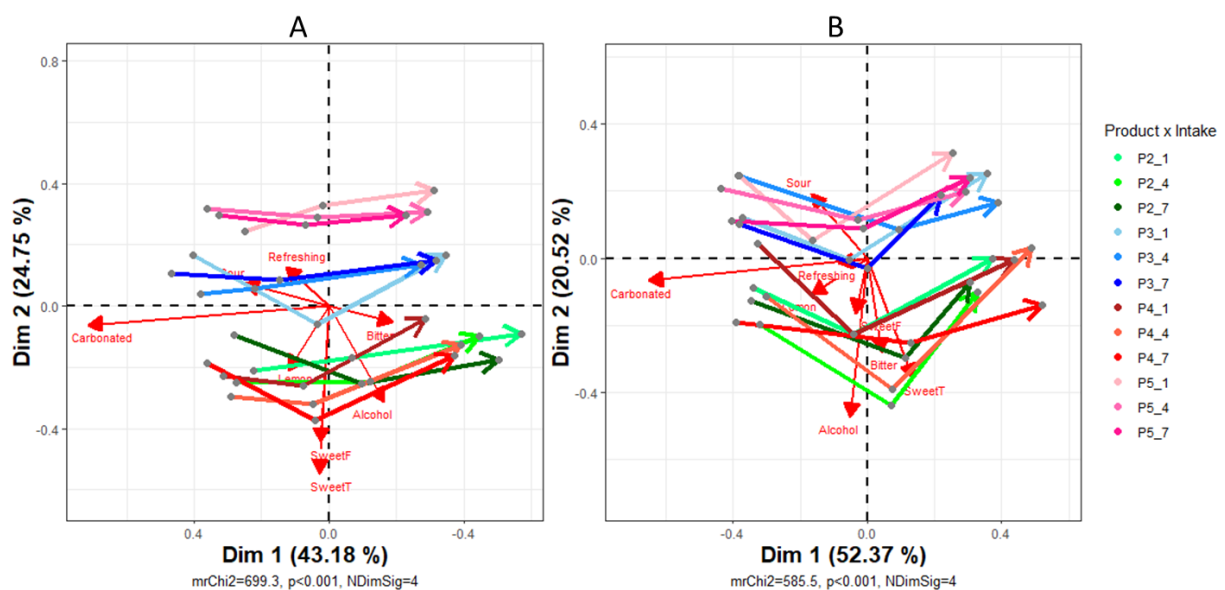


Figure 42. *mrCA* of 36 evaluations (product x sip x period) representing the product by sip sensory trajectories over periods on dimensions 1 and 2.

Figure 42.A (left) corresponds to the RET panel, and Figure 42.B (right) corresponds to the SIM panel. “ $mrChi2$ ” is the statistic of the $mr\chi^2$ for the dimensionality test, “ p ” is the p -value of the dimension test, “ $NDimSig$ ” is the number of significant dimensions. The sensory trajectory of each product was represented by a two-segment arrow from period 1 to 2 and then from period 2 to 3. Red arrows (with the triangle arrowheads) indicate the relative importance and correlation between the descriptors and the dimensions.

For the two panels, the first dimension depicted the evolution from Carbonated to Bitter over the three successive periods of perception, whereas the second dimension opposed the products mostly characterized by Sour and Refreshing to the products mostly characterized by SweetT, SweetF and Alcohol. Bitter was perceived more in period T3 in panel RET and in period T2 in panel SIM. The trajectories were straighter in the RET panel than in the SIM panel, meaning that the sips were not that different (for RET panel) in dimension two within the same product. The same number (4) of significant dimensions was observed, denoting a similar complexity of differences. However, discrimination was better with panel RET (mr χ^2 of 699.3 vs. 585.5 for SIM). The period trajectories exhibited a clear evolution of perception, meaning that the descriptors were not selected in the same way in periods T1 (more citations of Carbonated), T2 (more citations of SweetF), and T3 (less citation in average). However, these trajectories seem to be rather similar across products. Products P2 and P4 showed similar trajectories for both panels. Products P3 and P5 also showed similar trajectories for panel SIM but were more separated for panel RET.

There was almost no evolution of perception across sips regardless of the product, except for a tendency with product P4 in period 1 with the panel SIM (detailed results are presented in supplementary materials). Therefore, in the rest of this article, sips will subsequently be pooled and considered to be independent observations.

Figure 43.A (left) corresponds to the RET panel, and Figure 43.B (right) corresponds to the SIM panel. “mrChi2” is the statistic of the mr χ^2 for the dimensionality test, “p” is the p-value of the dimension test, “NDimSig” is the number of significant dimensions. The sensory trajectory of each product was represented by a two-segment arrow from period 1 to 2 and then from period 2 to 3. Red arrows indicate the relative importance and correlation between the descriptors and the dimensions.

Figure 43 shows the same sensory interpretation as Figure 42. The RET panel was still more discriminative (632.24 vs. 499.07) and exhibited more complex differences (eight vs. five significant dimensions).

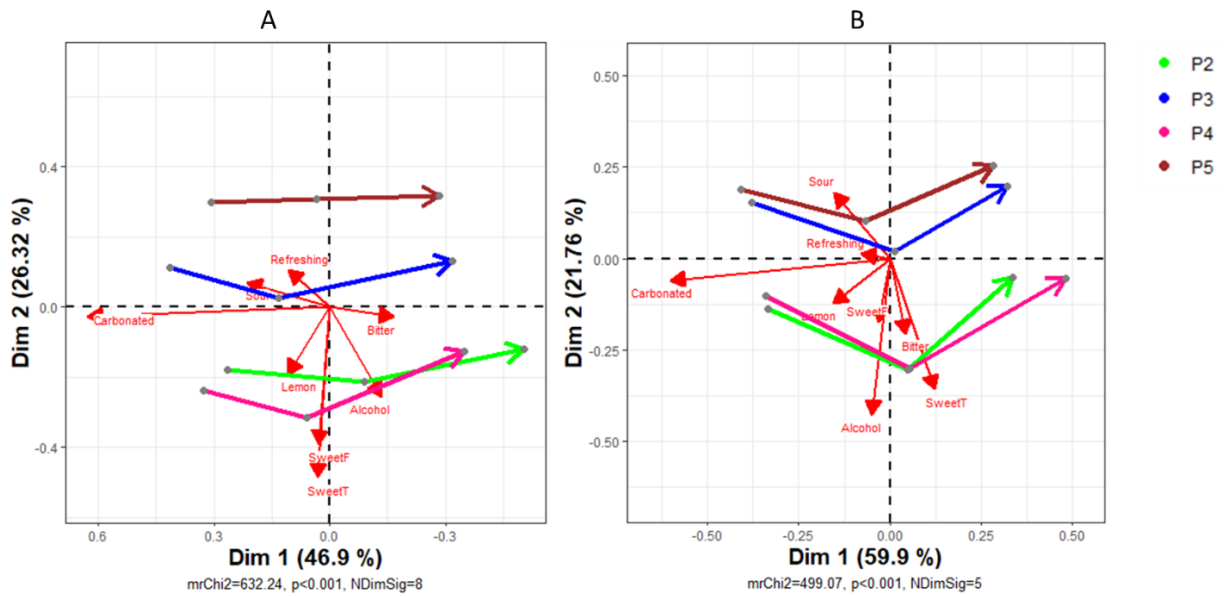


Figure 43. mrCA of 12 evaluations (product x period, sips as independent observations) representing the product sensory trajectories over periods on dimensions 1 and 2.

3.2.1. Comparison of products by period

Figure 44 illustrates the differences between products within each period or all periods pooled, with citations averaged over periods. Both panels discriminated the four products (P3 and P5 on dimension 3 and only in T3 and overall, not represented in panel SIM). The $mr\chi^2$ statistics show that panel RET was still more discriminative than SIM in every period. RET and SIM panels were more discriminative in period T1 than in periods T2 and T3.

Table 37 and Table 38 show the percentages of citations by attribute, product and period for the two panels, focusing on between-product comparisons. It should be noted that the results are presented in a single table for convenience, but the multiple-response hypergeometric tests have been made by period. Thus, the multiple-response hypergeometric tests per cell (Table 37 and Table 38) should be interpreted by block (T1, T2, T3 or all periods) because each block summarizes the result of one mrCA. For example, *Alcohol* was cited in P2 significantly more than in other products during period T1 in the two panels.

The first three blocks of the tables allowed us to compare each product to the others inside each period. P2 was more often described as *Alcohol* during periods T1, T2, T3 (RET, SIM), more *Bitter* in period T3 (RET) or T2 (SIM), more *SweetF* during periods T1, T2, T3 (RET), and more *SweetT* during period T1 (RET). P3 was more often described as *Lemon* during period T3 (RET), more *Refreshing* during period T2 (RET), and more *Sour* during periods T1, T2, T3 (RET,

SIM). P4 was more often described as *Lemon* during periods T2 (RET, SIM), T1, T3 (RET), more *SweetF* in periods T1 and T2 (RET, SIM) and more *SweetT* in periods T1, T2, T3 (RET, SIM). P5 was more often described as *Carbonated* in periods T1 (RET) and T3 (RET, SIM), *Refreshing* in periods T2 and T3 (RET) and *Sour* in periods T1 and T2 (SIM).

The last block of the tables allowed us to compare each product to the others without considering temporality. P2 was more often described as *Alcohol* (RET, SIM) and *SweetF* (RET). P3 was more often described as *Sour* (RET, SIM) and *Lemon* (RET). P4 was more often described as *Lemon*, *SweetF* and *SweetT* (RET, SIM). P5 was more often described as *Carbonated* (RET), *Refreshing* (RET) and *Sour* (SIM). If we ignored periods, products were still discriminated, but some differences were visible only during specific periods (for example, *Bitter* for product P2 during T3 and *Refreshing* for P3 in T2 in panel RET, *Bitter* for product P2 during T2 or *Carbonated* for Product P5 during T3 in panel SIM) and would have been missed without considering a temporal approach.

The two panels were in overall agreement except for P5, which was perceived as more *Sour* only by panel SIM. Overall, panel RET highlighted more differences (34) between products than panel SIM (22). The citation rates for *Alcohol* seemed to follow the alcoholic contents of the products.

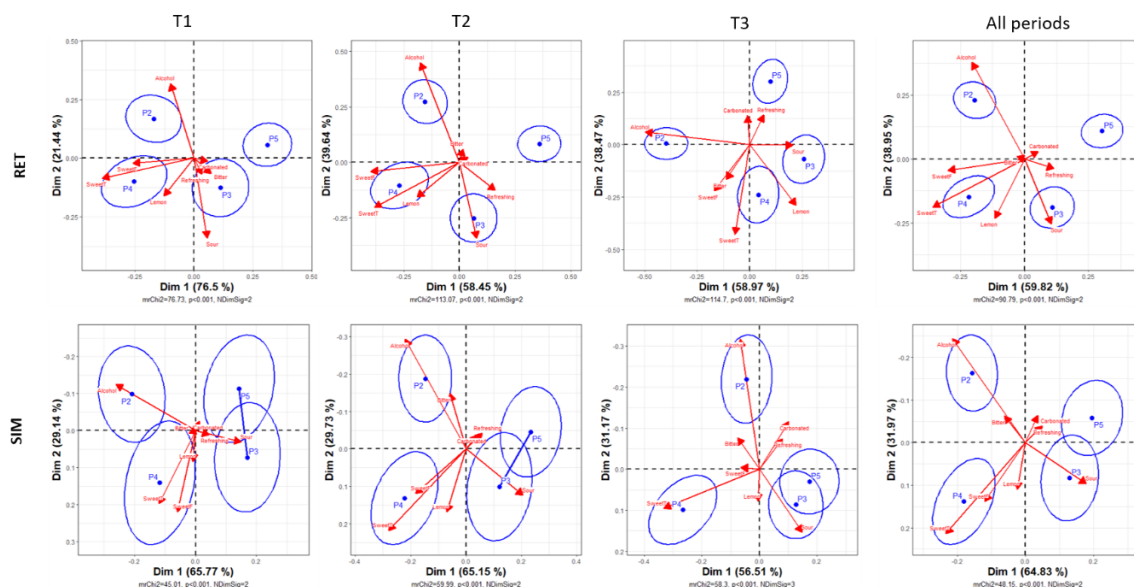


Figure 44. mrCA of the four products, by period and averaged over all periods on dimensions 1 and 2, for panels RET (first row) and SIM (second row). T1 (first column), T2 (second column), T3 (third column) and “All periods” (fourth column) correspond to the periods. “ $mr\chi^2$ ” is the statistic of the $mr\chi^2$ for the dimensionality test, “ p ” is the p -value of the dimension test, “ $NDimSig$ ” is the number of significant dimensions. Red arrows indicate the relative importance and correlation between the descriptors and the dimensions. Blue lines connect not discriminated products (whatever the represented dimensions).

Table 37. Percentages of citations by attribute (rows), product and period (columns) for panel RET.

	T1				T2				T3				All periods			
	P2	P3	P4	P5	P2	P3	P4	P5	P2	P3	P4	P5	P2	P3	P4	P5
Alcohol	33.3*	20.8	25.4	24.7	50.0*	24.3	35.8	33.3	46.9*	18.1	26.7	27.1	43.4*	21.2	29.5	28.5
Bitter	19.1	22.9	20.1	22.9	31.9	29.5	29.9	31.9	33.7*	28.5	33.0	24.7	28.1	27.1	27.8	26.7
Carbonated	67.4	70.8	67.0	72.2*	51.0	50.7	49.7	52.8	25.0	23.6	21.9	27.8*	47.9	48.3	46.2	50.7*
Lemon	59.0	60.8	67.4*	52.4	60.1	65.3	70.1*	51.7	43.1	62.5*	62.9*	45.8	54.2	62.9*	66.7*	50.0
Refreshing	30.6	34.0	32.3	33.3	30.2	39.9*	33.0	40.3*	26.0	28.8	25.7	32.3*	29.2	34.4	30.2	35.4*
Sour	32.6	47.2*	43.4	41.0	30.6	50.4*	42.7	41.3	24.0	36.1*	31.9	33.0	29.2	44.4*	39.6	38.2
SweetF	19.4*	13.9	21.5*	9.7	17.4*	13.5	20.8*	5.6	17.7*	12.2	16.3	8.7	18.1*	13.2	19.8*	8.0
SweetT	21.5*	14.6	26.0*	6.9	18.4	19.8	26.4*	7.6	17.0	16.3	21.9*	6.9	18.8	16.7	24.7*	7.3

Significant (alpha=0.05) multiple-response hypergeometric tests per cell for product comparison are indicated by an asterisk.

Table 38. Percentages of citations by attribute (rows), product and period (columns) for panel SIM.

	T1				T2				T3				All periods			
	P2	P3	P4	P5	P2	P3	P4	P5	P2	P3	P4	P5	P2	P3	P4	P5
Alcohol	39.2*	19.9	30.4	24.7	44.0*	26.2	34.4	28.5	33.3*	21.3	22.3	20.4	38.5*	22.0	29.3	25.0
Bitter	23.1	22.0	23.1	21.8	35.9*	28.7	30.8	30.6	26.7	22.7	25.6	21.8	28.6	24.5	26.4	25.0
Carbonated	71.8	71.6	70.3	72.9	46.2	46.5	44.0	47.9	27.8	23.8	21.3	29.6*	48.7	47.5	45.1	49.7
Lemon	64.5	67.4	69.2	63.7	60.4	65.6	70.7*	59.9	49.1	55.3	52.8	49.3	57.9	61.7	64.5*	58.5
Refreshing	32.6	35.8	33.7	35.2	34.8	35.5	32.6	37.0	31.9	31.2	26.0	31.0	33.0	33.7	31.1	34.9
Sour	29.7	40.8*	33.3	39.1*	27.5	39.0*	30.8	39.8*	22.3	33.0*	23.8	29.6	26.0	36.9*	29.7	36.6*
SweetF	13.6	15.3	19.1*	10.9	14.3	13.1	18.3*	10.2	13.6	14.9	13.6	8.8	13.6	13.5	17.6*	10.9
SweetT	16.5	15.3	20.9*	11.3	21.6	20.2	30.4*	14.1	18.7	14.9	30.4*	14.4	19.1	17.4	26.7*	12.7

Significant (alpha=0.05) multiple-response hypergeometric tests per cell for product comparison are indicated by an asterisk.

3.2.3. Comparison of periods by product

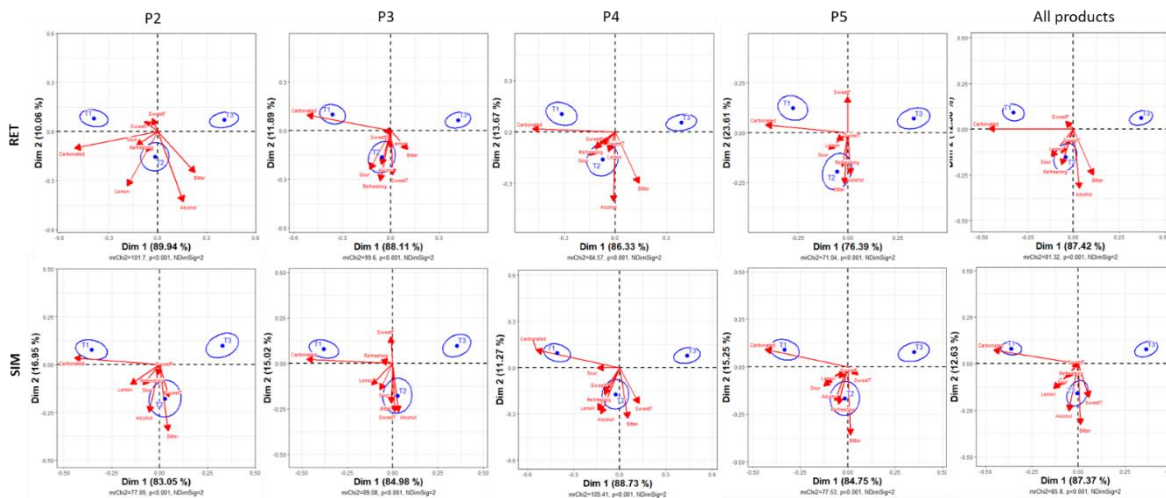


Figure 45. *mrCA* of the three periods, by product and averaged over all products on dimensions 1 and 2. RET (first row) corresponds to the retrospective panel, SIM (second row) to the simultaneous panel. P2 (first column), P3 (second column), P4 (third column), P5 (fourth column) and “All products” (fifth column) correspond to the products. “*mrChi*²” is the statistic of the *mr* χ^2 for the dimensionality test, “*p*” is the *p*-value of the dimension test, “*NDimSig*” is the number of significant dimensions. Red arrows indicate the relative importance and correlation between the descriptors and the dimensions.

Figure 45 shows that all periods were separated, denoting a clear temporality, but this temporality was globally the same regardless of the product. Overall, there is no difference in discrimination between the two panels, RET being more discriminative for P2 and P3, and SIM being more discriminative for P4 and P5 and for “All products”. The period T1 was always associated with Carbonated, while T3 was the period having the less multidimensional differences.

Table 39 and Table 40 show the percentages of citations by attribute, product and period for the two panels, focusing on between-period comparisons. The fifth block illustrated the overall product temporality. The interest of investigating temporal differences was confirmed, as products were overall more often described *Alcohol* in period T2 (RET, SIM), more *Bitter* in period T2 (RET, SIM) and T3 (RET), more *Carbonated* in T1 (RET, SIM), more *Lemon* in T1 (SIM) and T2 (RET, SIM), more *Refreshing* in T2 (RET), more *Sour* in T1 and T2 (RET, SIM), more *SweetF* in T2 (SIM). Only the citation rates of *SweetF* were constant over the periods.

Table 39. Percentages of citations by attribute (rows), product and period (columns) for panel RET.

	P2			P3			P4			P5			All products		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
Alcohol	33.3	50.0*	46.9*	21.2	24.3*	18.1	25.4	35.8*	26.7	24.7	33.3*	27.1	26.0	35.8*	29.9
Bitter	19.1	31.3	34.7*	22.6	28.6	30.6*	20.5	31.3*	31.3*	23.3	32.6*	23.6	21.5	30.9*	29.9*
Carbonated	68.40*	50.4	24.3	72.6*	49.7	21.9	65.6*	50.7*	23.6	70.8*	53.5	28.8	69.4*	51.0	24.7
Lemon	60.42*	61.5*	43.4	63.2	67.0	63.2	65.3	68.4*	62.2	50.7	50.7	45.5	59.7	61.8*	53.5
Refreshing	31.3	31.6	26.7	35.1	41.7*	30.6	31.6	31.3	24.0	32.6	39.2*	31.6	32.6	36.1*	28.1
Sour	32.3*	29.9	23.6	46.5	49.3*	35.4	44.1*	44.1*	32.6	41.3	42.0*	33.3	41.0*	41.3*	31.3
SweetF	19.4	17.4	17.7	14.2	13.5	12.2	21.2	20.8	16.3	9.4	5.6	8.7	16.0	14.2	13.5
SweetT	21.2*	18.1	16.7	13.9	19.4*	15.3	26.4	26.7	22.6	7.3	8.0	7.3	17.4	18.1	15.6

Significant (alpha=0.05) multiple-response hypergeometric tests per cell for period comparison are indicated by an asterisk.

Table 40. Percentages of citations by attribute (rows), product and period (columns) for panel SIM.

	P2			P3			P4			P5			All products		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
Alcohol	38.8	44.3*	33.3	19.5	27.3*	21.3	30.8	33.7*	22.3	25.0	27.5*	20.4	28.1	32.6*	24.2
Bitter	24.2	35.9*	26.7	23.8	29.1*	22.7	21.6	30.4*	25.6	20.4	30.3*	21.8	22.8	31.6*	23.9
Carbonated	71.1*	45.4	27.8	70.9*	44.0	23.8	71.1*	45.4	21.3	73.6*	49.7	29.6	70.9*	46.0	25.3
Lemon	65.9*	60.8	49.1	69.5*	66.7	55.3	67.4*	70.0*	52.8	61.6*	59.2	49.3	66.0*	63.5*	50.9
Refreshing	33.0	34.1	31.9	36.5	33.0	31.2	33.3	34.1	26.0	34.9	39.1*	31.0	35.1	34.7	29.8
Sour	28.9	28.2	22.3	39.7	40.8	33.0	34.4*	29.7	23.8	40.1*	38.4	29.6	35.1*	34.0*	26.7
SweetF	13.6	13.9	13.6	15.3	12.1	14.9	19.1	19.1	13.6	10.9	10.9	8.8	14.4	13.7	12.3
SweetT	16.1	21.6*	18.7	14.5	20.6*	14.9	21.3	30.4*	30.4*	11.6	13.7	14.4	15.4	21.1*	19.3

Significant (alpha=0.05) multiple-response hypergeometric tests per cell for period comparison are indicated by an asterisk.

The four first blocks of the tables allowed us to compare each period to the other ones by product. For P2, period T1 was more often described as *Carbonated*, *Lemon* (RET, SIM), *Sour* and *Sweet* (RET); period T2 was more often described as *Alcohol* (RET, SIM), *Lemon* (RET), *Bitter* and *SweetT* (SIM); and period T3 was more often described as *Alcohol* (RET). For P3, period T1 was more often described as *Carbonated* (RET, SIM) and *Lemon* (SIM); period T2 was more often described as *Alcohol* and *SweetT* (RET, SIM) and *Refreshing* and *Sour* (RET); and period T3 was more often described as *Bitter* (RET). For P4, period T1 was more often described as *Carbonated* and *Sour* (RET, SIM), *Lemon* (SIM); period T2 was more often described as *Alcohol*, *Bitter* and *Lemon* (RET, SIM), *Carbonated* and *Sour* (RET), *SweetT* (SIM); period T3 was more often described as *Bitter* (RET), *SweetT* (SIM). For P5, period T1 was more often described as *Carbonated* (RET, SIM), *Lemon* and *Sour* (SIM); period T2 was more often described as *Alcohol*, *Bitter* and *Refreshing* (RET, SIM), *Sour* (RET). These results showed that despite similar evolutions, some attribute temporal changes were product dependent (*SweetT*, *Refreshing*, *Sour*).

Panel RET highlighted about the same number of differences (34) as panel SIM (32), but some conclusions were different depending on the panel, some significant attributes switching by a period between RET and SIM.

3.2.4. Feedback of the consumers about the task

The analysis of the feedback on the simultaneous task showed that 29 consumers found the task easy, and seven found it difficult. However, some consumers complained about "the pace of rating being too fast", "the response time being too short", and "the time to select an impression being too long", and some found it "a little difficult to drink while being aware of the difference in change" or were "distracted by timing", "pressed for time", and "wanted slightly more time". Other expressed limitations found in panel SIM included "too fast," "I'm in a hurry," "a little quick," and "difficult timing".

The analysis of the feedback on the retrospective task showed that 28 consumers found the task easy, and eight found it difficult. Only two consumers reported verbatim statements related to memory: "forgot the taste" and "difficult to match the timing".

4. Discussion

4.1. Statistical methodology

The sensory trajectories, between-sip, between-product and between-period comparisons have been performed using the mrCA paradigm, proving its versatility. The approach proposed here is based on a step-by-step strategy of comparison as in Mahieu, Visalli, Thomas, et al. (2020). It has the advantage of clearly identifying the sources of differences. In this study, the differences between the products were larger than the differences between the periods, and there was no difference between the sips. Several limitations can be noticed. As there was no observed difference between the sips, the combination “consumer x sip” was considered as an observation, artificially increasing the degrees of freedom. It is fairly common when drawing TDS or TCATA curves, but possible alternatives would have been to sum or average the citations over the sips, however, in this case T-Table 40 could not have been interpreted in terms of citation rates. The mrCA is by nature multidimensional; thus, it takes dependence between the attributes into consideration, but it does not allow us to statistically compare the magnitudes of the citation rates of the attributes between them. It does not allow us to test whether a citation rate is large enough to be considered important or applicable at the panel level. Using a generalised linear model such as in Visalli et al. (2020) or Weerawarna et al. (2021) could have also been considered as a univariate alternative.

4.2. Retrospective vs. concurrent measurement

The two panels (RET and SIM) were able to temporally discriminate the products. As in Visalli et al. (2020), retrospective and concurrent temporal measures allowed us to draw relatively similar conclusions. This was congruent with the conclusions of Van Den Haak et al. (2003) but contrary to the observation of Liang et al. (2018) about the decrease in taste sensitivity due to the memory load that was not observed with panel RET. However, this decrease was observed for intensities, not for citation rates.

However, some differences have been observed. The RET panel highlighted more differences between the products, as observed in different scientific contexts by Kuusela & Paul (2000); Ryan & Haslegrave (2007); Whyte et al. (2010). Notably, the differences in sweetness between products seem to have been better captured with panel RET. Indeed, the second dimension of Figure 43.A captured the Brix measure gradient (Degrees Brix measures the sugar content

of an aqueous solution): P4 (6.33) > P2 (6.07) > P3 (4.82) > P5 (2.63). However, as the “sensory reality” is unknown, it is not possible to conclude the superiority of retrospective measures. Indeed, the differences were small, and no repeatability measurements were performed to quantify the level of noise. Thus, it was not possible to conclude if these differences were due to randomness or to the task, and only hypotheses can be formulated. No definition of the descriptors was given to the consumers. It is therefore possible that some sensory terms have been interpreted differently depending on the context. Indeed, differences were observed with Refreshing in Table 37 (RET) but not in Table 38 (SIM). The consumers in panel RET could have interpreted the descriptor as “thirst-quenching”, whereas, as it was evaluated earlier by the consumers in panel SIM, it could have been interpreted as something related to the temperature or the dryness of the drink.

The SIM panel had limited duration and allowed us to check the applicable attributes simultaneously to the tasting; it could have increased the difficulty of the task. The experimenters pre-tested the evaluation time of eight seconds for periods T1 and T2 and found it well adapted to this type of carbonated beverage. Keeping the beverage in mouth for eight seconds was even found quite long, but the experimenters wanted the in-mouth and after-swallowing evaluation times to be the same. As there was no difference in the number of citations between the two panels, it can be supposed that consumers from SIM panel had sufficient time to check the attributes. In addition, for panel SIM, the number of different descriptors used tends to decrease over sips. It is thus possible that the cognitive load was increased by the simultaneous task and influenced perception, as reported by Wal & Dillen (2013). To reduce the cognitive load and/or the boredom, the consumers could for example have reported dominant sensations instead of applicable ones.

Most of the discriminant attributes were the same regardless of the panel, but some appeared more important at different periods. As SIM was expected to be more temporally accurate than RET, it can be supposed that the switches of periods between the two panels could be due to temporal inaccuracy due to the memorization process involved in the retrospective choice of the period. This could corroborate the conclusions about the memory effect identified by Daniel and Katz (2018) and Aldrovandi et al. (2015). However, the switches of descriptors between panels and periods went both ways, e.g. *SweetT* for P2 from T1 in RET to T2 in SIM, and *Sour* for P5 from T2 in RET to T1 in SIM, suggesting that the changes were

related to variability rather than to the method. Moreover, if the consumers of panel RET had randomly chosen the periods for the descriptors, it is unlikely that a better discrimination would have been observed. One of the most notable differences was observed with *Bitter*, which was most cited for products P2 and P3 during period T3 by panel RET and in period T2 by panel SIM. For this descriptor, RET seems to be in agreement with previous studies. Dietz et al. (2022) showed that bitterness was more important at the end of the tasting, and Higgins et al. (2021) demonstrated that the bitterness peak was after 20 seconds. This observation remains to be confirmed because the products (beers) were not flavoured in the mentioned studies.

The analysis of consumers' feedback did not allow to conclude about the difficulty of the task. However, it should be noted that in this study, the evaluation times were relatively short. Retrospective evaluation could be unsuitable for products having lingering tastes or flavours for longer durations, which could cause difficulty in memorizing those sensory perceptions.

Regarding these results, it can be assumed that concurrent and retrospective measurement could align with Kahneman's theory (O'Brien, 2012) about two systems of thinking: the first one being fast, intuitive and emotional; the second one being slower and more deliberative; each one having its own advantages and drawbacks for developing consumer methodologies. The most appropriate measure probably depends on the objective. The concurrent measure was the most spontaneous and likely the most suitable for studying physiological processes involved in temporal perception. The retrospective measure was presumably the most cognitive and analytical, but also the closest to the remembered experience and thus could be more important to explain later consumer choices. Outside of food science, the superiority of remembered experience in predicting choices was observed by Kahneman et al. (1993) and Wirtz et al. (2003). If no similar study has been conducted with food choices, the importance of cognitive processes for determining answers to food cues was largely studied as demonstrated in the review of Higgs (2016). However, in this study, the retrospective measure was made just after the tasting, and it could be interesting to redo the experiment with different delays after the tasting to study what is really recalled of the products and how it impacts consumers' choices.

4.3. Periods and temporality

Retrospective measurement imposes temporal measurements by period. In Visalli et al. (2020), the chosen periods were quite subjective, as “attack”, “evolution” and “finish” were not precisely defined and could evoke different moments between the panellists. Here, this potential subjectivity was limited using specific and meaningful tasting points to delimitate the periods: “in mouth before swallowing,” “immediately after swallowing” and “aftertaste.” This study, as well as those of Visalli et al. (2020) and Mahieu, Visalli, Thomas, et al. (2020), proved that measurements by period captured temporal differences within the product. Indeed, the two protocols used in this study could be seen as “discrete time TCATA” with the advantage of no need for fading (Ares et al., 2016) to help the subjects to unselect the applicable attributes and with the possibility to include a larger number of descriptors (only with the retrospective protocol). Using periods, the temporal precision was expected to be lower than that of a continuous time measurement. However, the temporal precision of TDS and TCATA has not been extensively studied. In this way, it should be interesting to test the variability of the continuous time-dependent conclusions, for example, by using replicates or bootstrap, and to see if continuous time measures truly brought additional usable information compared to periods. In any case, if the temporal precision is not of crucial importance or is not the main source of difference between the studied samples, the discrete time TCATA methods (both concurrent and retrospective, but particularly retrospective) should be considered as useful temporal measurements in consumer studies.

5. Conclusion

This study compared the temporal description obtained from consumers at home in two conditions: concurrently with the tasting (panel SIM) and retrospectively with the tasting (panel RET). Consumers were asked about the applicable attributes for four full portions of lemon-flavoured carbonated alcoholic drinks at three specific time periods: “in mouth before swallowing”, “immediately after swallowing”, and aftertaste. Data were analysed using multiple-response correspondence analysis framework applied at different levels: product, period and sip. RET and SIM captured the differences between products and periods, with the differences between the products being larger than the differences between the periods. In both methods, no sip effect was observed. Perception of the products were identical in sips 1, 4 and 7. Overall, the consumers of RET and SIM agreed on the relative between-product

differences, with RET showing more differences than SIM. The consumers of RET and SIM only partially agreed on the within-product differences. The results suggest that the retrospective temporal evaluation could better discriminate the products and point to the need for more research about the temporal precision of the data.

III.1.4. Relevance of free comment to describe wine temporal sensory perception: An application with panels varying in culture and expertise

*This article was published in Food Quality and Preference:
<https://www.doi.org/10.1016/j.foodqual.2022.104785>*

J.M. Cardebat: Funding acquisition, Review & Editing; M. Dubois: Conceptualization, Data Curation, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing - Review & Editing; N. Georgantzis: Writing - Review & Editing; F. Ric: Writing - Review & Editing; P. Schlich: Conceptualization, Methodology, Writing - Review & Editing; M. Visalli: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization.

Wine is a complex product, and numerous sensory evaluation methods have been tested to characterize it. Among these, new sensory analysis evaluation methods have been developed to allow consumers to describe products using their own vocabulary. Recently, Free Comment Attack-Evolution-Finish (FC-AEF) was introduced to add the temporal aspects of tasting to the free description of wine. This method has been rarely used thus far, but as FC-AEF does not influence consumers by presenting them with predefined lists of attributes, it is of special interest to study the semantic aspects of sensory perception related to expertise and culture. FC-AEF was used to collect temporal data about two Bordeaux and two Rioja wines. Three panels of participants evaluated the products: French consumers at home (n=106), Spanish consumers at home (n=98), and international wine students (“connoisseurs”) in a sensory lab (n=47). Textual data were processed to extract relevant groups of sensory words and then used to characterize wines with and without considering the temporal aspects of tasting. The three panels were able to discriminate between the two Rioja and the two Bordeaux wines; however, only the French panel perceived differences between the two Bordeaux wines. Only the connoisseurs’ panel succeeded in perceiving temporal differences within the wines. However, the panels disagreed on the nature of the sensory differences between the samples. The impact of expertise on the ability to describe wines seems quite clear and in favour of experts. It is more difficult to conclude the impact of culture, as little agreement was observed in the wine descriptions of the consumer panels. An original framework based on a combination of semantic and statistical criteria was used to extract relevant sensory information from the wine comments. This article provides some methodological answers to challenges raised by the analysis of free comment data and its application for the semantic investigation of differences in perception observed in wine descriptions collected with different panels.

1. Introduction

Wine is a complex product, and describing its perception is a question that dates back to antiquity (Shapin, 2016). Amerine & Roessler (1976) were pioneers in the sensory evaluation of wine, being among the first to attempt to transform subjective judgements into objective descriptions (Shapin, 2016). Since then, many sensory analysis methods have been developed (Lesschaeve & Noble, 2022). Traditional descriptive methods used with trained panels rely on the use of a specific list of sensory attributes based on standardized descriptive vocabulary. These methods are time-consuming and thus money-consuming due to the duration of the training to reach a consensus about the use of the attributes (Sáenz-Navajas et al., 2016). Moreover, some authors have reported limitations about the use of such predetermined lists for wine characterization. Perrin & Pagès (2009) highlighted the risk of forgetting important attributes. Lawrence et al., (2013) wondered if imposing attributes could lead the panellists to select attributes that do not exist in the evaluated sample.

Thus, new sensory analysis evaluation methods have been developed that allow untrained panellists to describe products using their own vocabulary (Sáenz-Navajas et al., 2016). Among these methods, free comment (FC) has been used in several studies related to wine. Sauvageot et al. (2006) used FC to make an expert panel describe Chardonnay wine qualities. They concluded that the method required in-depth examination to become a useful technique in food science. Bécue-Bertaut et al. (2008) proposed a statistical framework to analyse FC data and applied it to consumers' descriptions of Spanish wine. They concluded with the interesting potential of FC. Lawrence et al. (2013) used FC to make a professional panel characterizing Cabernet Franc wines. They highlighted several advantages of FC over methods using a list of attributes: the method is faster and allows new information and specific characteristics of wines to be introduced. The second assertion was confirmed by Coulon-Leroy et al., (2017), who used Mixed Profiling, a method combining a predefined list of attributes with FC, to characterize Cabernet Franc and Gamay wines with wine professionals. Schüttler et al. (2015) used FC with wine experts to characterize the typicality of Riesling wines and found conclusions in agreement with former sensory studies. Vidal et al. (2015) used FC to show that consumers accurately understand the meaning of wine astringency using a limited vocabulary, whereas the use of terms extracted from a mouthfeel wheel was not relevant. Mahieu et al. (2020a) used FC and Check-All-That-Apply methods with consumers to characterize French wines and

concluded that there was better product discrimination using FC. FC does not consider temporality in the perceived sensations. However, it is common for wine experts to describe their perceptions as a temporal “attack-evolution-finish” sequence (Grainger, 2009; Harrington, 2008; Osterland, 2012; Spence & Wang, 2018). This sequence can be measured with a predefined list of descriptors using the Attack-Evolution-Finish (AEF) method described in Visalli et al. (2020). AEF was adapted to replace the list of attributes by FC using the so-called Free Comment AEF method (FC-AEF) proposed by Mahieu et al. (2020). FC-AEF provided additional information compared to AEF; however, to date, it has only been used in one consumer study to evaluate dark chocolate.

FC and FC-AEF do not require training, so the methods can be used both with consumer and expert panels. However, these methods, like all verbalization tasks, are intrinsically related to semantics. Many authors have reported an influence of the level of expertise of the tasters on verbalization tasks related to wine. Lehrer (1975) found that nonexperts generate useless descriptions due to the lack of consensus in word applications. Lawless (1984) demonstrated that experienced subjects used more terms with concrete references to describe wines. Hughson & Boakes (2002) suggested that differences in performance between novices and experts were due to a lack of vocabulary. Croijmans & Majid (2016) concluded that experts have only a limited and domain-specific advantage in communicating about flavours. Sáenz-Navajas et al. (2016) reported difficulty for consumers in describing wines. Croijmans et al. (2019) showed that wines were described differently by experts and by novices, with experts being more in agreement but using more specific source-based terms compared to novices. Bianchi et al. (2021) suggested that such differences can also be due to different understandings of wine lexicons. In addition to expertise, cultural aspects can affect wine description. Sáenz-Navajas et al. (2013) showed a cross-cultural agreement in wine discrimination based on aroma properties of Spanish and French trained panels, but this agreement was not confirmed for the in-mouth description. More generally, several authors (Ares, 2018; Rodrigues & Parr, 2019) pointed out the need for methodological attention with regard to these aspects.

This article presents analyses of part of the data described in Visalli et al. (2022). It focuses on the FC-AEF data describing two Bordeaux and two Rioja wines collected using consumer panels (France and Spain) and an international panel of connoisseurs. The methodological

aspects of the use of FC-AEF for wine characterization under different conditions are studied. The objective of this study was to demonstrate that data collected using free comment are usable to describe complex products including a temporal dimension and provide interesting semantic information to understand the difference in perception between panels.

2. Materials and methods

2.1. Protocol

Below is a summary of the protocol. A detailed description of the protocol can be found in the data paper, “A dataset on the sensory and affective perception of Bordeaux and Rioja red wines collected from French and Spanish consumers at home and international wine students in the lab.” (Visalli et al., 2022).

2.2. Participants

Three panels were recruited. The first panel (WC panel) was composed of wine connoisseurs. Fifty students of the School of Wine and Spirits Business of Burgundy School of Business in Dijon were recruited through a mailing. They were native from different countries and selected based on their knowledge about wines (they were at least holders of WSET certificate (Wine and Spirits Education Trust) Level 2 Award in Wines). The second and third panels were composed of regular wine consumers. One hundred fifty French consumers (FR panel) and 150 Spanish consumers (SP panel) were recruited from a database belonging to Qualtrics, a panel recruitment agency, through online questionnaires. The selection criteria included (i) being available to participate in a 20-minute online study involving the at-home tasting of four wines; (ii) having consumed red wine at home within the past month; (iii) possessing at least one wine glass at home; (iv) owning a computer with a webcam and an internet connection (Chrome, Firefox or Edge browser); (v) agreeing to be video recorded during the tasting; and (vi) agreeing to provide a postal address for the shipment of samples. Quotas on individual characteristics (age, gender, employment status, education level, family income) were established to balance the two consumers panels.

The research was carried out following the Declaration of Helsinki and received approval from the Burgundy School of Business Research Ethics Committee (reference of the application: CERBSB2022-9). The participants were informed of the conditions and validated an informed consent form. They received a compensation worth 20 euros (either 20€ through Lydia or, if

they won the experimental auction, the bottle of wine and 20€ minus the price of the bottle; see data paper for more details).

The consumers in the FR and SP panels were both divided into three groups of 50: the first had no other information than the label (group “no rating information”), the second received information about expert ratings (group “expert rating information” with scores from the Wine Advocate), and the third received information about peer ratings (group “peer rating information” with scores from Vivino).

2.3. Samples

Four non-organic red wines produced in France or Spain in 2018 were tasted in this study. They were all in the same price range, between 14 and 16 euros for a 75cl bottle. W1 (winery: Château Féret Lambert, composition: Merlot 90%, Cabernet 10%, alcohol content: 14.5°) was an AOC (Appellation d’Origine Contrôlée: protected designation of origin) Bordeaux Supérieur, W2 (La Louvière, Merlot 40%, Cabernet Sauvignon 60%, 13.5°) was an AOC Pessac Léognan, W3 (Bhilar, Tempranillo 85%, Grenache 10%, Viura 5%, 13.5°) and W4 (Miguel Merino, Tempranillo 100%, 14.5°) were Rioja DOP (Denominación de Origen Protegida: protected designation of origin). In order to ease the logistics of home-use tests, the wines were repackaged in 20 ml PET screw bottles in an inert atmosphere devoid of oxygen to avoid the risk of oxygenation and ensure the preservation of the organoleptic qualities of the wines (patented technology). The screw bottles were labelled with original information (origin, designation, producer, vintage, alcohol content) displayed in a standardized way, then shipped by postal mail to the participants of the consumer panels (FR, SP).

2.4. Experimental procedure

The participants received an email containing an individualized URL to invite them to connect to the TimeSens version 2 web application (INRAE) using a web browser (Chrome, Firefox, or Edge were recommended to ensure maximum compatibility with the web app) on their computer. The FR and SP Panels completed the experiment at home; the WC panel completed the experiment in the sensory lab of the Burgundy School of Business (Dijon, France).

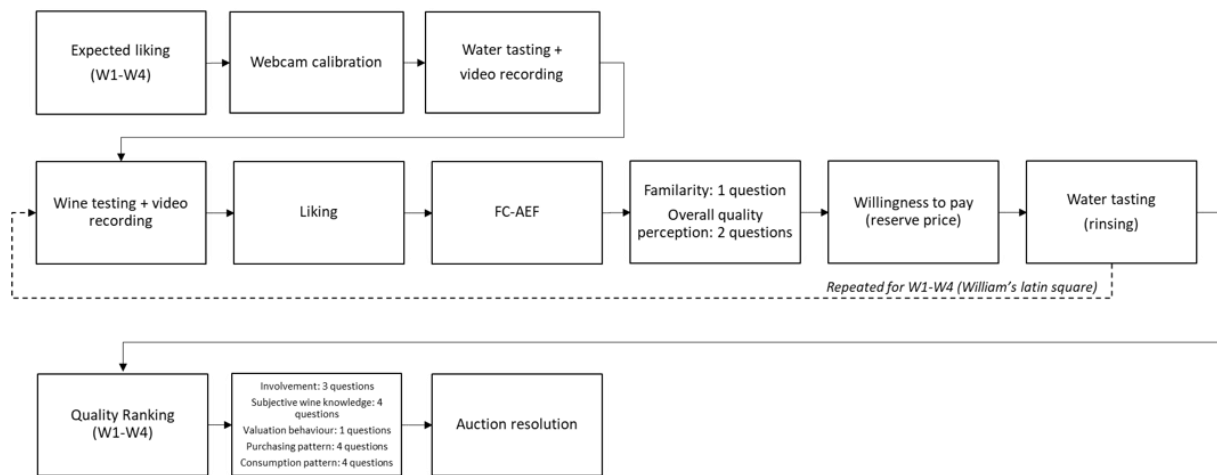


Figure 46. Summary of the experimental procedure.

Figure 46 summarizes the experimental procedure for the three panels (see data paper for more details). This article specifically focuses on the results from wine descriptions using FC-AEF; other results will be presented elsewhere.

Expected liking and water tasting

The participants first had to rate their expected liking before tasting the four wines using a 7-point hedonic scale. Then, they were invited to calibrate their webcam. They were instructed to pour some water into their wine glass and then to look at the water, swirl and sniff it. Then, they were invited to take a small mouthful and at the same time click on the button to start the video recording. More details about this part of the experimental procedure can be found in the data paper.

Wine tasting

After the water tasting, the participants were instructed to empty their wine glass and then pour the appropriate wine sample into the glass. The order of presentation of the samples was balanced over participants based on a William's Latin square. The participants were instructed to look at the wine and then swirl and sniff it. Then, they were invited to take a small mouthful and at the same time click on the button to start the video recording (it was the only tasting of this wine). The label corresponding to the wine sample they had to taste was displayed on the screen. The participants in the "expert rating information" group had the following supplementary information to the right of the sample label: "This wine has been rated $x/100$ by the Wine Advocate – Robert Parker". The participants in the group "peer rating information" also had supplementary information: "This wine has been rated $x/5$ by

consumers – website Vivino”. After 10 seconds, they had to rate their liking on a 7-point hedonic scale. Then, the FC-AEF task was explained. The participants were informed that they had to retrospectively describe the sensations they perceived in mouth during the tasting in chronological order. Three periods were defined to summarize the tasting: “at first”, “after a few moments” and “at the end of the tasting”. For each period, they had to describe their sensations (tastes, aromas) using their own words. They were informed that they could use the same words in different periods or enter “nothing” if they perceived nothing. A fictive example with chocolate was displayed to help them understand the concept. Then, they had to perform the FC-AEF task for the sample they tasted (see Figure 47).

Wine description

université
BORDEAUX

23/54

Describe the sensations you perceived during the tasting (tastes, aromas) in chronological order. Use your own words to answer. You can use the same word at different times. If you haven't perceived anything for a period, enter "nothing".

At first I perceived this wine to be

Then after a few moments I perceived it to be

And at the end of the tasting I perceived it to be

NEXT

Figure 47. Screen for FC-AEF measurement.

The participants next rated their familiarity with the wine, their appreciation of the wine quality, and their expectation about how other people would like the wine using a 5-point Likert scale. Then, they had to indicate their willingness to pay for a 75-cl bottle of the tasted wine. Finally, they were instructed to empty and rinse their glass, and the procedure for wine tasting was repeated for the three other wine samples.

Wine ranking and BDM auction resolution

After having tasted the four wines, the participants had to rank them, with wine ranked first being the most qualitative and wine ranked fourth the least qualitative. Then, several questions were asked about their involvement in wine, their subjective wine knowledge, their consumption pattern, their purchasing pattern, and their valuation behaviour. Finally, one of the wines was randomly drawn for each subject, and the auction was resolved.

2.5. Data analysis

Statistical analyses were performed using R 4.1.0 software (R Core team, 2020).

2.5.1. Characterization of panels

The main characteristics of participants were reported within each panel (number, gender, age). The mean scores of variables likely to explain differences in terms of descriptive ability between panels (product category involvement, subjective wine knowledge, consumption pattern, and purchasing pattern) were computed, and differences between panels were evaluated using a one-way ANOVA model. Tukey LSD post hoc tests were performed when the panel effect was significant ($\alpha=0.05$, means followed by the same letter were not significantly different).

2.5.2. Selection and grouping of words, by panel

FC data were processed by a custom program. First, all French and Spanish wine descriptions were translated into English using Google Translate. Second, the descriptions were cleaned: the letters were transformed to lower-case, and accented characters and punctuation marks were removed. Third, using regular expressions (Erwig & Gopinath, 2012) associated with a custom lexicon of words related to sensory perception, words or groups of words were extracted from descriptions and lemmatized to singular form. Fourth, quantifiers associated with the extracted lemmas were searched and summarized as “not”, “high” or “low”. Example: “Fruité, acidité et équilibré, peu tannique, de belle rondeur.” “Fruity, acidity and balanced, low tannin, nice roundness”, “fruity acidity and balanced low tannin nice roundness”, “fruity acid balanced tannic round”, “fruity, acid, balanced, tannic_low, round”.

The results obtained with the program were manually checked and corrected by the authors of the article. Only words related to objective sensory perception (taste, flavour, mouthfeel) were retained.

The number of different words used by each panel was counted, and the average number of words used per product/participant/period was calculated. The differences in citations were evaluated using two ANOVA models. The first one (i.e. *Number of words = Participant (nested in Panel) + Panel + Product + Period + dual interactions + error*) was carried out to test whether the number of words used by period to describe the perception depended on the panel, the wine, or the period. The second one (i.e. *Total number of different words = Participant (nested in Panel) + Panel + Product + Panel x Product + error*) was carried out to test whether the total number of distinct words (all periods combined) used to describe the perception depended on the panel or the wine. Tukey LSD post hoc tests were performed for significant factors ($\alpha=0.05$, means followed by the same letter were not significantly different).

For each panel independently, the lemmas related to sensory words remaining after data pre-treatment were aggregated into groups of sensory words based on an ascending hierarchical classification using the methodology presented in Greenacre (1988) and applied to FC as in Mahieu et al. (2020). This step ensures a grouping of words having the same “semantic meanings” and the same “statistical properties”. All words in the same group were thereafter considered equivalent and interchangeable for a given panel in the context of the study. To facilitate reading, the groups of words were identified by short labels representative of the sensory concept. Only groups of words cited by at least 5% of the participants for one wine were kept.

2.5.3. Between and within product comparisons, by panel

The usual χ^2 framework was not well suited to analyse multiple-response collected with free comment (Mahieu et al., 2021). To overcome this limitation, Mahieu et al. (2021) developed a framework and an R package (MultiResponseR, Mahieu, 2021) for the multidimensional analysis of contingency tables derived from FC data. This framework presents a lot of features for the purpose of product comparisons within each panel. Indeed, it includes a test of dimensionality, a test for pairwise product comparisons, and a test of significance of product by descriptor associations.

Two types of analyses have been carried out for each panel: (a) between products, to compare the overall (all periods combined) differences of product characterization, and (b) within product, to compare the temporal differences of product characterization between the periods (attack, evolution finish) for a given product. For (a), three contingency tables (one by panel: “FR”, “SP”, “WC”) were computed, the row variables being the four products (“W1” to “W4”) and the column variables being the group of words (the groups of words cited in several periods by the same participant counting as one). For (b), 12 contingency tables (one for each panel x product: “FR-W1” to “FR-W4”, “SP-W1” to “SP-W4” and “WC-W1” to “WC-W4”) were computed, the row variables being the three periods (“T1”, “T2” and “T3”) and the column variables being the group of words. Then, the differences between the row variables were tested using the procedure summarized below.

(i) The dimensionality of the dependence between the row and column variables of the contingency tables was tested using dimensionality tests (2000 simulations) based on the multiple-response χ^2 framework.

(ii) When at least one dimension was significant ($\alpha=0.05$), the multiple-response correspondence analysis (mrCA) was computed on the contingency tables.

(iii) The outputs of mrCA were displayed using a standard biplot, and 95% confidence ellipses were computed with a total bootstrap procedure (Cadoret & Husson, 2013) with 2000 simulations. Procrustes rotations were performed on the significant dimensions.

(iv) For each pair of row variables of the contingency tables, total bootstrap tests were performed on the significant dimensions to assess the significance of the difference. When significant ($\alpha=0.05$), multiple-response hypergeometric tests (2000 simulations, $\alpha = 0.05$) were performed to test if the descriptors were cited in a proportion that was significantly greater than the overall average citation proportion. The results of the multiple-response hypergeometric tests are summarized in tables with cells representing the citation rates.

The similarities between the three product configurations obtained with between products mrCA (a) were compared and tested using the RV coefficient (Robert & Escoufier, 1976) and standardized RV coefficient (Josse et al., 2008) for each pair of panels. The tables of citation rates (rows being the group of words and columns being the product) obtained as a result of the hypergeometric test were vectorized, and then correlations between “FR”/“SP”,

“FR”/“WC” and “SP”/“WC” were computed (keeping only the matching group of words) to test the similarity of the descriptions. Finally, the differences in conclusions obtained from the three panels were qualitatively described and interpreted.

3. Results

3.1. Characterization of panels

Table 41 summarizes the characteristics of participants by panel. Forty-seven (94%) students completed the study in the WC panel, 106 consumers (70%) in the FR panel, and 98 (65%) in the SP panel. Due to dropouts, the two consumer panels were finally not balanced in their compositions.

Table 41. Characteristics of participants, by panel.

	FR, n=106	SP, n=98	WC, n=47
Age: 18-40	28%	39%	79%
Age: 41-60	35%	45%	15%
Age: 61+	37%	16%	6%
Gender: male	49%	48%	36%
Gender: female	51%	52%	62%
Gender: other	-	-	2%

As expected, the mean scores were similar between the consumer panels and different from the WC panel for product category involvement (FR: 4.07, a; SP: 4.18, a, WC: 4.73, b), consumption pattern (SP: 2.09, a, FR: 2.15, a, WC: 2.47, b) and purchasing pattern (SP: 3.25, a, SP: 3.29, a, WC: 4.38, b). However, unexpected differences were observed for subjective wine knowledge (FR: 3.08, a, SP: 3.44, b, WC: 3.68, b). The participants in the WC panel were more involved (interest in wine, discussion about wine, pleasure to buy wine) and had a greater budget dedicated to wine, but they consumed wine less frequently compared to the consumer panels. The participants in the WC and SP panels declared being more knowledgeable about wine than the participants in the FR panel.

3.2. Selection and grouping of words, by panel

One hundred thirty-two different words related to objective sensory description were used by the WC panel, 121 by the FR panel, and 105 by the SP panel (177 distinct words, including 36 hapaxes).

Table 42. Average number of words used per participant to describe wine samples, grouped by panel (WC, SP, FR); product (W1 to W4); and period (T1 to T3).

	W1				W2				W3				W4			
	T1	T2	T3	ALL	T1	T2	T3	ALL	T1	T2	T3	ALL	T1	T2	T3	ALL
WC	2.02	2.00	1.88	4.79	1.85	2.19	1.88	4.74	1.70	1.84	1.98	4.49	2.02	2.04	2.36	5.19
SP	1.22	1.27	1.19	2.77	1.19	1.30	1.26	2.88	1.21	1.27	1.22	2.88	1.27	1.32	1.27	3.01
FR	1.21	1.21	1.29	2.50	1.24	1.26	1.19	2.56	1.26	1.23	1.21	2.54	1.18	1.28	1.32	2.61

ALL corresponds to the number of different words in all periods combined.

When considering the number of words by period, the main factors Panel ($F=309.5$, $p<0.001$), Product ($F=2.9$, $p=0.034$), and Participant ($F=5.7$, $p<0.001$) were significant, and the interactions were not. Table 42. Average number of words used per participant to describe wine samples, grouped by panel (WC, SP, FR); product (W1 to W4); and period (T1 to T3). shows that the participants in the WC panel used significantly more words by product/period (1.96, b) compared to the FR and SP panels (1.21, a and 1.24, a). Small but significant differences were observed in the number of words by product: W3 (1.41, a) and W1 (1.45, a) had fewer words by period than W2 (1.47, ab) and W4 (1.55, b).

When considering the total number of distinct words and all periods combined, only the main factors Panel ($F=338.1$, $p<0.001$) and Participant ($F=6.2$, $p<0.001$) were significant. The participants in the WC panel used significantly more words in all periods combined (4.80, c) compared to the SP (2.87, b) and FR (2.50, a) panels.

As a statistical warning, it is to be noticed that many tests were run to make these comparisons and many variables were correlated with each other, thus the overall study-wise error rate is probably inflated. However, the F statistics related to the Panel factor in both comparisons are large enough to interpret the Panel effect as the major source of difference between the number of words used to characterize the perception.

One hundred and five words (59% compared to total words, 74% compared to total words without hapaxes) have been retained and aggregated into 28 groups of sensory words.

Table 43. Groups of words obtained after Greenacre classification and cited by at least 5% of the panel for at least one wine, by panel.

Group label	WC	FR	SP
acid	acid, acid_high, acid_low, fresh, sour	acid, acid_high, acid_low, fresh, sour	acid, acid_low, lively, sour
alcohol	alcohol, alcohol_high, alcohol_low, warm	alcohol, alcohol_high, warm	alcohol, alcohol_high, warm
bitter	bitter	bitter, bitter_low	bitter, bitter_low
body		body, robust, thick	body, character, robust, thick
caramel	balsamic, chocolate		
dry	dry		Dry
earthy	earthy, forest, mushroom		
floral	floral, violet		
fresh			Fresh
fruity_dry	dried_fruit		
fruity_fresh	berry, blackberry, blackcurrant, blueberry, cherry, crunchy, raspberry, strawberry	berry, blackberry, blackcurrant, cherry, grape, raspberry, strawberry	berry, blackcurrant, blueberry, cherry, crunchy, raspberry
fruity_mature	black_fruit, cooked_fruit, jam, maturity, plum, ripe, ripe_fruit		
fruity_overall	fruity, red_fruit	fruity, red_fruit	fruity, red_fruit
herbaceous	herbaceous, bell_pepper		
intense		aromatic, intense, odourous, pronounced, tasty	aromatic, intense, odourous, tasty
light	bland, intense_low, light, thin, watery	bland, fluid, light, not_strong	bland, light, not_body, watery
long	long	long	
pungent	egg, pungent, sulphur	pungent, vinegar	
spicy	pepper, spicy		spicy
spicy_sweet	cinnamon, clove, liquorice		liquorice
strong	not_balanced, strong	strong	strong
sugary		sugary	sugary
sweet	sweet, round	sweet, round	sweet, round
tannic_high	astringent, drying, green, harsh, rough, sticky_tannic, tannic, tannic_high, young	astringent, astringent_high, drying, fleshy, green, harsh, harsh_low, rough, tannic, tannic_high, young	harsh, harsh_low, rough, tannic, tannic_high, young
tannic_low	easy, tannic_low, smooth, soft	easy, rough_low, smooth, soft, tannic_low, velvety	easy, not_astringent, smooth, soft, tannic_low
vegetal	leave, vegetal, vegetable		
woody_aromas	burnt_toast, cedar, cigare_box, cocoa, coffee, oak, smoky, tobacco, vanilla	burnt_toast, cocoa, coffee, oak, smoky, tobacco, vanilla	cocoa, coffee, oak, smoky, vanilla, tobacco
woody_overall	woody	woody	woody

Table 43 shows that some groups of words were relatively similar regardless of the panel: *acid*, *alcohol*, *bitter*, *fruity_fresh*, *fruity_overall*, *sweet*, *strong*, *tannic_high*, *tannic_low*, *woody_aromas*, and *woody_overall*. Among these groups, it is interesting to note that few modulators have been retained (i) because few have been effectively used by the participants, and (ii) because of the disagreement in the use of the modulators. For example, the citation rates of the attributes *acid*, *acid_high*, *acid_low* were highly correlated (WC and FR panels),

so the ascending hierarchical classification aggregated them in the same cluster and therefore considered them as synonymous. It was also true for *alcohol* and *bitter*, but not for *tannic*, with two groups being retained (low vs. high).

WC was the only panel to perceive *caramel*, *earthy*, *floral*, *fruity_dry*, *fruity_mature*, *herbaceous*, and *vegetal aromas*, suggesting an increased ability due to expertise in recognition and/or the verbalization of perceived aromas. However, some aroma dimensions seemed to be perceived in a very generic way dissociated from the sensory terms that are expected to constitute them. *Fruity_overall* was perceived differently from other fruity groups (*fruity_fresh*, *fruity_mature*, *fruity_dry*). It was the same with *woody_overall*, which was perceived differently from *woody_aromas*.

SP used *fresh* differently from WC and SP, who associated *fresh* with *acid*. Only the two consumer SP and FR panels used terms related to *intensity*, *body*, and *sugary* (which was different from *sweet*), possibly denoting a different understanding of these terms.

3.3. Between and within product comparisons, by panel

Figure 48.A shows that the FR panel multidimensionally discriminated between the four wines. W1 has a greater citation rate for *body*, *intense*, W2 for *long*, *strong*, and *tannic_high*. W3 was perceived very differently: it has greater citation rates for *alcohol*, *fruity_fresh*, *sugary*, *tannic_low* but also *pungent*. W4 was perceived in between W2 and W3, it has a greater citation rate for *light*. All wines were perceived as equally *bitter*, *fruity_overall*, *sweet*, and *woody* (overall and aromas).

Figure 48.B shows that the SP panel multidimensionally discriminated the two Rioja wines (W3 and W4) but not the Bordeaux (W1 and W2). W1 has greater citation rate for *body*, *woody_overall*, W2 for *acid*, *dry*. W1 and W2 were also perceived (non-significantly) more *intense*, *long*, *tannic_high*, *woody_aromas*. W3 has greater citation rates for *fruity_fresh*, *light*, *sweet*, *tannic_low*. W4 has a greater citation rate for *alcohol*, *bitter*, *fruity_overall*, *light*, *spicy_sweet*, and *sugary*.

Figure 48.C shows that WC panel multidimensionally discriminated the two Rioja wines (W3 and W4) but not the Bordeaux (W1 and W2). W1 has a greater citation rate for *dry*, *fruity_fresh*, *fruity_aromas*, W2 for *acid*, *long*, W1 and W2 for *tannic_high*, *vegetal*. W3 has greater citation rates for *pungent*, *spicy_sweet*, *strong*. W4 has a greater citation rate for *alcohol*, *bitter*, *caramel*, *earthy*, *floral*, *fruity_mature*, *light*, *sweet*, *tannic_low*.

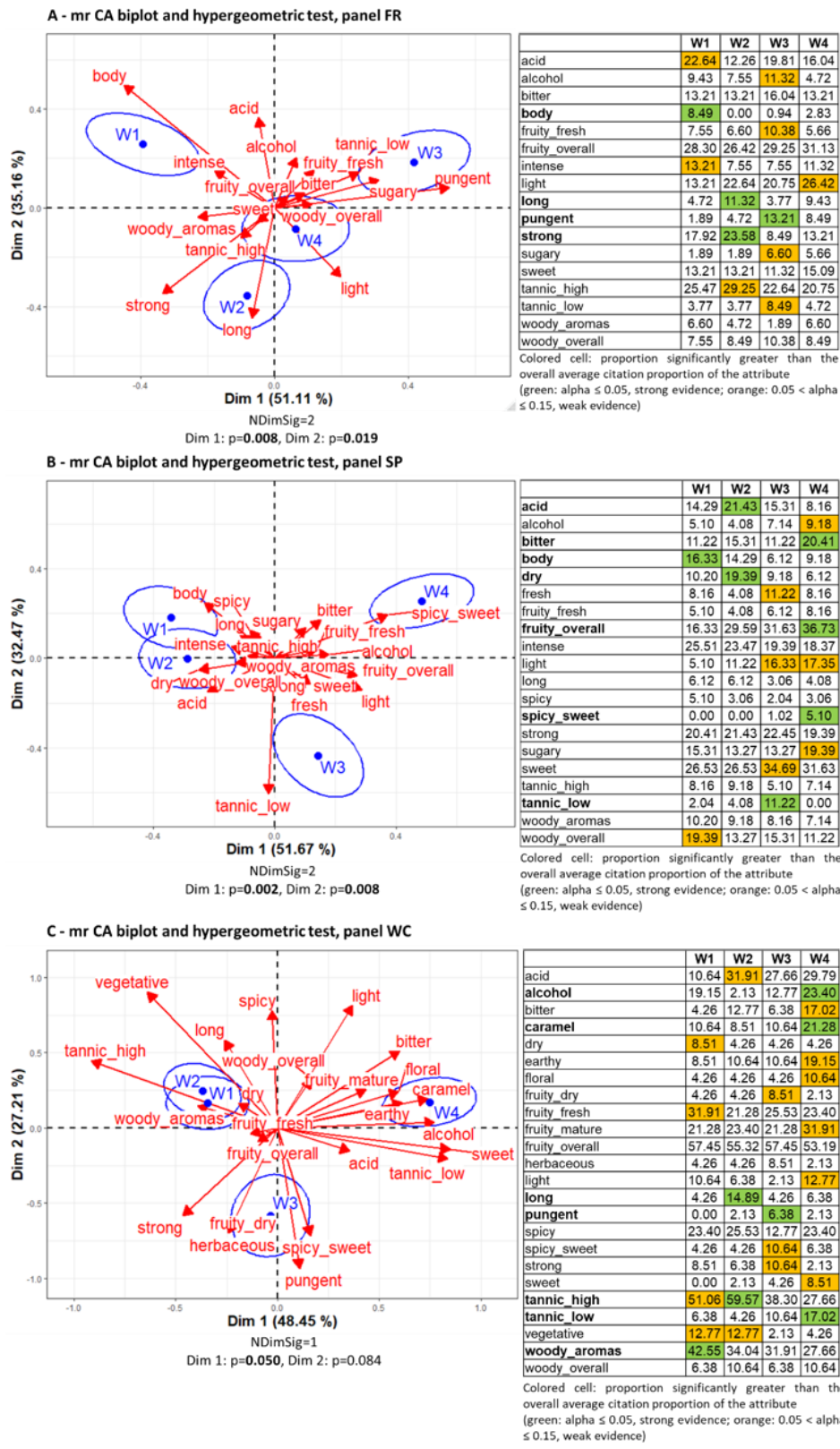


Figure 48. mrCA biplot and table summarizing the results of the hypergeometric test between products. A: FR panel, B: SP panel, C: WC panel. “NDimSig” is the number of significant dimensions. “p” is the p-value of the dimension test. Red arrows indicate the relative importance and correlation between the descriptors and the dimensions. Blue circles represent the 95% confidence ellipses. Products having no ellipse intersection were perceived differently.

Regarding the common characteristics across the panels, the two Bordeaux W1 and W2 were more associated with terms related to medium or full-bodied wines (common with Vivino reviews and Wine Advocate for W1), and W4 was more perceived as a low-bodied, fruity wine (common with Wine Advocate review). Some major differences exist between the panels. SP and WC had greater citation rates for *acid* with W2 (common with Vivino reviews), while FR had greater rate for *acid* and *dry* for W1. SP and WC had greater citation rates for *alcohol* with W4, while FR had a greater rate for W3, W1 having the greater alcohol content. There was no agreement on *dry*, *sweet*, and *sugary*, questioning the use of these words. *Dry* was opposed to *sweet* in WC panel (where Vivino opposed dry to liquorous and sweet to acid), while it was associated with *acid* in SP panel and not used in FR panel. *Sweet* and *sugary* (used only by consumer panels) seem to measure different sensory dimensions. *Light* and *strong* were not fully opposed, which can traduce different clusters of perception or a hedonic use of the words.

The RV coefficient between the product configurations of WC and SP was significant (RV=0.97; RVs=2.03, p=0.04), denoting a great similarity. Indeed, the two panels agreed regarding product discrimination. The RV coefficients between the product configurations of WC and FR (RV=0.78; RVs=-1.27, p=0.79) and SP and FR (RV = 0.80; RVs=-0.46, p=0.62) were non-significant. The FR panel discriminated all product pairs and perceived W2 as closer to W4 than to W1, contrarily to SP and WC which did not discriminate W1 and W2.

The coefficients of correlation between the vectorized citation rates (for group of words common between two panels) were significantly different from zero between WC and FR (r=0.51, p<0.001) and between SP and FR (r=0.39, p=0.001), and non-significantly different from zero between WC and SP (r=0.10, p=0.47).

These results show that even if they perceived the same differences between products, the WC and SP panels did not characterize these differences in the same way. The two consumer panels were closer in terms of general perception and use of terms but disagreed on the characterization of product differences.

Figure 49.A shows that for the FR panel, only W1 has significant differences between periods (thus other mrCAs are not presented). The wine has a greater citation rate for *acid* and *sweet* in period T1 (attack), *intense* and *strong* in period 2 (evolution), and *bitter* in period T3 (finish).

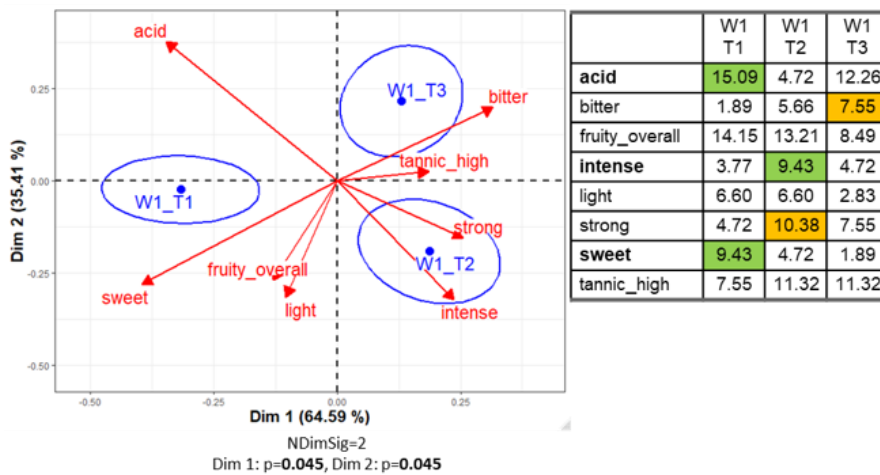
However, the citation rates are quite low, denoting a lack of agreement on temporal perception.

Figure 49.B and Figure 49.C show that, for the SP panel, only W2 and W3 have significant differences between periods (thus, other mrCAs are not presented). As with the FR panel, the citation rates are quite low, denoting a lack of agreement on the temporal perception. There was no obvious product signature, except perhaps for *strong*, which was cited more in T2 (common with FR), and *bitter*, which was cited more in T3 (nonsignificant for W2).

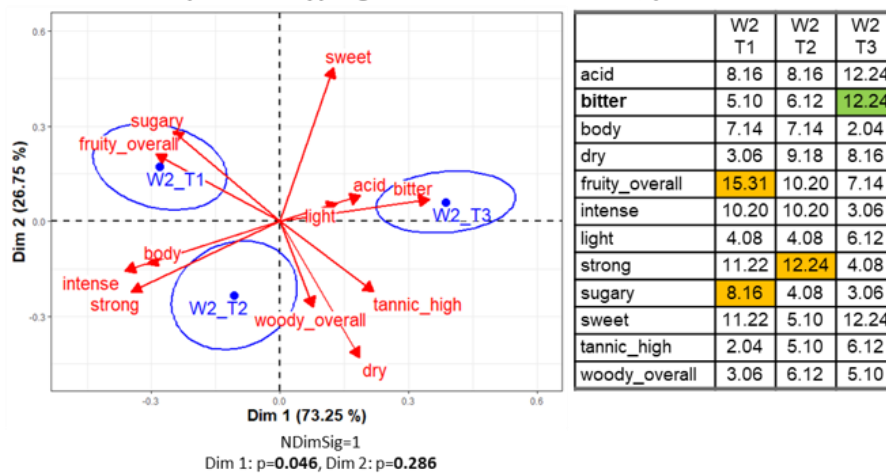
Figure 50.A-D shows that for the WC panel, all periods have significant differences between periods. The citation rates are higher than for the FR and SP panels, denoting a better consensus and thus a better ability to discriminate the periods. *Fruity (fresh, mature, dry, overall)* was cited more in periods T1 (W1, W3, W4) and T2 (W2). *Caramel* was cited more in T2 (W1) and T3 (W3). *Earthy* (W1) and *herbaceous* (W3) were cited more in T3. *Tannic (high, low)* was cited more in period T3 (W1, W3, W4). W3 received more citations of *tannic_high* and *tannic_low*, which seems to be contradictory; however, this disagreement could probably be interpreted as *tannic_medium*. It is interesting to note that W1 and W2 were not discriminated. Thus, additional analysis was run. The results from the WC panel were aggregated into three contingency tables, one by period, the row variables being the four products and the column variables being the group of words. Then, the differences between periods within each panel x product were tested using the procedure described in 2.2.3.

Figure 51.A-C shows that period T1 was the period during which the products were the most discriminated. W1 and W2 were better discriminated in this period, W1 received more citations of *light* and W2 more of *spicy*, *tannic_high*, and *vegetal*. Multidimensional differences were not significant in periods T2 and T3; however, there was a small effect ($p=0.06$) for T3, confirming that when including the temporal aspect, W1 and W2 were perceived differently (finishing more with *alcohol*, *fruity_fresh*, and *woody_ aromas* for W1).

A - mr CA biplot and hypergeometric test, wine W1, panel FR



B - mr CA biplot and hypergeometric test, wine W2, panel SP



C - mr CA biplot and hypergeometric test, wine W3, panel SP

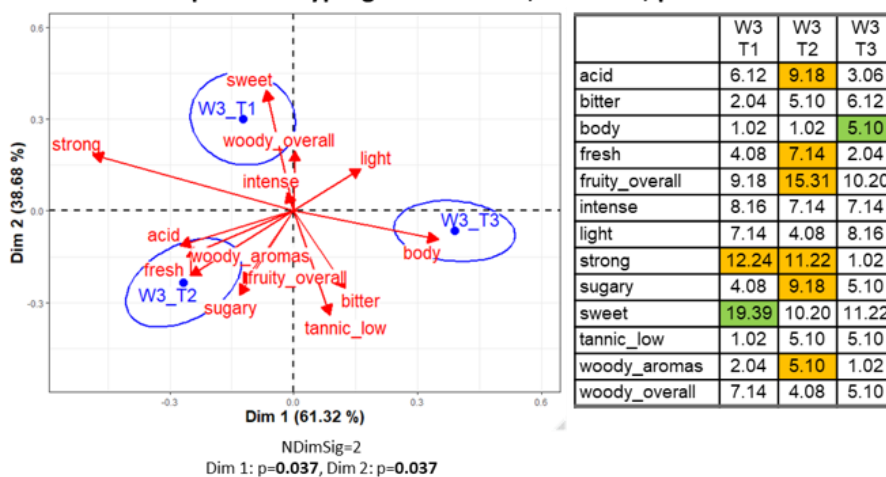


Figure 49. mrCA biplot and table summarizing the results of the hypergeometric test within product (significant mrCA only). A: wine W1, FR panel; B: wine W2, SP panel; C: wine W3, SP panel. “NDimSig” is the number of significant dimensions. “p” is the p-value of the dimension test. Red arrows indicate the relative importance and correlation between the descriptors and the dimensions. Blue circles represent the 95% confidence ellipses. Periods having no ellipse intersection were perceived differently.

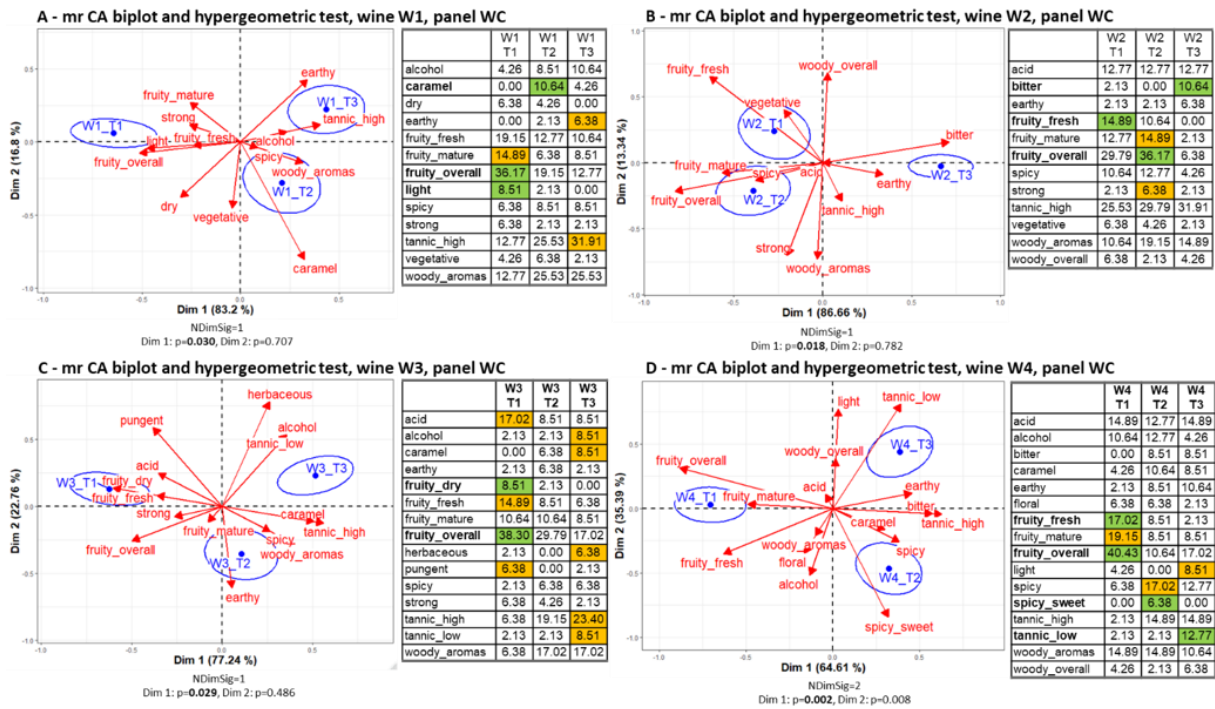


Figure 50. mrCA biplot and table summarizing the results of the hypergeometric test within product for the WC panel. A: wine W1, B: wine W2, C: wine W3, D: wine W4. “NDimSig” is the number of significant dimensions. “p” is the p-value of the dimension test. Red arrows indicate the relative importance and correlation between the descriptors and the dimensions. Blue circles represent the 95% confidence ellipses. Periods having no ellipse intersection were perceived differently.

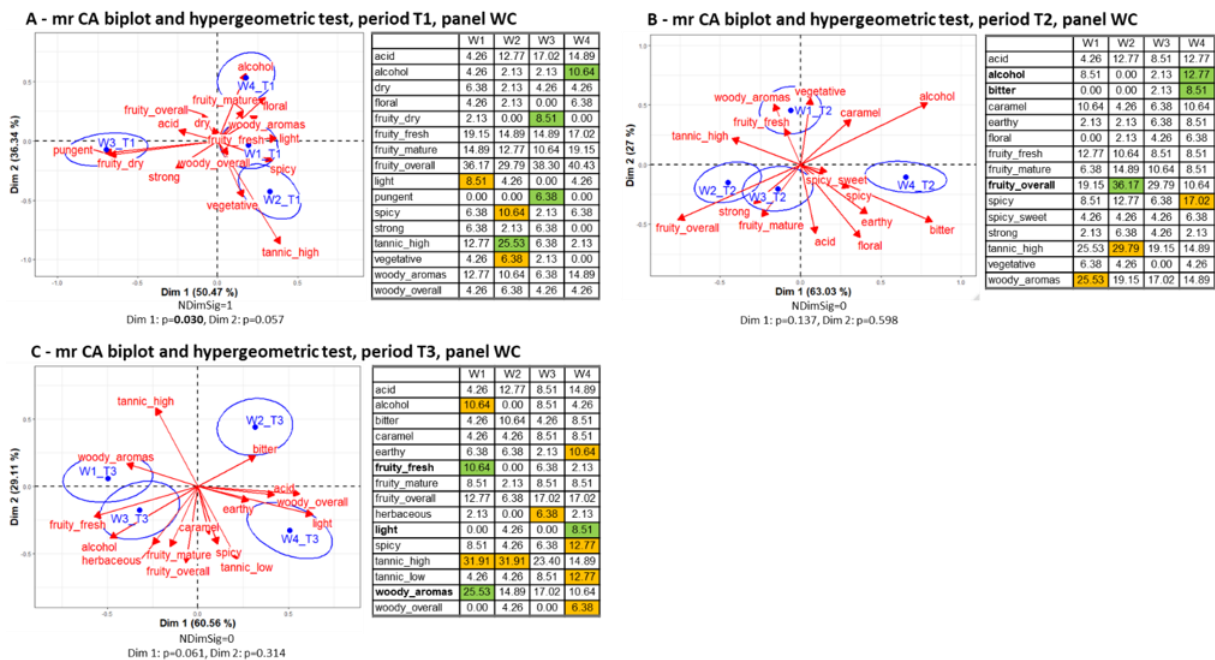


Figure 51. mrCA biplot and table summarizing the results of the hypergeometric test between products for the WC panel. A: period T1, B: period T2, C: period T3. “NDimSig” is the number of significant dimensions. “p” is the p-value of the dimension test. Red arrows indicate the relative importance and correlation between the descriptors and the dimensions. Blue circles represent the 95% confidence ellipses. Products having no ellipse intersection were perceived differently.

4. Discussion

4.1. Effectiveness of FC-AEF as a method for wine description

The objective of this study was to test the effectiveness of FC-AEF to describe wines. To this end, FC data were processed to extract relevant groups of sensory words from consumers' descriptions. These groups of words were then used to discriminate between products with and without considering the temporal aspects of sensory perception. Analysis of FC-AEF data relies heavily on textual data analysis. The methodology used performed well but can still be improved, notably by enriching the lexicons of sensory terms used or by including terms not related to an objective sensory description that were not considered in this study.

The hierarchical classification already includes the possibility for words to be grouped with other words sharing the same meaning but with the additional constraint of being statistically correlated. Examples of words not retained because of the absence of consensual definition and/or classification include for example: "*secondary aroma*", "*clean*", "*patchouli*", etc. Examples of words not retained because they were not classified in the same parent cluster (here *fruity*) include "*banana*", "*peach*", "*apple*", etc. We believe that forcing these elements to aggregate is more likely to add noise than information to the data. Moreover, it is part of the original data analysis strategy which tries to be less subjective than hand-made grouping. As the data are available in Visalli et al. (2022), other data pre-processing approaches can be tested and compared. With the described approach, a larger number of words (more than 100) was generated compared to the study of Hayward et al. (2020), which generated approximately 70 words. The mean number of terms used to describe wines was also slightly greater (more than 2.5 vs. approximately 2.2) compared to the same study.

Many of the aroma terms used by the participants to describe the wines can be found on Noble's wheel of aromas (Noble et al., 1987), which was also observed by Hayward et al. (2020). However, this does not mean that the terms were interpreted in the same way. Indeed, the words were aggregated into 28 groups sharing similar meanings based on the approach of Mahieu et al. (2020). Interestingly, these groups of words do not fully correspond to the classification of the wheel of aromas, notably *fruity*, *woody*, and *spicy*. This confirms the point of view of Candau & Wathelet (2011), who stated that olfactory "categories" are not categories from the point of view of classical and prototype theories. It is also possible that

fruity_overall and *woody_overall* reflect general expectations about wines rather than specific aroma perceptions. *Spicy/spicy_sweet* could either be relative to a generic/specific differentiation or a subclassification of spices. Contrary to the observation of Vidal et al. (2015), *tannic* seems to be consensually perceived (including low vs. high-intensity levels) and verbalized among the panels. However, *tannic* was the only group of words being differentiated regarding levels of intensity. This means that even if semantic differences in intensities were perceived at an individual level, these differences (*acid, bitter, alcohol*) were not consistent, and only presence/absence made sense at the panel level. This may be due to the absence of training of the participants or to the fact that the task of rating was not explicitly required. More generally, the groups of attributes can be seen as the most consensual “meta-descriptors” used to describe perception, which includes the physiological and cognitive aspects of sensory perception. In this sense, these meta-descriptors can be relevant in explaining the sensory dimensions of wines that lack definition, such as body (Runnebaum et al., 2011) and freshness (Morata et al., 2020), or that can be interpreted in different ways, such as light (Hayward et al., 2020).

Both the consumer and WC panels were able to use FC-AEF without familiarization and to discriminate Bordeaux wines from Rioja wines without considering the temporal aspects. The three panels coincided in perceiving the Bordeaux wines as medium- to full-bodied, with Riojas being perceived as lighter and fruitier. Other differences between products were not consensual among the three panels. This could be due to expertise and/or cultural effects (see 4.2), but the two consumer panels presented a relatively low agreement at the panel level. Thus, even if statistical evidence in sensory differences between wines were observed, the results are to be moderated because the citation rates were rather low for the attributes concerned (8 to 23% in FR panel, 5 to 21% in WC panel). As always, questions about the validity and reproducibility of the results cannot be ignored.

Only the WC panel perceived temporal differences during the three phases of tasting (attack, evolution, finish). It is possible that only the WC panel had significant results because some of the participants may have misunderstood the instructions and instead evaluated the wines using more traditional methods (using the olfactory evaluation to describe the attack, the taste and retro-nasal evaluations in mouth to describe the evolution, and the aftertaste evaluation after swallowing to describe the finish). However, this seems unlikely given the

results presented in Figure 51. As Mahieu et al. (2020) succeeded in identifying temporal changes in solid products (chocolates), it may not be related to FC-AEF, as other temporal methods have sometimes also failed (Wang et al., 2021b), but rather to the difficulty of measuring the temporal aspect of a complex product such as wine (Lockshin & Corsi, 2019).

It nevertheless seems reasonable to consider that using FC-AEF to obtain insight into the temporal description of wines requires a minimum of training.

4.2. Effectiveness of FC-AEF as a method for understanding differences in sensory description between different panels

As an example of application, FC-AEF was used for investigating sensory differences perceived by panels varying in culture and expertise. Indeed, FC-AEF allows us to investigate the semantics behind the differences of perception without influencing the answers, avoiding some issues identified by Ares (2018) related to data collection such as the use of scales and “common” predefined vocabulary. For this purpose, the temporal dimension is presumably pointless; however, FC-AEF probably encouraged the participants to focus on the entire tasting experience. Thus, it would be interesting to compare (in terms of richness of vocabulary) the results obtained with the approach of Mahieu et al. (2020) focusing on sensory modalities (vision, texture, aromas) rather than temporal periods of perception.

Overall, the sensory descriptions of the wines differed depending on the panel, suggesting an impact of expertise and cultural factors. Even if the term “expert” may be too strong for describing the student WC panel, the results of this study confirm the superior ability of experienced tasters in wine description. The participants in the WC panel used almost twice as many words in all periods combined compared to the SP and FR consumer panels (4.80 vs. 2.87 and 2.50). Only three groups of sensory words were perceived (for at least one wine) by more than 25% of participants in the FR panel (*fruity_overall, light, tannic_high*) and in the SP panel (*fruity_overall, intense, sweet*), versus eight in the WC panel (*acid, fruity_fresh, fruity_mature, fruity_overall, spicy, tannic_high, woody_aromas*). These results are congruent with the findings of Lawless (1984), Urdapilleta et al. (2011) and Croijmans et al. (2019).

Two limitations must be mentioned. First, the setting effect was combined with the expertise effect, as the students performed the experiment in a sensory lab. Thus, they could have been more focused on the task than the consumer participants, who performed it at home in less controlled conditions. Second, the tasting was not blind; thus, it is not possible to definitely

conclude if expertise increased the sensory perception ability or if knowledge influenced the participants in a way that led to more consensual descriptions regarding the expectations about the wines (e.g. *vegetal* as a typical characteristic of Cabernet Sauvignon grape variety for Bordeaux wines).

Only the FR panel was able to discriminate the two Bordeaux wines. Moreover, each panel identified some specific wine characteristics. These concerns included characteristics being perceived by only one consumer panel, such as *long* or *pungent* (FR panel) or *dry, fresh, spicy* and *liquorice* (SP panel) or differences in the characterization of the same wines (*acid, fruity_overall*). Sáenz-Navajas et al. (2013) previously reported the use of different terminology to describe red wine characteristics by French and Spanish trained panellists. Valentin et al. (2021) showed that providing information on the origin of the wines affected more French than New Zealand participants. However, in this study, the cultural differences should be interpreted with caution, as other factors (in addition to the potential validity issue identified in 4.1) may also explain the differences. The participants in the SP panel considered themselves were more knowledgeable about wine than did the participants in the FR panel, and they used more words to describe the wines. In addition, despite the use of a framework combining semantics and statistics for the determination of relevant sensory groups of words, translation included the usual problems that could have affected the results.

5. Conclusion

FC-AEF was used as a sensory descriptive method for wine characterization. It allowed both consumer and expert panels to discriminate the samples, but describing temporal differences in wines was out of the consumers' reach. As FC-AEF requires participants to express their perception in natural language, it avoids some biases relative to the use of scales and attributes and can be useful for example to study expertise or cultural differences. However, minimal expertise seems to be necessary to obtain robust conclusions, and further studies are required to study the validity of the sensory and semantic differences observed with panels of consumers. The framework used to extract relevant sensory information from FC-AEF could be applied to online consumer wine reviews to compare the differences in conclusions between sensory analysis, wine apps, and wine critics.

III.1.5. Main findings of sections III.1

In the three presented articles, AEF was used as a sensory evaluation method asking consumers to report their temporal perception in different settings (in the laboratory or at home), without familiarization nor briefing other than on-screen instructions. Three variants of AEF were implemented with different measured concepts (dominance or applicability), different modes of data acquisition (list of attributes or free comment), and different types of periods (subjectively determined by the consumers or related to food oral processing steps). In the first two articles, the conclusions obtained with AEF-D and AEF-A were compared to those obtained with concurrent methods, TDS and discrete time TCATA, respectively. In the third article, the conclusions obtained with FC-AEF-A from different types of panels having various levels of expertise regarding the evaluated products were compared.

Inherently, the descriptions of within-product temporal evolutions were less accurate with AEF-D than with TDS (III.1.2). However, when reducing the TDS continuous temporal signal in three periods, the most cited attributes were the same in each period whatever the method, and AEF-D discriminated between products as well as TDS (III.1.2). These results raised questions about the granularity of temporal data, as they suggest that considering the temporality over only three periods could be enough to capture the main differences between products.

To better understand the nature of temporal differences in perception using concurrent and retrospective methods, the number of evaluation points over time was set at three in AEF-A and D-TCATA, and the periods were defined in the same way: “in mouth before swallowing”, “immediately after swallowing” and “aftertaste” (III.1.3). Data were analysed product-wise (to identify in each period the descriptors best characterizing the products) and period-wise (to identify the temporal sequence of perception of attributes within a product). The differences observed between the products were larger than the differences between the periods, themselves larger than the differences between the intakes. AEF-A discriminated between products better than D-TCATA. Although the most cited attributes were the same, they were not always cited during the same periods. This showed that the moment of the measurement (concurrent or retrospective) can influence how consumers report their temporal perception (III.1.3).

When describing complex products using FC-AEF-A, it was shown that although all panels were able to discriminate between samples, a minimum level of expertise was required to get information relevant to the temporal aspects of products (III.1.4).

Given the results observed in the two first articles, it can be considered that the retrospective measurements present face validity because they made it possible to describe and discriminate between products in a manner relatively similar to concurrent methods. Furthermore, it appears that retrospective methods simplify the collection and analysis of temporal data without causing a significant loss of temporal information (III.1.2 and III.1.3), except for the most complex products (III.1.4).

These results are encouraging, and it might have been tempting to stop there and conclude on the validity of the new methods, as most research has done (I.3.2, part 3.5). However, these findings are subject to interpretation potentially biased by many factors, including experimenter expectations. Indeed, it has been shown that discrimination between products mostly depended on non-temporal aspects, and that it was possible to conclude on temporal differences despite small effect sizes which may be non-reproducible (III.1.4). It was therefore important to study in depth all aspects related to the performances of temporal methods when used with consumers, and not just discrimination.

III.2. Performances of rapid temporal sensory evaluation methods

III.2.1. Context and aims

The nature of the differences observed between the products measured by rapid temporal methods has rarely been examined in the literature, however these differences can be explained in three distinct ways. Sensory attributes can be perceived in one product and not in another (III.1.4), independently of temporal aspects. In this case, the added-value of temporal methods is almost nil. Distinct sensory attributes can be perceived at different moments, but at the same moment in all products. In this case, the added-value of temporal methods is limited, as they only make it possible to characterize the dynamics of perception common to all products. Finally, the same sensory attributes can be perceived at different moments in different products. In this case, using temporal methods really makes the difference compared to static methods. The results of the scoping review (I.3.2, part 3.4.2) showed that confirmatory data analyses considering the temporal nature of data were only used in 25% of articles to conclude on the differences between products. Thus, the real ability of temporal methods to discriminate between products on the basis of temporal aspects is not well established, especially with measurements involving consumers.

It was observed that just because the concurrent methods collect data in continuous time did not mean that they were able to capture subtle temporal differences, since in most cases the differences observed were the same than those observed with the retrospective methods (III.1.2). Thus, it is likely that the temporal resolution of TDS and TCATA for the purpose of capturing temporal differences is below most expectations, and the magnitude of the temporal differences that the methods are able to capture remains to be documented. Moreover, observing differences does not mean that these differences actually exist. Indeed, the reliability of temporal measurements was ignored in the studies presented in Section III.1 as in most studies (see Section I.3.2, part 3.3.3). Neither the face validity of methods nor their ability to discriminate between samples guarantee the reproducibility of the conclusions, and it can be difficult to separate variations due to the actual differences between products from those due to systematic errors or bias in the measurements. This is why it is important to go further and deepen the conclusions drawn from the temporal data collected from consumers.

In this chapter, the performances of temporal sensory evaluation methods were studied. Three studies were carried out with the aim of evaluating the validity, reliability, temporal resolution and ability to discriminate between products of the TDS, TCATA and AEF methods.

The first article investigated the validity and the reliability of TDS, TCATA, and AEF-RATA (a variant of AEF-A including RATA scores) measurements collected from three panels of consumers (between-subject design), each having used a different method to evaluate in the laboratory different controlled temporal stimuli delivered using a gustometer.

The second article studied the ability of TDS, TCATA, AEF-D and AEF-A to capture within-product significant differences between attributes taking into account the heterogeneity in individual measurements. The article was based on data collected by four panels of consumers (between-subject design), each having used a different method to describe guacamoles, ice teas (in the laboratory), crisps and dark chocolates (at home).

The third article tried to determine which temporal sensory evaluation methods (TDS, TCATA, AEF-D, AEF-A), variables (citations, durations, times of citation) and statistical analyses were the most product-discriminative while ensuring valid differences. It was based on the data used in the first and second article.

III.2.2. Assessment of the reliability and validity of temporal sensory evaluation methods used with consumers on controlled stimuli delivered by a gustometer

This article was published in Food Quality and Preference:

<https://www.doi.org/10.1016/j.foodqual.2023.104942>

N. Béno: Conceptualization, Data Curation, Investigation, Methodology, Project administration, Resources, Validation, Writing - Review & Editing; L. Nicolle: Data Curation, Investigation, Methodology, Resources, Writing - Review & Editing; P.

Schlich: Methodology, Supervision, Writing - Review & Editing; M. Visalli: Conceptualization, Funding acquisition, Methodology, Software, Visualization, Writing – Original draft.

Temporal Dominance of Sensations (TDS) and Temporal Check-All-That-Apply (TCATA) are the most used qualitative temporal sensory evaluation methods. Considered as rapid methods, they are more and more used with consumer panels. However, no study has investigated in detail the validity and reliability of these methods when they are used with consumers. This study aimed to fill in the gap by comparing the results obtained from temporal measurements to controlled temporal stimuli delivered using a gustometer. A total of 149 consumers were randomly assigned to one of three panels, each panel using a different temporal sensory evaluation method among TDS (n=50), TCATA (n=50) and AEF-RATA (n=49). AEF-RATA is introduced as a new method allowing the consumers to retrospectively report the intensity of the applicable attributes in three defined periods. Using a gustometer, four single-compound solutions were delivered to the consumers to evaluate their recognition ability using free comment. Then, eighteen multiple-compound solutions (composed of two to five compounds varying in their sequence, concentration and duration of stimulation) were delivered to the consumers to evaluate their ability to use the three temporal evaluation methods. The compounds included sodium chloride (“salty”), saccharose (“sweet”), citric acid (“acid”), citral (“lemon”) and basil hydrosol (“basil”). The results show that consumers were able to report temporal sequences of perceived attributes congruent with the delivered sequence of compounds in the stimuli. Contrary to our hypotheses, despite different operational definitions, very few differences are observed between the TDS and TCATA measures, suggesting that most consumers interpreted applicability as dominance. Citation rates and durations of dominance/applicability are strongly correlated between them, but not with the concentration of the compounds. Individual repeatability is low, but does not impact reliability of the measures at the panel level. Consumers can be confused with attribute identification, which can make the interpretation of low citation rates tricky. TDS restores the signal the closest to the stimuli, except for durations for which TCATA performs better. AEF-RATA has no advantage over the other methods. These results open a discussion on what is possible to ask consumer panels and how not to overinterpret the results.

1. Introduction

1.1. Temporal measurements, from analytical to rapid methods

Numerous sensory evaluation methods have been developed to capture the temporality of perception in food products (Kemp et al., 2017). Time-Intensity (TI, Lee & Pangborn, 1986) was the first computerized method to emerge (Takagaki & Asakura, 1984), and it has since been widely used and studied (Cliff & Heymann, 1993). However, TI only measures one descriptor at a time, a major limitation that led researchers to develop other intensity-based methods, among them Progressive Profile (Jack et al., 1994), Dual Time Intensity (Duizer et al., 1996), Multi Attribute Time Intensity (Kuesten et al., 2013). However, none of these methods have been widely used, and it was only with the emergence of qualitative methods that TI started to be challenged. In the last decade, two methods have taken over in terms of usage: Temporal Dominance of Sensations (TDS, Pineau et al., 2009) and Temporal Check-All-That-Apply (TCATA, Castura et al., 2016).

TDS introduced the concept of dominance, i.e. the sensation that triggers the attention (one sensation at any time). At the time of its creation, TDS asked trained panellists to score intensities of the dominant sensations. Thus, TDS has been compared with other intensity based methods such as TI (Le Révérend et al., 2008; Pineau et al., 2009; Sokolowsky & Fischer, 2012) or descriptive analysis (DA) (Labbe et al., 2009; Ng et al., 2012; Saint-Eve et al., 2011), the “gold standard” of the moment. Deemed difficult, the task of scoring has since been abandoned because analyses based on durations of dominance computed from TDS data led to similar product discrimination than analyses based on intensities (Schlich, 2017).

TCATA, as a temporal extension of CATA (Adams et al., 2007), was presented as an alternative to TDS having the advantage of being able to track the applicability of all attributes instead of just the dominant one. The first applications of TCATA involved also trained panellists, and the method was compared to TDS, the most widely used temporal method at that time (Ares et al., 2015; Berget et al., 2020; Delompré et al., 2020; Nguyen & Wismer, 2022; Nguyen et al., 2018; Sharma & Duizer, 2019; Wu et al., 2019).

Now, TDS (Ares et al., 2017; Schlich, 2017; Silva et al., 2018; Thomas et al., 2015) and TCATA (Alcaire et al., 2017; Harwood & Drake, 2021; Maheeka et al., 2021; Vidal et al., 2017) are considered as rapid methods as they can be used with consumer panels. Going further in

simplifying the task of collection of temporal data from consumers, the Attack-Evolution-Finish (AEF) method asks panellists to check either dominant (AEF-D, Visalli et al., 2020) or applicable (AEF-A, Visalli et al., 2022) attributes retrospectively to the tasting. However, this shift of paradigm, from analytical to rapid measurements, may question about how accurately these methods measure what they are intended to measure and produce results that correspond to real properties when used with consumers. Indeed, authors like Stone (2018) still considers that subjects need to be above average users (and thus need to be trained) to be able to perform analytical tests. Moreover, the conclusions of most of the aforementioned comparisons of methods are too subjective to help sensory analysts making informed choices, and the need for more detailed investigations based on consumers studies is highlighted by some authors (Meyners, 2020b; Nguyen et al., 2018). Understanding the factors that determine the validity of sensory perception measures is of crucial importance to select the most appropriate sensory evaluation method (Sidel & Stone, 1993).

1.2. Validity of temporal measurements with consumers

It is not a question here of internal validity (i.e. the degree of confidence that the conclusions are not influenced by uncontrolled factors), nor external validity (i.e. the extent to which the results can be generalized). What is under consideration is the validity of the measurements (i.e. how accurately a method measures what it is intended to measure). Without contest, TDS, TCATA capture a signal whatever the type of panel (trained or consumer). However, “just because one obtains a graphical display or a series of tables with associated statistical significance does not mean it has any meaning or validity” (Stone et al., 2012). Four types of measurement validity can be distinguished: face validity, construct validity, content validity and criterion validity (Middleton, 2022).

Face validity refers to the extent to which a test appears to measure what it is intended to measure (Ward et al., 2013). To have face validity, a measure should be “clearly relevant for what it’s measuring, appropriate for the participants and adequate for its purpose” (Bhandari, 2022b). However, face validity is a subjective assessment considered as a weak form of validity at risk for research bias (Bhandari, 2022b). Regarding their large use and adoption, TCATA and TDS present face validity.

Construct validity is the extent to which the measurements used actually test what they are supposed to (Ginty, 2013). A construct is a theoretical concept that is not always directly measurable (Bhandari, 2022a). The construct should be associated with a good operational definition, otherwise the results may be compromised (Bhandari, 2022a). Intensity is the construct assessed in TI and in DA, it is defined as “the magnitude of the perceived sensation” (ISO 5492:2008, 2008). Dominance is the construct assessed in TDS and AEF-D. Several definitions exist for dominance. The team which introduced TDS defined dominance as “what attracts attention at a given time” (Schlich, 2017), assuming that the reasons may be diverse among individuals and across types of products. This general definition led to controversy (Hutchings et al., 2022; Meyners, 2020b; Varela et al., 2018). Applicability is the construct assessed in TCATA and AEF-A, but it is originated from Check All That Apply (CATA, Adams et al., 2007). An attribute is applicable when it is perceived in a product. Construct validity can be decomposed in convergent validity (i.e. the extent to which measures of the same or similar constructs actually correspond to each other, Bhandari, 2022a) and discriminant validity (i.e. the extent to which measures are unrelated or negatively related to measures of distinct constructs, Bhandari, 2022a). The inventors of TCATA defend that dominance and applicability are different concepts, which theoretically results in TCATA having a greater number of cited attributes, longer durations of perception and citation rates expressing intensities (Castura et al., 2016).

Content validity assesses whether a measurement is representative of all aspects of the construct (Rusticus, 2014). These aspects include an attribute, a time of beginning of perception and a duration of perception (perception referring here to dominant attributes for TDS and applicable attributes for TCATA).

Criterion validity examines the extent to which measurements correlate with external criteria and thus are predictive of real-life values (Piedmont, 2014). Intensity was long reported to correlate with concentration of sapid or aromatic compounds (Stevens, 1969). This is why intensities measured with trained subjects have often been considered as a “gold standard” to validate the outcomes of the new methods. Thus, the different measures derived from data collected with temporal methods should correlate with the release of the compounds in the stimuli.

1.3. Reliability of temporal measurements with consumers

Reliability is about the consistency of the measurements. If a measurement is not reliable, it is unlikely to be valid (Middleton, 2023). Stone (2018) considers that analytical tests require replication to have confidence in the reported result. Replications allow repeatability to be evaluated, i.e. the consistency of results collected with the same test on the same sample at different points in time. Replicated stimuli are often evaluated in TDS and TCATA involving trained panels. However, data from replications are most of times aggregated as if they were originated from different subjects in order to increase the number of observations to build TDS and TCATA curves. Repeatability of temporal methods has been very little studied (Castura et al., 2016; Hutchings, de Casanove, et al., 2017; Lepage et al., 2014; Meyners, 2011). Schlich (2017) considers that monitoring individual performances is not necessary while using TDS with consumers and without collecting intensities. However, if consumers are not necessarily expected to be repeatable at individual level, studying repeatability at panel level is important to conclude on the reliability of the temporal measurements.

1.4. Using a gustometer to deliver a controlled multi-attribute temporal stimulus

Validity of sensory measurements is difficult to establish as the “sensory reality” is unknown. This is why some authors agreed on the need to confirm findings using samples with known sensory differences (Ares et al., 2015; Meyners, 2020b). Attempts have been made on videos (Peltier et al., 2019; Velázquez et al., 2020). However, the conclusions can be difficult to transpose to more complex taste stimuli.

The gustometer (Andersen et al., 2019) is an instrument that enables to program the delivery of predetermined volume and concentration of liquid compounds to the tongue over a specified period of time. Such a device has been proven useful for studying the influence of cross-modal sensory interactions on flavour intensity (Niimi et al., 2015) or taste and aromas (Niimi et al., 2014). It has been used in combination with TI to study the impact of temporal fluctuations on enhancement of sweet taste (Burseg et al., 2010) or salty taste intensities (Busch et al., 2009). Using a gustometer could be of primary interest for studying how consumers transcribe their temporal perception of controlled temporal gustatory stimuli using different sensory evaluation methods.

1.5. Objectives

The primary objective of this article was to study the validity and the reliability of measurements of gustometer-delivered stimuli made with consumers using TDS and TCATA. The secondary objective was to test the ability of consumer to retrospectively report the evolution of perceived intensities introducing AEF-RATA, a new method combining AEF-A and Rate-All-That-Apply (Ares, Bruzzone, et al., 2014). The ultimate objective was to better understand what is actually being measured, regarding different aspects of temporal measurements: dominance vs. applicability, citation rates vs. durations vs. intensities, citation times vs. periods.

2. Materials and methods

Hereafter is a summary of the protocol. A detailed description can be found in Béno et al., 2023.

2.1. Participants

A total of 149 consumers (56 men and 93 women, between 21 and 65 years old) participated in this study. They were preselected from a population registered in the ChemoSens Platform's PanelSens database (declared to the relevant authority, authorization number 1148039). The inclusion conditions for participating in this study were as follows: being between 18 and 65 years old; not suffering from food or non-food allergies; not being pregnant or breastfeeding and not following a restrictive diet incompatible with the consumption of sugar or salt. The purpose of the study was explained via an information sheet sent by email. The consumers had to fill out an informed consent form. They were compensated for their participation (one session) with vouchers worth 10 euros. The selected consumers were randomly assigned to one of three panels, with a constraint of balance in gender and age between panels. Each panel used a different temporal method among AEF-RATA (n=49), TCATA (n=50) and TDS (n=50).

2.2. Stimuli

Different stimuli (i.e. liquid solution composed of one or more compounds) were delivered to the consumers using a gustometer model Burghart GU002. This device consists of five syringes

(2.5 ml) containing the sapid and aromatic compound solutions, and two syringes (2.5 ml) containing water. Up to five compounds can be loaded, mixed, and diluted by the gustometer. The selected compounds are described in Table 44.

Table 44. Description of the sapid and aromatic compounds used in the stimuli, and corresponding sensory attributes.

Compound	Attribute	Concentration
Sodium chloride	Salty	25 g/l
Saccharose	Sweet	250 g/l
Citric acid	Acid	10 g/l
Citral	Lemon	10 mg/l
Basil hydrosol	Basil	20 g/l

The attributes *Salty*, *Sweet* and *Acid* were selected because they were meant to be easily identifiable by consumers. *Umami* was not considered because it is unknown to most French people, and *Bitter* to avoid potential carryover effects. *Lemon* and *Basil* were retained because they can both be found in sweet or savoury dishes and thus be associated with the three flavours. *Lemon* was chosen in particular to observe potential interactions and/or confusion that may exist with *Acid*. *Basil* was selected because it was expected easy to discriminate compared to *Lemon* (fruity vs. vegetal) but not easy to identify. The concentrations of the compounds were chosen to be as close to iso-intensity as possible, based on previous studies and internal pre-tests.

The gustometer made it possible to control at each moment of the “tasting” what the consumer has in the mouth. Controlled by a computer, the gustometer delivered liquids in the form of sequences of pulses (70 µl every 100 ms) through a tube (350 µl/s flow rate). Three levels of concentration were chosen for the compounds: weak, medium, and strong, corresponding respectively to 10, 50 and 90% of the concentrated solution. As swallowing was made more difficult by the continuous flow of liquid, the stimulation durations should not exceed 30 seconds to limit the volume in the mouth to 10 ml and reduce the discomfort of consumers.

Four types of stimuli were designed to investigate the research questions: (i) single-compound stimuli of same concentrations; (ii) multiple-compound stimuli with same concentrations and no simultaneous delivery of compounds (S01, S02, S03, S08, see first column of graph panel in Figure 53); (iii) multiple-compound stimuli with same concentrations and simultaneous delivery of compounds (S09, S10, S11, S14, see first column of graph panel in Figure 54); (iv)

multiple-compound stimuli with varying concentrations and simultaneous delivery of compounds (S04, S05, S06, S07, S012, S013, see first column of graph panel in Figure 55).

2.3. Experimental procedure

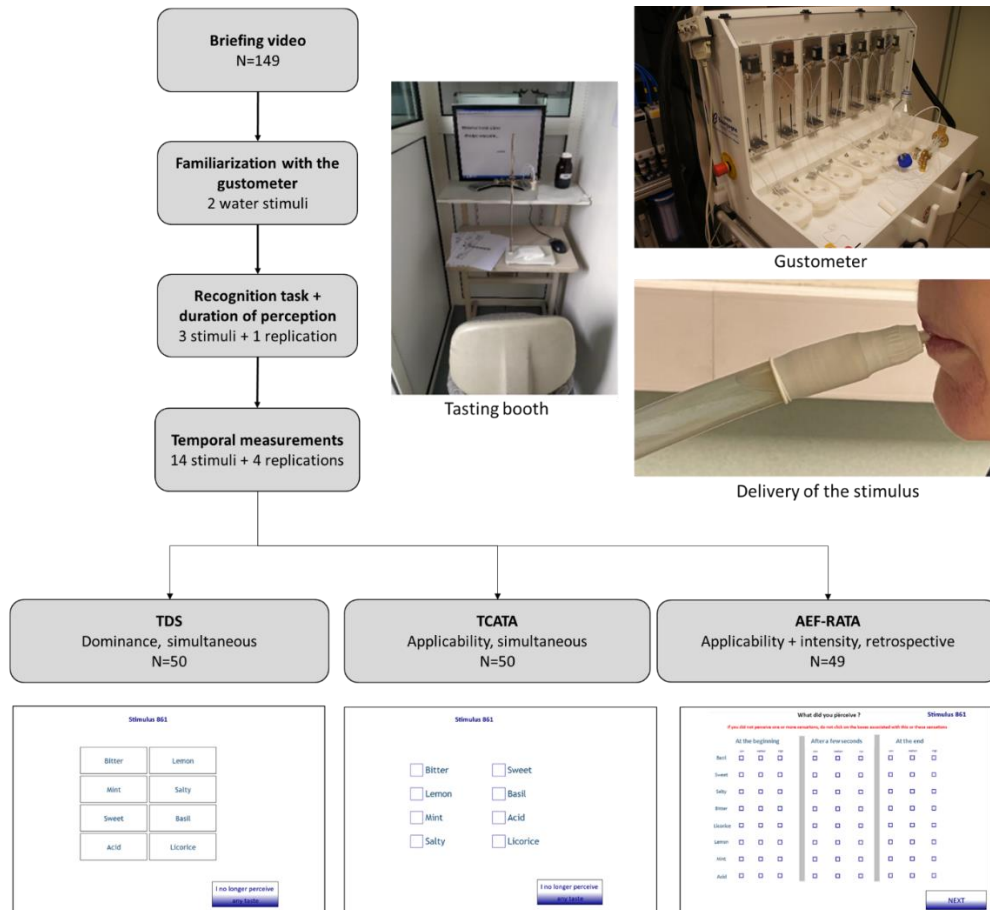


Figure 52. Overview of the experimental procedure.

Figure 52 summarizes the experimental procedure. The consumers participated in individual sessions in the olfactometry lab of ChemoSens. They were installed in a booth equipped with TimeSens© V2 software (INRAE, 2022). For technical reasons, it was not possible to trigger the gustometer from TimeSens. Thus, the delivery of the stimuli was manually triggered by the experimenter.

The session started with a familiarization with the gustometer. The consumers were orally instructed about how the gustometer works. They could ask all the questions they wanted to the experimenter. Then, they were invited to experiment the stimulation with the gustometer so that they get used to the device, and in particular to swallow at the same time as they received a liquid flow under pressure. Two water solutions were delivered during 20 and 30 s.

During the recognition task, four single-compound solutions (medium concentration level) were delivered to the consumers, each for 16 seconds at a constant flow rate of 350 $\mu\text{l/s}$: water for 4 s, compound for 8 s, then water for 4 s. Among the five compounds, each consumer only evaluated two sapid compounds and one aromatic compound. The stimulus presented at rank 4 was a replication of the stimulus presented either at rank 1 or 2. The order of presentation of the stimuli follows an incomplete William's Latin square design balanced at the panel level. The instructions for the task were presented on the screen, then illustrated in a tutorial video. To record their duration of perception, the consumers had to click on a button when they started to perceive and when they no longer perceived anything. Then, they had to self-report the sensation(s) they perceived using free comment. Between two stimuli, water was delivered by the gustometer during 10 seconds to clean the tubes and rinse the mouths of consumers.

During the temporal measurements, 18 multiple-compound solutions were delivered to the consumers (14 different and 4 replicated). The stimuli varied in number, sequence (with or without overlap), duration, and concentration of the compounds (weak, medium or strong). Each stimulus lasted 30 seconds (including water at the beginning and the end), and 10.5 ml were delivered at a constant flow rate of 350 $\mu\text{l/s}$. Since the objective was to compare the methods, the order of presentation was the same for all the consumers: S01, S07, S011, S04, S12, S14, S10, S02, S013, S03, S08, S09, S06, S05, S11_2, S02_2, S08_2, S14_2. This order was chosen by the experimenters in such a way as (i) to have the replicated stimuli at the end; (ii) to alternate fixed-concentration/variable-concentration stimuli; (iii) to alternate stimuli with 3/4/5 descriptors; (iv) to ensure that the stimuli do not start with the same descriptor as the previous stimulus. The instructions for the task were presented on the screen, then illustrated in a tutorial video. The same list of eight attributes was used in TDS, TCATA and AEF-RATA. This list included five attributes related to the compounds: *Sweet*, *Salty*, *Acid*, *Lemon* and *Basil*; and three distractors: *Bitter*, *Licorice* and *Mint*. The order of presentation of the attributes was randomized between the consumers but remained the same for a consumer within the session. For TDS and TCATA, the chronometer started when the consumers clicked on an attribute. The selected attributes, times and durations of perception of dominance (TDS) and applicability (TCATA) were recorded. For AEF-RATA, the consumers had to retrospectively rate the intensity they perceived (weak, medium, strong) for each applicable attribute and for each

of three periods: “at the beginning”, “after a few seconds”, “at the end”. Between two stimulations, water was delivered by the gustometer for 10 seconds to clean the tubes and rinse mouths of consumers. A five-minute break was imposed after the 9th stimulus.

The session lasted for approximately one hour.

2.4. Data analysis

By construction, and due to the technical constraint of the manual trigger of the gustometer by the experimenter, all citation times were left-standardized (i.e. 0 corresponds to the time of the first sensation checked in TDS and TCATA). For this reason, we also left-standardized the time of beginning of delivery of the compounds in the stimuli.

All statistical analyses were performed using R 4.1.0 software (R Core team, 2022) with a risk alpha of 5%. All ANOVAs were performed considering the subject factor as random, and they were followed by Tukey HSD post-hoc tests when significant effects were observed. All correlations were computed using Spearman’s correlation coefficient because of non-normally distributed data for several variables. The strengths of the relationship between two variables were interpreted based on Cohen effect size (Cohen, 2013), and effect sizes were considered large when the coefficients of correlation were greater than 0.5.

2.4.1. Evaluations of single-compound stimuli with free comment

2.4.1.1. Ability of consumers to identify tastes and aromas

The ability of consumers to identify basic tastes and aromas was evaluated on the single-compound stimuli with data from all panels combined. The attributes used for describing the stimuli were plotted as word clouds, and their mean number by consumer computed. Only sensory related words were retained.

A consumer was considered having recognized the compound when he/she has cited (among all the attributes used to describe the stimulus) the corresponding attribute or a related one (e.g. an attribute category related to vegetal for *Basil* or to fruity for *Lemon*, see list of words in Supplementary material). The percentage of recognition was computed by stimulus and by consumer. An ANOVA model was then adjusted to compare the percentages of consumers who did not identify the stimuli in each panel, i.e. $proportion = stimulus + panel + error$. The

percentage of matching attributes used for describing the two replicated stimuli was also computed by consumer.

2.4.1.2. Durations of perception

To verify that there was no difference between the three panels concerning the perceived durations of the attributes in order to be able to conclude later on differences attributable to the methods, two tests were carried out. An ANOVA model was then fitted to compare the mean durations of perception of the stimuli, i.e. *duration of perception = stimulus + panel + stimulus x panel + subject (nested in panel) + error*.

The repeatability of the durations of perception of the stimuli was assessed by computing the standard deviations between the two replicated stimuli, by consumer. Then, an ANOVA model was fitted to compare these mean standard deviations, i.e. *standard deviation of duration of perception = stimulus + panel + stimulus x panel + error*.

2.4.2. Evaluations of multi-compound stimuli with AEF-RATA, TCATA and TDS

2.4.2.1. Variables derived from the stimuli and from the measurements

To study the validity and reliability of the measurements, several variables were computed. Table 45 describes the variables (common to the three sensory evaluation methods) derived from the controlled stimuli and their hypothetically associated measures corresponding to the “chemical reality”. *Number of distinct attributes cited* and *Overall duration* were computed by panellist x stimulus, then averaged by stimulus. *First time/period of citation of the attribute* and *Duration of dominance/applicability of the attribute* were computed by panellist x stimulus x attribute, then averaged by stimulus x attribute.

For TDS, *Duration* was the time difference (for a given stimulus x subject) between the click on the attribute and the click on the next dominant attribute or the STOP button (end of perception). For TCATA, *Duration* was the time difference between the check and uncheck of the attribute or the STOP button if the attribute was never unchecked. When an attribute was cited as dominant or applicable more than once, all durations were summed. *Overall duration* was the time difference (for a given stimulus x subject) between the click on the first attribute and the click on the STOP button.

For AEF-RATA, attributes were considered applicable when their intensities were greater than zero. *First time* was the first period when the attribute was selected (1, 2, 3 or NA), and *Duration* was the number of periods when the attribute was selected (0, 1, 2 or 3). Two additional variables were computed for this method (*Mean intensity* and *Max intensity* of the attributes), then compared to *Mean* and *Max concentration* of the corresponding compounds (experimental unit: stimulus x attribute).

Table 45. Variables derived from the controlled stimuli and their corresponding measures.

Experimental unit	Variable name	Controlled	Measured
Stimulus	Number of attributes	Number of compounds present in the stimulus	Number of distinct attributes cited for the stimulus
Stimulus	Overall duration	Overall duration of the stimulation	Overall duration of perception of the stimulus
Stimulus x attribute	First time	First time of delivery of the compound	First time (TDS and TCATA) or period (AEF-RATA) of citation of the attribute
Stimulus x attribute	Duration	Duration of delivery of the compound	Duration of dominance/applicability of the attribute (in seconds for TDS and TCATA, in periods for AEF-RATA)
Stimulus x attribute	Citation rate	100% if present, 0 otherwise	Percentage of consumers having cited the attribute at least once
Stimulus x attribute	Max citation rate	Max concentration of the compound (hypothesized from Castura et al., 2016 and Stevens, 1969)	Max percentage of consumers having cited the attribute at a specific time

2.4.2.2. Face validity of the measurements

To assess the face validity of the measurements, the sequences of delivered compounds were visually compared with the usual representation of the TDS, TCATA and AEF-RATA data.

The evolution of concentrations of the compounds in the stimuli were plotted in the first columns of the panels of plots (Figures 53-55), the x-axis corresponding to time of delivery of the compound by the gustometer, and the y-axis corresponding to the concentration expressed in four levels: 0 (absence), 1 (10% of the concentrated solution), 2 (50% of the concentrated solution) and 3 (90% of the concentrated solution).

The evolution of citation rates computed at panel level (dominance or applicability) at each time (discretization: 0.1 s) or period were plotted using TDS curves (Pineau et al., 2009), TCATA curves (Castura et al., 2016) and AEF barplots (Visalli et al., 2020) in columns 2-4 of the panels of plots (Figures 53-55). The curves were smoothed using locally estimated scatterplot smoothing (LOESS, Cleveland, 1979). For TDS, the significance line was represented as a horizontal line with ordinate corresponding to the minimum value the dominance rate should equal to be considered as significantly higher than (1/number of attributes) based on a binomial test. Additionally, for the three methods a horizontal red line was drawn corresponding to the highest citation rate reached by distractors.

2.4.2.3. Construct validity of the measurements

The construct validity of the measurements was assessed by comparing each variable of Table 45 measured using a given method with the same variable measured using another method (i.e. TDS vs. TCATA, TDS vs. AEF-RATA and TCATA vs. AEF-RATA). The objective was to identify and quantify the similarities (convergent validity) and differences (divergent validity) between the different methods.

For variables computed by stimulus, coefficients of correlations were computed and paired Wilcoxon tests (because of non-normally distributed data for several variables) performed between the two vectors of 14 observations (replicates were not considered to avoid over-weighting the four replicated stimuli).

For variables computed by stimulus x attribute, coefficients of correlation and mean differences were first computed within stimulus between the two vectors of attributes. Then the means and standard deviations of the coefficients of correlation (14 observations) were computed, and paired Wilcoxon tests performed (14 observations). For these calculations, two vectors of attributes were considered when relevant: “expected attributes” corresponding to compounds present in the stimuli; and (ii) “unexpected attributes” corresponding to compounds not present in the stimuli.

2.4.2.4. Content validity of the measurements

The content validity of the measurements was assessed by comparing the different variables of Table 45 measured using the same method (i.e. *Citation rate* vs. *Max citation rate*, *Citation rate* vs. *Duration*). The objective was to test if all the variables derived from the data are

independent and thus not redundant. The same computations as in Section 2.4.2.3 were performed (only correlations and for expected attributes).

2.4.2.5. Criterion validity of the measurements

The criterion validity of the measurements was assessed by comparing the controlled and the measured variables of Table 45 (e.g. *Number of compounds present in the stimulus* vs. *Number of distinct attributes cited in TDS*). The objective was to test if the variables derived from the data are related to chemical properties of the stimuli and how much the measured values deviate from “reality”. Deviation includes random error (i.e. differences between observed and target values due to random variation) and systematic error (i.e. consistent differences between observed and target values). The same computations as in Section 2.4.2.3 were performed (only Wilcoxon tests).

2.4.2.6. Reliability of the measurements

To assess the repeatability at the individual level, the Jaccard index (Jaccard, 1901) was computed as the ratio $J(t_i) = (\text{number of attributes cited both in replicates 1 and 2 at } t_i / \text{total number of different attributes cited in replicates 1, 2 or both at } t_i)$ at each time/period t_i for all replicated stimuli, by method. In TDS, $J(t_i)$ equals 1 when the same attribute is dominant at t_i and 0 otherwise. In TCATA and AEF-RATA, $J(t_i)$ can be between 0 and 1. The individual reliability scores were computed as the averages of all $J(t_i)$ by consumer and stimulus. These scores can range between 0 (completely different temporal sequences) and 1 (exactly the same sequences). An ANOVA model was then adjusted to compare these scores, i.e. *individual reliability score = stimulus + method + stimulus x method + consumer (nested in method) + error*.

To assess reproducibility of the measurements at the panel level, the same computations as in Section 2.4.2.3 were performed (using differences in absolute values for Wilcoxon tests) to compare the same variables measured between the replicated stimuli.

2.4.2.7. Hypotheses

The following hypotheses were formulated.

H1A. Large effect sizes are expected for correlations between methods for the variables *Number of attributes*, *Citation rate*, *Max citation rate*, *First time*, *Overall duration* and *Duration*

of unexpected attributes (convergent validity). No large effect sizes are expected for correlations between TDS/TCATA and TDS/AEF-RATA for the variable *Duration of expected attributes* (divergent validity).

H1B. Significant differences (positive differences in favour of applicability-based methods) are expected for Wilcoxon tests between TDS/TCATA and TDS/AEF-RATA for the variables *Number of attributes*, *Citation rate of expected attributes*, *Max citation rate of expected attributes*, *Duration of expected attributes* (divergent validity). Non-significant differences are expected for the other variables, nor between TCATA and AEF-RATA (convergent validity).

H2. No large effect sizes are expected for correlations between the different variables measured with the same method, except for *Citation rate* vs. *Maximum citation rate* (content validity).

H3A. Large effect sizes are expected for correlations between controlled and measured variables (criterion validity).

H3B. No significant differences are expected for the Wilcoxon tests between the variables except for *Duration of expected attributes* with TDS which is not expected to correspond to duration of stimulation (criterion validity).

3. Results

Data are available in Béno et al., 2023.

3.1. Evaluations of single-compound stimuli with FC

3.1.1. Identification of tastes and aromas

Table 46. Percentage of citations. Rows: stimulus, columns: perceived sensation.

	Acid	Basil	Bitter	Lemon	Salty	Sweet
Citric acid	63%	0%	19%	29%	<1%	14%
Sodium chloride	7%	0%	14%	0%	78%	6%
Saccharose	2%	<1%	1%	0%	2%	87%
Basil hydrosol	<1%	10% (54%*)	12%	<1%	4%	3%
Citral	4%	0%	4%	45% (64%**)	0%	10%

*: Vegetal terms. **: Fruity terms.

Table 46 shows that the three sapid stimuli were rather well identified, especially for *Sweet* (87%). The *Salty* and *Acid* stimuli have lower identification rates (78 and 63%, respectively).

Compared to the sapid stimuli, the flavour stimuli have lower identification rates: 45% for *Lemon*, and only 10% for *Basil* (64 and 54% when considering the fruity or vegetal related terms, respectively). *Acid* is the attribute having the highest rate of misidentification, being identified as *Lemon* by 29% of consumers (surprisingly, the opposite was not true), *Bitter* (19%) and *Sweet* (14%).

When considering the fruity or vegetal related terms for *Lemon* and *Basil*, 30% of the consumers correctly identified the four stimuli, 35% three, 23% two, 10% one, and 2% zero. The capacity of identification depends on the stimulus ($F=10.77$, $p=0.003$) but not on the panel ($F=0.46$, $p=0.65$). Only 32% used the same exact words to describe the two replicated stimuli, while 33% used completely different words. The list of words used to describe the stimuli is available in Supplementary material.

3.1.2. Durations of perception

In average the tastes and flavours are perceived during 12.2 s (standard deviation: 6.4 s) while they are delivered during 8 seconds, with no significant difference between the five stimuli nor between the three panels (see detailed results in Supplementary material).

3.2. Evaluations of multi-compound stimuli with AEF-RATA, TCATA and TDS

3.2.1. Face validity of the measurements

Figure 53 shows that whatever the method, the results observed at panel level are consistent with stimuli when there is no simultaneous delivery of compounds. The sequences of attributes having the highest citation rates correspond to the sequences of compounds delivered in the stimulus. For AEF-RATA, the compounds delivered in the first seconds and last seconds have the corresponding attributes with the higher citation rates in periods A and F, respectively. There are less intersections between curves in TDS compared to TCATA. Durations of perception are longer than durations of stimulation.

As a tendency, TCATA and AEF-RATA reach higher maximum citation rates than TDS, but it is impossible to know if it is due to the fact that several attributes can be quoted at a given time or to a superior agreement. For these methods the citation rates of unexpected attributes (horizontal red lines) are also higher, notably for *Lemon* which can reach a large citation rate (greater than 40% on Figure 53.F) but also for distractors (25% on Figure 53.F) while it remains under the significance level in TDS. This result confirms that the poor identification ability

observed in Section 3.1.1 (confusion between *Acid* and *Lemon*) can affect the validity of the results.

Despite iso-intensity, *Sweet* was perceived by more consumers whatever the method (red curves and bars in Figure 53.A, Figure 53.B, Figure 53.C, Figure 53.D). In Figure 53.B (and Figure 53.E, to a less extent), the maximum citation rates are not related to concentrations of the compounds. As a general trend, the citation rates seem to peak in the middle of compound release periods. These results suggest that maximum citation rates are functions of durations of stimulation and easiness of identification of the compound.

Figure 54 shows that despite simultaneous delivery of compounds, the shapes of the TDS and TCATA curves are very similar if exception is made for the heights of the citation rates, greater in TCATA. Observations made on Figure 53 are still true for the citation rates: they do not depend on the concentration. *Sweet* still has the highest citation rate, even when the compound was delivered at low concentration level (Figure 54.B). *Lemon* also reaches high citation rates when delivered at low concentration level (Figure 54.A), but *Acid* (Figure 54.B, Figure 54.C and Figure 54.D) and *Salty* (Figure 54.B) do not (an do not reach significance in TDS). Except for these attributes, the sequences of attributes having significant dominance rates correspond to the sequences of delivered compounds. This is also true for TCATA, but the durations of applicability do not always match with the durations of delivery: on Figures Figure 54.A and Figure 54.B, while *Lemon* was delivered during all the stimulation, its citation rate decreases over time notably when new easy-to-identify compounds are delivered.

AEF-RATA highest citation rates in periods A and F still correspond to attributes perceived in the first and last seconds. As in TCATA, the citation rates in AEF-RATA are not related to the concentrations. However, when the duration of the stimulation is short as on Figure 54.B, the distinction between the periods is blurred, as *Sweet* reaches relatively high citation rates in periods A and E. For TCATA and AEF-RATA, when the number of compounds increases, the distractors are more cited. It seems that low confidence should be placed in interpreting the evolution of attributes with citation rates below 25%, which interestingly corresponds roughly to the TDS significance level of citation rate.

Figure 55 confirms previous observations about the similarity of shapes of TDS and TCATA curves. Surprisingly, the citation rates of *Sweet* are not systematically higher than the others (Figure 55.A and Figure 55.B), which suggests that the variations in concentrations could

disturb the ability to identify the sensations. Citation rates go up if the concentration goes up after going down (Figure 55.B), but essentially if there is no competition with a more concentrated compound (see *Basil* on Figure 55.C and *Lemon* on Figure 55.D). These results suggest that the novelty is more important than the concentration for triggering dominance and applicability. More surprising, in TCATA the delivery of new compounds makes some consumers unchecking attributes for which the compound is still present. When the complexity of the stimuli increases (Figure 55.E and Figure 55.F), the curves are less aligned with the stimuli, and short duration variations are not perceptible in citation rates (Figure 55.F). AEF-RATA cannot report the variations of concentration within a period. Sometimes the barplots seem closer to the stimulus than the other methods (notably for *Basil*, Figure 55.C), but citation rates for unexpected attributes are overall larger.

The TDS and TCATA curves and the AEF-RATA barplots are “relatively similar” to the signal of the controlled stimuli, the methods present face validity. However, as expected it is very difficult to draw definitive conclusions based on such a subjective interpretation.

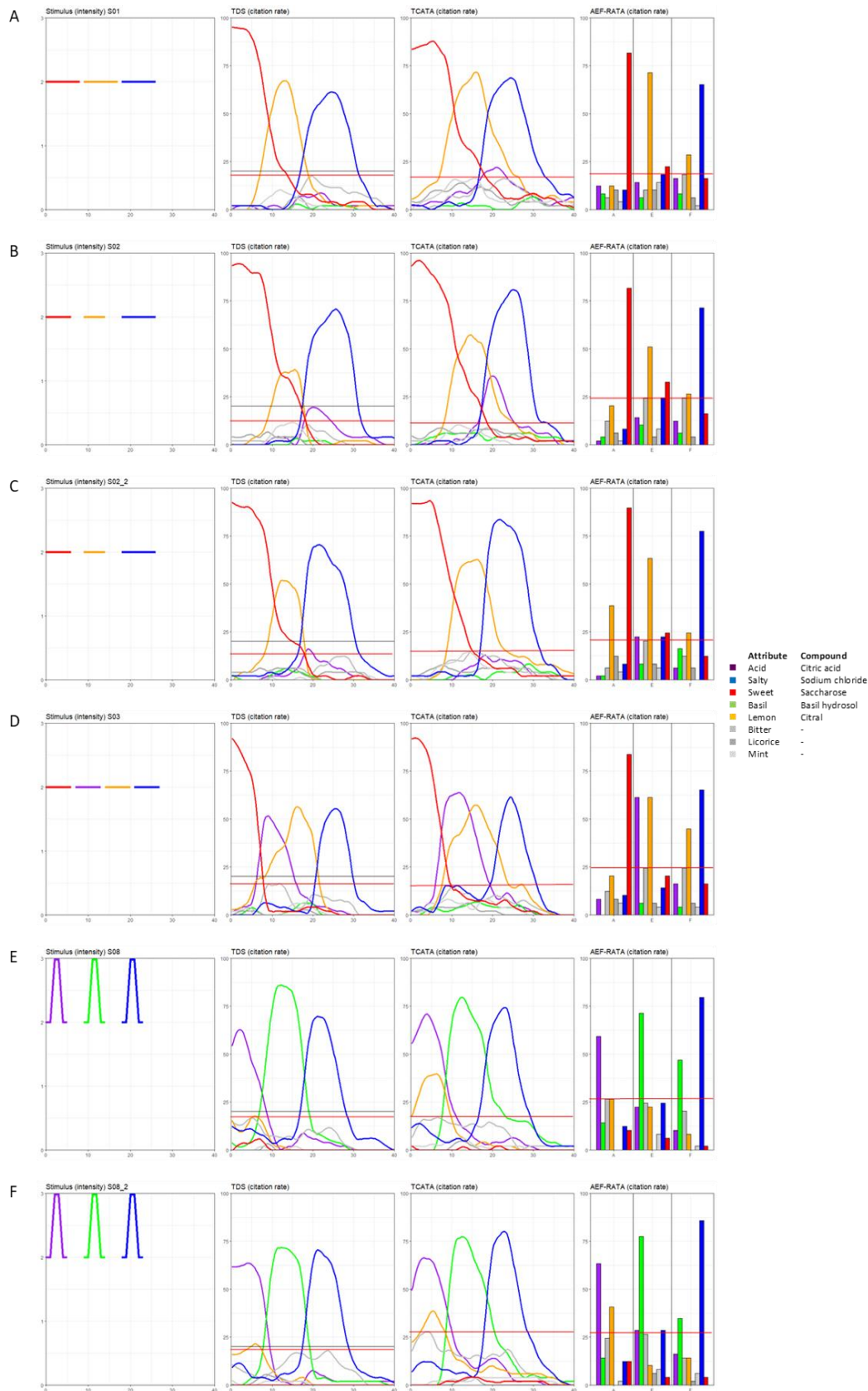


Figure 53. Evaluations of stimuli with same concentrations and no simultaneous delivery of compounds. From left to right: evolution of concentration of compounds in the stimuli over time, TDS curves, TCATA curves and AEF-RATA barplots. A: S01; B: S02; C: replication of S02; D: S03; E: S08; F: replication of S08.

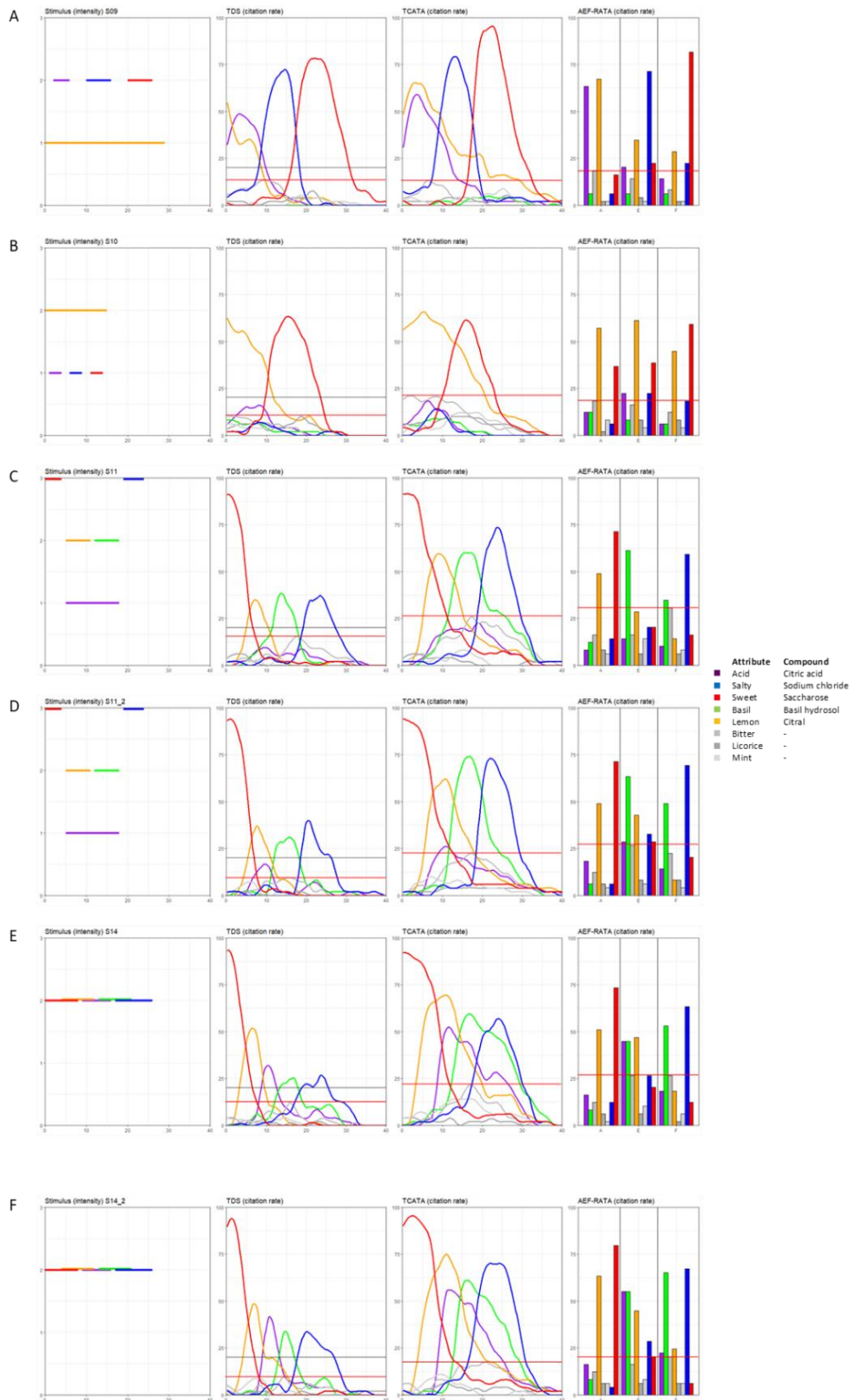


Figure 54. Evaluations of stimuli with constant concentrations and simultaneous delivery of compounds. From left to right: evolution of concentration of compounds in the stimuli over time, TDS curves, TCATA curves and AEF-RATA barplots. A: S09; B: S10; C: S11; D: replication of S11; E: S14; F: replication of S14.

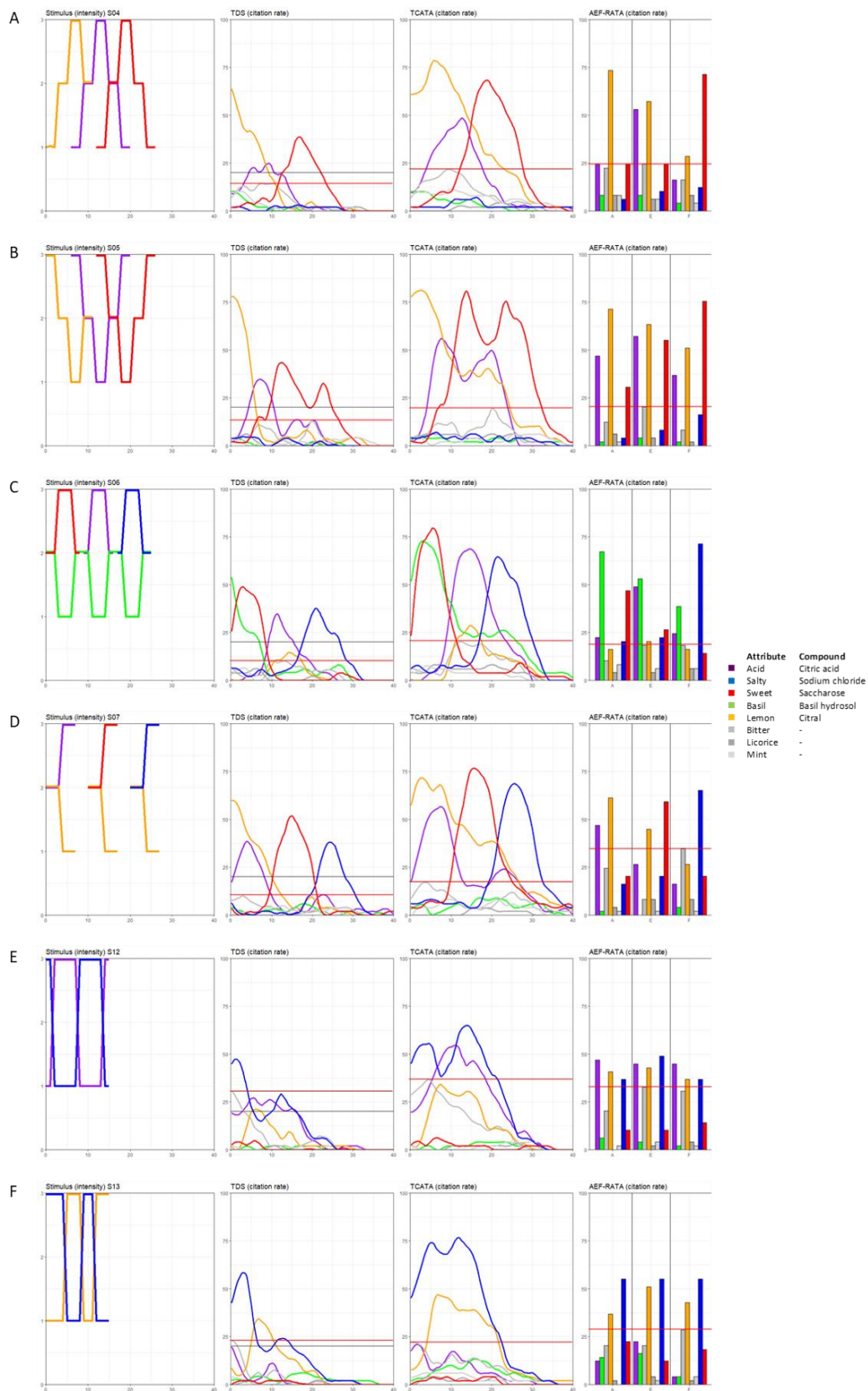


Figure 55. Evaluations of stimuli with varying concentrations and simultaneous delivery of compounds. From left to right: evolution of concentration of compounds in the stimuli over time, TCATA curves, TDS curves and AEF-RATA barplots. A: S04; B: S05; C: S06; D: S07; E: S12; F: S13.

Table 47. Construct validity of the measurements.

Variable	TDS vs. TCATA		TDS vs. AEF-RATA		TCATA vs. AEF-RATA	
	Correlation	Difference	Correlation	Difference	Correlation	Difference
Number of attributes	0.99	-0.42**	0.97	-0.19*	0.97	+0.23*
Overall duration	0.95	-0.60 s**	-	-	-	-
First time – expected attributes	0.86 +/- 0.14	-0.21 s	-	-	-	-
Duration – expected attributes	0.46 +/- 0.14	-2.76 s***	-	-	-	-
Duration – unexpected attributes	0.53 +/- 0.07	-0.97 s***	-	-	-	-
Citation rate – expected attributes	0.64 +/- 0.15	-5.36%*	0.77 +/- 0.12	+1.69%	0.77 +/- 0.10	+7.05***
Citation rate – unexpected attributes	0.53 +/- 0.09	-5.30%***	0.62 +/- 0.08	-5.44%***	0.53 +/- 0.10	-0.13%
Max citation rate – expected attributes	0.66 +/- 0.14	-7.37%*	0.75 +/- 0.08	+0.25%	0.71 +/- 0.13	+7.62%*
Max citation rate – unexpected attributes	0.63 +/- 0.08	-4.28%***	0.62 +/- 0.08	-4.55%***	0.64 +/- 0.08	-0.27%

Correlation: Spearman coefficient of correlation +/- standard deviations (for variables computed by stimulus x attribute). Difference: Mean value of the difference (first minus second method). P-values of the Wilcoxon test: <0.001***, <0.01**, <0.05*.

Table 48. Content validity of the measurements.

Variable	Correlation TDS	Correlation TCATA	Correlation AEF-RATA
Citation rate vs. Max citation rate	0.88 +/- 0.06	0.90 +/- 0.03	0.83 +/- 0.06
Citation rate vs. Duration	0.75 +/- 0.07	0.76 +/- 0.10	0.72 +/- 0.14
Max citation rate vs. Duration	0.76 +/- 0.11	0.78 +/- 0.10	0.59 +/- 0.15

Correlation: Spearman coefficient of correlation +/- standard deviations (computation including only expected attributes).

Table 49. Criterion validity of the measurements.

Variable	TDS		TCATA		AEF-RATA	
	Correlation	Difference	Correlation	Difference	Correlation	Difference
Number of attributes	0.80	-0.19	0.76	0.23	0.84	0.00
Overall duration	0.61	1.91*	0.56	2.51**	-	-
First time	0.97 +/- 0.02	0.86 s	0.97 +/- 0.02	1.07 s*	0.96 +/- 0.02	-
Duration – expected attributes	0.03 +/- 0.19	-4.21 s**	0.16 +/- 0.23	-1.45 s	0.16 +/- 0.25	-
Duration – unexpected attributes	-	0.94 s***	-	1.91 s***	-	-
Citation rate – expected attributes	-	-24.20%***	-	-18.83%***	-	-25.89%***
Citation rate – unexpected attributes	-	15.19%***	-	20.50%***	-	20.63%***
Max citation rate	0.17 +/- 0.21	-	0.32 +/- 0.20	-	0.30 +/- 0.19	-

Correlation: Spearman coefficient of correlation +/- standard deviations (for variables computed by stimulus x attribute). Difference (method minus controlled stimulation): Mean value of the difference. P-values of the Wilcoxon test: <0.001***, <0.01**, <0.05*.

3.2.2. Construct validity of the measurements

Table 47 shows the results of the correlations and Wilcoxon tests between the same variables measured using different methods. Large effect size correlations are observed whatever the variable and the compared methods, except for the correlation between the duration of expected attributes in TDS and TCATA.

Significant differences are observed between TDS and TCATA. As expected, due to the difference between dominance and applicability the number of attributes and the durations/citation rates/max citation rates of expected attributes are lower in TDS. However, the durations/citation rates/max citation rates of unexpected attributes are also lower in TDS, confirming results observed in Section 2.3.1 about the larger number of “false positives” due to higher citation rates in TCATA and AEF-RATA.

There was no reason for overall duration of perception to differ depending on the sensory evaluation method, because in TDS and TCATA this duration corresponded to the moment when the panellist no longer perceives anything. However, the overall duration was lower in TDS, which shows that the temporal evaluation method has an impact on the reported durations. Unexpected significant differences are observed between TCATA and AEF-RATA. While the two methods are supposed to measure the same concept, citation rates and max citation rates are higher in TCATA compared to AEF-RATA (only for expected attributes). More attributes are cited and citation rates/max citation rates are higher in TCATA.

These results partially validate the hypothesis H1A: all the variables derived from TDS, TCATA and AEF-RATA measurements are strongly correlated, except durations. The hypothesis H1B is not validated: there are differences in durations and citation rates between TDS and applicability-based methods, but some of these differences are due to higher citation rates in TCATA/AEF-RATA for both expected and unexpected attributes. The concept of applicability is slightly different from the concept of dominance, and divergent validity is only observed for biased measures.

3.2.3. Content validity of the measurements

Table 48 shows the results of the comparison between the different variables measured using the same method. All the measured variables are correlated with large effect sizes. This is contrary to our hypotheses as we expected citation rates and durations to be less correlated

as they were assumed to be related to different physical properties of the stimuli. In the same way, contrary to our hypotheses, AEF-RATA intensities are correlated with durations and citation rates (see Supplementary material).

The hypothesis H2 is not verified: citation rates, durations and intensities (AEF-RATA) are not independent measures.

3.2.4. Criterion validity of the measurements

Table 49 shows the results of the comparison between the control and measured variables. Contrary to our hypotheses, no large correlations are observed between measured intensities and concentrations (AEF-RATA), nor between measured durations and durations of delivery (TDS, TCATA), nor between maximum citation rates and concentrations (all methods). The overall durations of perceptions are longer than the durations of the stimulations, confirming results observed in Section 3.1.2. Overall durations measured with TDS are in average closer to overall durations of stimulation than those measured with TCATA. In TCATA, times of citation are delayed compared to times of delivery in the stimuli. As expected, durations measured with TDS are not representative of durations of stimulation, while they were closer when measured with TCATA. Both TDS and TCATA measure durations significantly greater than zero for unexpected attributes, TCATA overestimating more than TDS these durations. All methods significantly underestimate the citation rates of expected attributes, TCATA being in average closest to the stimuli. All methods significantly overestimate the citation rates of unexpected attributes. Contrary to our hypotheses, AEF-RATA intensities are poorly correlated concentrations of the compounds (see Supplementary material).

The hypothesis H3A is only verified for the number of attributes, the overall duration and the first times of citation. This result means that consumers are able to detect and report sequences of attributes with a fairly good temporal resolution. However, the results aggregated at the panel level do not robustly quantify the magnitude of the temporal change in chemical compositions of the stimuli, whether directly (durations and intensities) or indirectly (citation rates). H3B is not verified: systematic error of measures (false positive and false negatives) are observed for dominant/applicable attributes, with an error rate between 15 and 25% for citation rates, and a systematic underestimation of attribute durations.

3.2.5. Reliability of the measurements

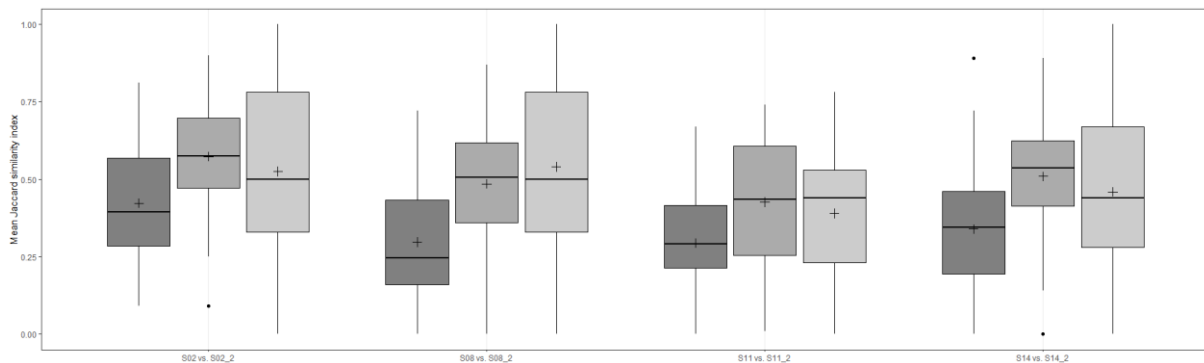


Figure 56. Distribution of the individual repeatability scores averaged over times (TCATA, TDS) or periods (AEF-RATA) of the consumers, by product and method. Plus symbols: means. Dark grey: TDS, medium grey: TCATA, light grey: AEF-RATA.

Figure 56 shows the distribution on individual repeatability scores computed based on Jaccard index. Stimuli ($F=9.17$, $p<0.001$) and method ($F=5.67$, $p<0.001$) effects are significant, but not the interaction ($F=2.04$, $p=0.06$). TDS (mean=0.34, group a) is less repeatable than AEF-RATA (0.48, b) and TCATA (0.50, b). The repeatability is better in S12 (0.51, c) than in S08 and S14 (0.40, b) than in S11 (0.37, a). TDS was less repeatable at individual level than the applicability-based methods. The individual repeatability decreases when complexity of the stimuli increases. For variables aggregated at the panel level (i.e. Number of attributes, Overall duration, First time, Duration, Citation rate, Max citation rate, AEF-RATA mean and max intensity), no significant differences were observed whatever the variable and method (see Supplementary materials).

4. Discussion

4.1. Limitations

Before discussing the results of this study, it is necessary to report the limitations that may limit the generalizability of the conclusions. Most were due to technical constraints related to the use of the gustometer. The stimuli were restricted to liquid solutions, excluding *de facto* some sensory modalities, notably texture attributes that are important in temporal perception (Devezeaux de Lavergne et al., 2017). The stimuli were limited in duration and in number of compounds, which could have benefited the TDS and AEF-RATA methods (limited number of co-occurrences of attributes). The way of tasting, with a pipe in the mouth which delivered a liquid continuously, was not ecological. Some consumers may have been disturbed by the

device and the particular way of swallowing. Particularly, taste perception could have been affected because for some consumers the position of the velum may have been modified under these conditions compared to a classic tasting (Sinding et al., 2021).

The design of the experiment should also be considered. For the purpose of studying the criterion validity, the sensory perception was compared to the “chemical reality” of the stimuli, considering “expected” and “unexpected” attributes. However, with regard to potential physiological interactions between the compounds, it is understood that the “sensory reality” remains unknown. Notably, Veldhuizen et al. (2018) reported complex interaction between acid, sweet and lemon. Cognitive phenomena may also have affected the measurements. The stimuli were designed to try to better understand what triggers dominance and applicability. This led to the creation of unrealistic sensory stimuli compared to usual food products, which may have confused consumers. Twenty-four stimuli were presented to the consumers during approximately one hour: it could have generated a certain weariness or a habituation which could have impacted the results despite the precautions taken in the experimental procedure.

4.2. Temporal measurements collected from consumers with TDS, TCATA and AEF-RATA present face validity

Temporal measurements include several aspects related to the sensory perception of a stimulus: the selection of an attribute with regards to the operational definition of the construct (dominance/applicability), the moment of the selection (time/period of beginning of dominance/applicability), the duration of dominance or applicability (in seconds/periods). These aspects are generally evaluated with data aggregated at the panel level by plotting TDS/TCATA curves or AEF-RATA barplots representing the averaged temporal evolution of the perception of the stimuli over panellists. In this study, the visual comparison of these plots with the temporal evolution of the concentration of the compounds delivered in the stimuli (Figure 53 - Figure 55) confirms again that TDS and TCATA, but also AEF-RATA present face validity. In general, studies comparing TDS and TCATA concludes that the two methods identify the most relevant changes in the sensory characteristics of the evaluated samples with TCATA providing better discrimination or more detailed information (Berget et al., 2020). Most of them stopped at this stage, often reaching the convenient conclusion that as the methods are conceptually different, they are expected to differ in the provided information

which makes the measurements complementary. However, whether to choose the most appropriate method (Nguyen et al., 2018) or to interpret the results (Meyners, 2020), it is important to go further and investigate when and how the measurements differ or converge.

4.3. Construct validity is unclear: the concepts of dominance and applicability are interpreted similarly by most consumers and are essentially related to new perceptions

Except to a certain extent for durations, no divergent correlation was observed between TDS, TCATA and AEF-RATA whatever the aspect of the measures considered (Table 47). This shows that the concepts of applicability and dominance have been similarly interpreted by most consumers, despite different operational definitions. Whatever the method, the maximum citation rates are functions of durations of stimulation and easiness of identification, and novelty seems more important than concentration to trigger the selection of both dominant and applicable attributes (see Section 3.2.1). For TDS, this result confirms the conclusions of Varela et al. (2018) who observed that change of dominant attribute was triggered when the perception of an attribute fades and/or when a new sensory attribute is perceived. Hutchings et al. (2022) postulated that the selection of attributes was an instinctive reaction in response to a perceived sensation whatever the definition of dominance given to consumers. It seems that this postulate can be extended to applicability and TCATA, which supports the hypotheses of Nguyen et al. (2018) and Meyners (2020). They assumed that even under TCATA instructions, due to necessarily rapid choices, subjects could use a competitive evaluation between attributes like in TDS, making “most applicable” and “most striking” attributes similar.

More surprisingly, in TCATA, the delivery of new compounds leads some consumers to uncheck the attributes corresponding to compounds that are still present. On some stimuli (Figure 54.A and Figure 55.D), this could be the result of a reinforcement of *Lemon* by *Acid*: when *Acid* was no longer delivered, by a contrast effect *Lemon* was less perceived (Veldhuizen et al., 2018). However, the same effect is observed with *Basil* (Figure 55.C), and in general, the durations of perception of the attributes are shorter than the durations of delivery of the compounds in multiple-compounds stimuli (Table 49) but not in single-compounds stimuli (Supplementary material). Thus, it would seem that either a new perception or interactions between attributes could trigger the end of applicability periods.

4.4. Consumers can describe temporal sequences of attributes, but their lack of precision in identifying attributes can threaten criterion validity

The results of this study confirm that consumers are able to perceive and report temporal changes in the stimuli. The correlation between times of citation and times of delivery of the compounds is almost perfect, and the offset is of the order of one second (Table 49), which corresponds to the average reaction time (Kelling & Halpern, 1988). It is to be noticed that a left standardization of the times was required to align the delivered and perceived signals. This transformation does not distort the sequences, unlike left-right standardization (Michael Meyners & Castura, 2018).

Times of citation and overall durations measured with TDS are closer to their corresponding controlled variables than TCATA, which suggests that TDS is a less cognitive task. Indeed, in TCATA they have to maintain in their working memory the actual state of the attributes (checked or unchecked), and they have more choices to make at each time which can result in different behaviours (Michael Meyners & Castura, 2018). In AEF-RATA, the periods of citation also correlate with the times of delivery (Table 49), which confirms that consumers are able to memorize their perception over a short duration when the number of sensations is limited and concurs with the conclusions of Daniel & Katz (2018).

Whatever the method (Table 49), citation rates varying between 15 and 20% are observed for attributes not corresponding to compounds present in the stimuli. Some high citation rates can be due to associations between *Acid* and *Lemon*, but attributes such as *Mint* or *Licorice* also reach high values. These “distractor” attributes, although probably eligible to describe fresh Basil (Du et al., 2023; Patel et al., 2021), were almost not cited during the recognition task and would not have been included in the list of attributes if we were based on the terms generated by the panel. This result suggests that with consumer panels citation rates can be affected by report effects between non-orthogonal attributes (e.g. *Lemon* and *Acid*) or confusion effects induced by the list of attributes. Stone (2018) claims that about 30% of any population cannot discriminate differences at better than chance. Thus, it should be interesting to investigate if this is due to a subset of “bad” panellists, or to compare the results to those obtained with a trained panel from which a better discrimination between basic taste and flavour would be expected. Moreover, in this study, the stimuli were quite simple compared to real food products, so it is likely that “false positive” or unexpectedly high citation

rates have greater chance to occur with uncontrolled stimuli including substantial noise (Berget et al., 2020).

The citation rates for attributes not corresponding to compounds present in the stimuli is lower in TDS compared to TCATA and AEF-RATA. This can be explained by the fundamental difference between dominance and applicability. When consumers have more freedom (applicability), they may be tempted to check more attributes to be sure not to miss anything. On the contrary, in a situation of forced choice (dominance), they probably take no risk and choose the most obvious attribute, which in turn can imply not citing attributes corresponding to compounds actually present in the stimulus.

4.5. Content validity is unclear: durations of dominance/applicability, citation rates and intensities are similar measures but they are poorly related with the concentrations of the compounds

Most statistical analyses of TDS and TCATA data are based on durations or citation rates. These measures are not independent, as evidenced by the large correlation coefficients (Table 49). This relationship can be easily explained. Attributes not perceived have a duration of zero, thus the lower is the citation rate of an attribute the shorter is its duration. Moreover, the shorter is the duration, the more difficult it is for the citation rate to reach a high percentage due to disagreement and random noise around the citation times of the attribute. This is very important for the interpretation of the results, because it means that analysing durations without considering the sequences of attributes (shown by TDS/TCATA curves) would be equivalent to analysing static citation rates as if TDS/TCATA was no more than CATA data. Moreover, this study show that mean durations are not related to physical phenomenon as they do not correlate with delivered durations while perceived sequences (time of citations of the attributes) match with delivered sequences. This suggests that attribute durations are not necessarily the most appropriate variable to analyse in order to get insights on temporal perception, unless if they are computed by periods.

With AEF-RATA, intensities are also correlated to citation rates and durations (Table 48), but not with concentrations of the compounds (Table 49), which concurs with the conclusions of Kelling & Halpern (1988). Likewise, TDS and TCATA citation rates are poorly correlated with concentrations (Table 49).

4.6. Measurements made with TDS and TCATA are poorly reliable at the individual level but acceptable at the panel level

At the individual level, the repeatability score ranges between 0.3 and 0.5 (Table 48). This means that considering any given time/period, consumers are in average consistent in their attribute selection once in two or three. This seems relatively low, but it is demanding for consumers to be repeatable in the choice of the attributes and the times of choices, especially for complex stimuli. TDS is less repeatable than TCATA and AEF-RATA, which could suggest that there is some randomness in the selection of the attributes when several compete for dominance. This rather low individual repeatability does not seem to affect the reliability of some measures analysed at panel level. This probably means that most errors are random and compensate each other on average.

4.7. Comparison of methods and recommendations

The objective of this study was not to draw definitive conclusions on the methods, but to try to give objective elements to understand what is possible to ask consumer panels and how not to overinterpret the results. In view of the results, some general trends are observed and, on this basis, recommendations can be made for further studies using temporal methods and involving consumers.

TDS is the method the less reliable at the individual level. Regarding the controversy with the definition of dominance that results in a poor operational definition, and knowing that novelty is the main trigger for dominance, a non-ambivalent instruction to give to the participants could be simply to click on the attribute as soon as they perceive it (without much precision about intensity or attention).

Compared to TDS, TCATA is expected measuring more perceptions and more interpretable durations (compared to the chemical reality), but in this study the differences are not obvious. More investigations are necessary to determine if consumers are really capable of reporting more perceptions and appropriate durations of applicability (with or without fading, Ares et al., 2016).

AEF-RATA was introduced in an attempt to ask consumers for intensities of applicable attributes, but the results suggest that such a quantitative measurement should remain in the domain of trained panels. With simple stimuli, it seems that low temporal resolution methods

such as AEF present no advantage compared to TCATA for tracking temporal changes of applicability within-product, except possibly for presenting to consumers longer lists of attributes or using free comment instead of lists of attributes. However, the relevance remains to be demonstrated, FC-AEF having only been used in two studies (Mahieu et al., 2020; Visalli, Dubois, et al., 2022), and the method was proven effective to measure differences in temporal perception only in the first one.

Whatever the method, for consumers, reporting events related to new perceptions is feasible, but the quantification of the temporal perception (either duration or intensity) is far more difficult. Thus, the quantitative interpretation of results should be made with caution. If large citation rates or durations increase the confidence in the actual presence of the compound related to the perceived attribute, differences in citation rates or duration are not necessarily due to differences in intensity (see Section 3.2.1). Too much importance should not be given to attributes having low citation rates or short durations in the interpretation of the results. In TDS curves, important attributes are most often determined based on the significance level based on the binomial test. In TCATA curves and AEF-RATA barplots, there is no equivalent. The results of this study demonstrate that an arbitrary threshold should be fixed to eliminate “false positive” attributes. Empirical evidences from this study suggests this threshold could be set between 20 and 25% for citation rates, but more investigations are required to determine the most appropriate level. With controlled stimuli, times of first citations are valid and more reliable than citation rates and durations. Thus, it could be envisaged to use alternative statistical analysis methods less used at the moment but which do not depend on citation rates and durations but on citation times (Visalli et al., 2023), sequences (Castura & Li, 2016), transitions (Lecuelle et al., 2018) or temporal modelling of categorical data (Peltier et al., 2023).

A number of elements of the protocol can also be considered to limit the risks due to poor identification ability of the consumers. Regarding the list of attributes, presenting definitions with examples should help even if the name of the attributes seems to be auto-explicit. For flavour attributes, sensory categories (e.g. fruity, vegetal, etc.) seems to be better candidate than specific attributes. The attributes should also be chosen to be as orthogonal as possible to avoid potential confusions. Adding one or two replicated samples can also be interesting to quantify the noise in the temporal measurements. Regular consumption of the food to be

studied may not be the only selection criterion. A phase of selection on the capacities to recognize attributes could be useful to make it possible to retain only consumers that are actually able to perform a sensory evaluation task.

5. Conclusion

In this study, results obtained from temporal measurements were compared to controlled temporal stimuli delivered using a gustometer. While TDS, TCATA and AEF-RATA presents face validity evaluated on subjective evaluation of the usual graphic outputs, a more detailed investigation questions about the other aspects of validity. Indeed, considering the chemical reality of the stimulus as the “gold standard” for comparison provides insight into what is actually measured with temporal methods. This allows going further than the usual conclusion about "complementary measures" when differences are found. Thus, it seems that to draw robust and valid conclusions from TDS or TCATA, it is not reasonable to interpret results as those obtained with quantitative evaluation methods measuring robust concepts related to physical properties such as sensory intensities. Both TDS and TCATA can be used with consumer panels to gain insight on temporal sequences of perceived sensations, but subtle differences in mean durations or in citation rates should be interpreted with caution. Indeed, these variables are intrinsically linked and are functions of several parameters including duration of stimulation and concentration, but also easiness of identification of the attribute, number of compounds simultaneously being present in the mouth, novelty, etc. This uncertainty combines with the noise brought by individual variability in temporal measurements, which makes it even more important to consider the nature and the size of the observed differences. In general, in sensory analysis, discrimination is sought, thus liberal tests, more powerful, are used to find a statistically significant result. Of course, there is no serious issue such as with drug testing, but this power is not necessarily a good thing with studies involving untrained panellists as it could result in reporting wrong (or at least uncertain or not reproducible). As the TDS and TCATA tasks primarily consists for consumers to report attribute perception times, in the future the focus may be more on analysing these temporal events rather than seeking discrimination based on quantification of citation rates, durations or intensities that can be better achieved with other types of methods or panels.

Supplementary material

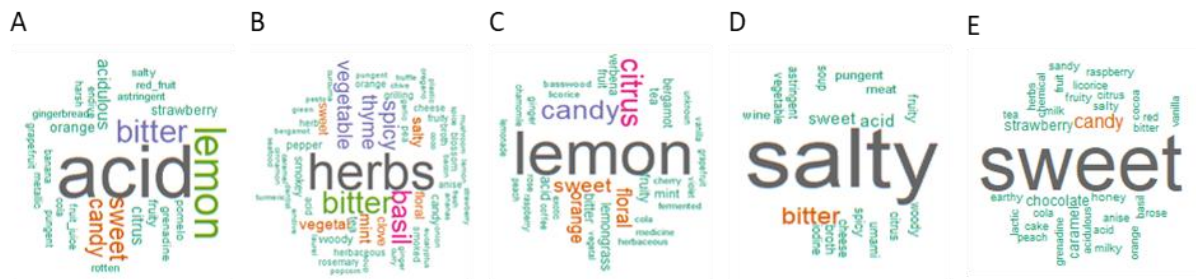


Figure 57. Word cloud of the words used to describe the compound. A: Citric acid. B: Basil hydrosol. C: Citral. D: Sodium chloride. E: Saccharose. The words are plotted with size and colour chosen based on the frequency.

Table 50. Words associated with “Fruity” and “Vegetal”.

Fruity	Vegetal
Fruity, Orange, Strawberry, Bergamot, Raspberry, Fruit, Lemonade, Citrus, Lemon, Peach, Exotic fruit	Vegetal, Mint, Grilled herbs, Thyme, Vegetable, Tea, Soup, Broth, Herb(s), Lemongrass, Herbaceous, Chive, Verbena, Mint, Basswood, Endive, Pea, Rosemary, Eucalyptus, Laurel

Table 51. Comparisons of mean durations of perception and mean standard deviations of durations of perception between the stimuli.

	Duration, in seconds	Standard error in duration between replicated stimuli, in seconds
Acid	12.3 [10.7, 13.9]	3.0 [1.2, 3.8]
Basil	12.6 [11.0, 14.2]	4.7 [2.8, 6.6]
Lemon	12.1 [10.5, 13.7]	2.5 [0.7, 4.3]
Salty	12.2 [10.6, 13.9]	1.9 [-0.1, 3.9]
Sweet	12.1 [10.5, 13.8]	2.1 [0.4, 4.0]
F-stimulus (p-value)	0.13 (0.97)	0.37 (0.83)
F-panel (p-value)	0.52 (0.59)	0.66 (0.53)
F-interaction (p-value)	0.83 (0.57)	0.64 (0.74)

Between square brackets: 5% confidence intervals.

Table 52. Content validity of AEF-RATA intensity measurements.

Variable	Correlation
Duration vs. Mean intensity	0.91 +/- 0.04
Duration vs. Max intensity	0.63 +/- 0.14
Citation rate vs. Mean intensity	0.67 +/- 0.14
Citation rate vs. Max intensity	0.69 +/- 0.15
Max citation rate vs. Mean intensity	0.55 +/- 0.15
Max citation rate vs. Max intensity	0.62 +/- 0.14

Correlation: Spearman coefficients of correlation +/- standard deviations (computation including only expected attributes).

Table 53. Criterion validity of AEF-RATA intensity measurements.

Variable	Correlation
Mean intensity	-0.02 +/- 0.30
Max intensity	0.26 +/- 0.33

Correlation: Spearman coefficients of correlation +/- standard deviations (for variables computed by stimulus x attribute).

Table 54. Reliability of the measurements.

Variable	Difference TDS	Difference TCATA	Difference AEF-RATA
Number of attributes	0.07	0.14	0.13
Overall duration	0.53 s	0.64 s	-
First time	0.68 s	0.28 s	0.04 p
Duration	0.10 s	0.24 s	0.10 p
Citation rate	2.23%	4.43%	5.10%
Max citation rate	3.60%	3.06%	7.14%
Mean intensity	-	-	0.02
Max intensity	-	-	0.01

Difference: Mean values of the absolute differences between the replicated stimuli (p-values of the Wilcoxon test: <0.001***, <0.01**, <0.05*).

III.2.3. Temporal precision and resolution of TDS and TCATA used with consumers

*This article has been **submitted to Food Quality and Preference** and is under revision.*

S. Cordelle: Project administration. B. Mahieu: Conceptualization, Methodology, Writing - Review & Editing. C. Peltier: Conceptualization, Methodology. P. Schlich: Conceptualization, Methodology, Writing - Review & Editing. M. Visalli: Conceptualization, Methodology, Software, Validation, Formal analysis, Data Curation, Writing - Original Draft, Visualization.

Temporal Dominance of Sensations (TDS) and Temporal Check-All-That-Apply (TCATA) make it possible to capture temporal changes in the perception of sensory attributes within a product. However, temporal measurements bring new sources of uncontrolled variability that can make them imprecise, particularly with consumer panels. This article aims to demonstrate the importance of considering this imprecision when interpreting temporal data. First, the levels of imprecision due to non-repeatability or disagreement were assessed. The precision of TDS and TCATA was shown to be low, mainly due to disagreement caused by high heterogeneity in total durations of perception. Then, a new framework of analysis has been introduced. It allowed testing of differences in the median times of selection or unselection of attributes. A new representation of TDS and TCATA data has been proposed to visualize and objectify uncertainty in periods of perception. The conclusions obtained with the new framework were compared to those obtained with the usual analyses based on the evolution of citation rates over time. It was shown that the interpretation of TDS/TCATA curves and difference curves can be tricky without guidelines regarding the minimum size of differences that can be considered significant and interpretable. Finally, it was shown that the temporal resolution of the methods was about 10% of the total duration of the perception. These findings question the ability of TDS and TCATA to capture subtle temporal differences. Therefore, recommendations were made for interpreting temporal data in a way that robust conclusions can be drawn.

1. Introduction

The sensory perception of food is acknowledged as a temporal phenomenon (Dijksterhuis & Piggott, 2000). Thus, over the last 60 years, several evaluation methods has been developed to measure sensory perception (Kemp et al., 2017). Among these methods, most record the perception of panellists in continuous time, simultaneously to the tasting (Visalli, Schlich, et al., 2022). Continuous Time-Intensity (TI, Lee & Pangborn, 1986) was the first method and has long been the temporal method of reference. TI only allows the measurement of one attribute at a time, which is expensive and prevents the study of the interactions between attributes. Therefore, other methods have been developed. Although the use of multi-attributes intensity-based temporal methods remains limited in research, the use of qualitative temporal methods is increasing over the years (Visalli & Galmarini, 2023). These methods replaced intensity measurements with dominance (Temporal Dominance of Sensations, TDS, Pineau et al., 2009) or applicability (Temporal Check All That Apply, TCATA, Castura, Antúnez, et al., 2016) measurements. They allow temporal perception to be recorded not only from trained panellists but also from consumers, which reduces the time and cost of developing a trained panel. More important, they produce results that are likely closer to the natural perception of consumers.

The challenge for temporal methods is to provide data that meaningfully reflect a complex real-world experience using a task that consumers can perform (Castura, 2018). This requires finding the best compromise between ecological assessment, reliability and validity. Recent studies have attempted to evaluate the validity of temporal measurements on controlled stimuli: videos displayed to children or liquid solutions delivered on consumers' tongues using a gustometer. Velázquez et al. (2020) demonstrated face validity of TCATA. They also showed that if TCATA provided more detailed temporal descriptions of videos compared to TDS, both methods failed to capture temporal changes in product perception when used with real food products (milk desserts). Visalli, Béno, et al. (2023) confirmed the face validity of the methods and demonstrated that consumers could report attribute sequences congruent with simple controlled temporal stimuli. However, they also highlighted issues with other aspects of validity, in particular criterion validity. Criterion validity is related to accuracy, i.e. the degree of proximity of a measurement to the "true" value.

Accuracy can be affected by several factors, including the precision of the measuring instruments (i.e. in sensory analysis, the panel using the temporal method). Precision here refers to the degree of closeness between the different evaluations made by the panellists for the same products. Precision is inversely related to the variability of measurements made by panellists. Compared to static measurements, dynamic measurements introduce new sources of uncontrolled variability linked to the temporal dimension of the data. Indeed, the qualitative assessment of the attributes is associated with temporal events, i.e. times of selection or unselection of the dominant (TDS) or applicable (TCATA) attributes. An attribute is unselected when the consumer selects a new dominant attribute (TDS), unchecks an applicable attribute (TCATA), or clicks on the button to report the end of the tasting (stop button).

A first source of imprecision is the disagreement between the panellists. The variability inherent to sensory evaluation, and especially to heterogeneity in temporal measurements cannot be removed entirely by training or controlled experimental setups (Dietz, Yang, et al., 2022; Esme et al., 2020). This variability can be explained by differences in individual behaviour due to oral physiology (Engelen & Van Der Bilt, 2008), oral processing strategies to manipulate foods (Devezeaux de Lavergne et al., 2017; Doyennette et al., 2019; Varela et al., 2021), salivary dilution (Fischer et al., 1994), reaction time (Veldhuizen et al., 2010), gustatory working memory (Daniel & Katz, 2018; Lim et al., 2022), etc. TDS is a method that is assumed to deal with variability between subjects (Schlich, 2017). However, due to this variability, large differences can be observed in the times to first attribute selection or times to click on the stop button (Hutchings et al., 2014; Pineau et al., 2012; Visalli et al., 2020). These differences are likely to affect the times of selection/unselection of the attributes and can thus be a source of uncertainty in the periods of perception of dominant/applicable attributes observed at the panel level. If there are segments of consumers, the average panel view may not match anyone's perception (Cardot et al., 2019).

A second source of imprecision is the non-repeatability of the panellists. With trained panels using static sensory evaluation methods, panellists' agreement and repeatability are often verified before they evaluate samples. With studies involving TDS and TCATA measurements, less than 10% of articles reported something related to these indicators of panel performances (Visalli & Galmarini, 2023), and only a few methodological articles addressed these questions

(Castura, Antúnez, et al., 2016; Frascolla et al., 2023). Consumers are not expected to be individually repeatable as trained panellists, and their repeatability should typically not be controlled before measurements. However, documenting consumers' repeatability *a posteriori* is important to get an idea of the relative amount of information compared to noise in the data (Visalli, Béno, et al., 2023).

The ability of temporal methods to discriminate between temporal events within and between products depends on the precision of the measurements. The more precise the measurements, the more the methods will be able to detect short-duration temporal differences between two times of selection/unselection of attributes. Thus, understanding the differences between methods in terms of panellist agreement and repeatability might help to explain the differences in discriminability (Meyners, 2020). Discriminability depends on the resolution of the measuring tool. In metrology, the resolution is the smallest change that can be measured by an instrument. For example, it is not possible to robustly measure differences of the order of a millimeter with a ruler graduated in centimeters. In the context of sensory perception, the temporal resolution relates to the sensitivity to temporal difference between two sensations (Gotow, 2019). When temporal perception is measured with panels of subjects using TDS or TCATA, it therefore seems reasonable to consider that temporal differences with durations lower than the temporal resolution of the method used may be at best uncertain, at worst meaningless. The temporal resolution of TDS and TCATA is not documented. Indeed, it is not because at the individual level these methods allow data to be collected in continuous time during tasting that they can reliably capture short-duration temporal signals at the panel level.

TDS and TCATA curves are usually used to represent the within-product evolution of citation rates of attributes evaluated as dominant or applicable (y-axis) over time (x-axis). Citation rates reflect at the panel level the agreement between the panellists for a given attribute at any discretized time. TDS and TCATA difference curves enable to compare the temporal evolution of the difference in citation rates of dominant/applicable attributes between two products (or two set of products). To be interpretable, these plots require a minimum of homogeneity between panellists in temporal perception, a situation which is not always verified. Indeed, even with TI, with only one attribute assessed at a time, high resolution in the temporal dimension can make it difficult to separate variations due to products from those

due to differences in perception between the panellists (Overdo-López et al., 2005). To remove differences in times to first attribute selection and times to click on stop, times can be standardized (Lenfant et al., 2009). Time standardization transforms the time information in seconds into a percentage of the total duration of the product evaluation. With TCATA, Dietz, Yang, et al. (2022) showed that time standardization only marginally reduced noise in data. It is therefore unclear whether temporal standardization of data improves the temporal resolution of the methods.

The uncertainty on the y-axis of TDS/TCATA curves/difference curves is controlled, because tests of comparison of citation rates (comparison to the chance level in TDS, or to the average citation rate measured in the other products with TCATA) are carried out at each time for each attribute. However, whether the data are standardized or not, variations around times of selection/unselection of attributes are not directly observable on the TDS/TCATA curves/difference curves, and the uncertainty on the x-axis (times of begin/end of perception) is not controlled. As the temporal resolution of TDS and TCATA is unknown, it may be difficult to objectify and robustly interpret significant periods of dominance/applicability observed on TDS/TCATA curves or difference curves for short durations, particularly those with the lowest magnitudes. The subjectivity of TDS data interpretation was illustrated by Galmarini (2023) who presented the results of a survey administered to members of the European Sensory Network. The participants (57, all reporting experience with TDS) were asked to describe TDS curves of three chocolates (see Supplementary material, Annex 1). Although these survey data come from an uncontrolled study and should be interpreted as such, a large heterogeneity was observed in the way the curves were described. Some people have interpreted only large changes in dominance, while others have interpreted citation rates barely exceeding the significance level and for very short durations. Among the three presented curves, two represented replicated samples of the same chocolate. More than 60% of participants concluded that the samples were different based on a visual comparison of the two sets of curves.

This work is a continuation of the research of Visalli et al. (2020, 2022, 2023) aimed at determining what typical precision and temporal resolution of TDS and TCATA are when data are collected from consumers. The objectives of this study were: (i) to evaluate the levels of imprecision of TDS/TCATA due to the heterogeneity of temporal measurements among

consumers (non-repeatability) or between consumers (disagreement), with or without time standardization; (ii) to propose a new framework of analyses common to TDS and TCATA making it possible to test and represent the differences between times of selection and unselection of attributes, within and between products; (iii) to compare the conclusions obtained with the new framework with those obtained with usual analyses (TDS/TCATA curves and difference curves); (iv) to determine the temporal resolution of the methods based on (ii). In order to obtain the most generalizable conclusions possible, the investigation was carried out on several TDS and TCATA datasets (Visalli, Cordelle, et al., 2022). These datasets included consumer evaluations of samples in four product categories (crisps, dark chocolates, guacamoles and ice teas) of varying sensory complexity. Based on the results, recommendations are provided for interpreting temporal data to draw robust conclusions.

2. Material and methods

2.1. Protocol

We now provide a summary of the protocol used in this study. A detailed description of this protocol is provided by Visalli, Cordelle, et al. (2022).

2.1.1. Participants

A total of 436 subjects who reported being regular consumers of the categories of products participated in this study. They were preselected from a population registered in the ChemoSens Platform's PanelSens database (declared to the relevant authority, Commission Nationale Informatique et Libertés - CNIL - authorization number 1148039) and compensated for their participation with a voucher worth 15 euros. The research was carried out following the Declaration of Helsinki. The consumers were informed of the conditions for participating and they signed an informed consent written form. They were able to withdraw from the study at any time without giving a reason. The products tested were commercial products, safe for consumption.

The consumers were randomly assigned to one of six panels with balanced gender and age. Each panel used a different temporal method and participated in three sessions: one in the lab and two at home. This study primarily focuses on the results of the two panels that used TDS (n=68 to 70) and regular TCATA (n=72 or 73).

2.1.2. Products and attributes

The chosen products were commercial products that were accepted by a majority of consumers, easy to prepare for the experimenters and varying in composition, texture and sensory complexity.

Each panel had to describe all the products from each of the four product categories using a single sensory evaluation method. The products were evaluated over three sessions, one in the lab and two at home on two different days. Within each product category, one product was replicated (code suffixed by “_rep”).

Table 55. Products evaluated by the consumers.

Session	Location	Product category	Codes	Sensory complexity	Attributes
1	Lab	Guacamoles	G1 to G4	High	Smooth/Fat, Sour, Bitter, Avocado, Lemon, Spicy, Onion/Shallot, Aromatic/Herb, Thick, Tomato, Fluid, Sweet
1	Lab	Ice teas	IT1 to IT4	Low	Peach, Tea, Watery/Diluted, Bitter, Artificial, Sweet, Astringent/Rough, Sour
2	Home	Crisps	C1 to C4	Low	Crackly/Hard, Crispy, Potato, Roasted, Salty, Bland, Melting, Fat, Sticky/Pasty
3	Home	Dark chocolates	CH1 to CH4	High	Cocoa, Woody/Roasted, Sweet, Artificial, Sticky, Dry/Powdery, Bitter, Fruity, Floral, Melting/Fat, Sour, Astringent

Table 55 shows the session, location, product category, codes, sensory complexity and attributes used to describe the products. Sensory complexity was evaluated using an adaptation of the questionnaire of Schlich et al. (2015) and the level of complexity determined according to the results of a PCA of the averaged scores of the items of the questionnaire (read more in Visalli, Cordelle, et al., 2022).

To select the attributes, a pretest was performed by an internal panel of six persons (ChemoSens staff, untrained but familiar with temporal methods). The selected attributes were presented in a random order on the screen but this order was constant for each

consumer across the evaluations of the same product category. The list of attributes was the same for TDS and TCATA.

The samples were evaluated in a fixed order indicated by the number in the sample code. The replicated sample was always presented at the fifth rank. This design is unusual; this study was purely methodological and thus designed with the objective of comparing temporal methods, not products. In this specific case, fixing the order of testing minimized the variance introduced by different orders of presentation across the panellists (Bell, 2013; Humes et al., 2013). However, with this process, an order effect may have affected the sample descriptions despite a forced 30-second break and mouth-rinsing before assessing each sample.

2.1.3. Experimental procedure

All consumers were invited to a collective briefing at the lab in groups of 16. The temporal method was explained, and the sensory attributes were orally presented (only for food categories evaluated in laboratory). A demonstration of the task was done by the panel leader by video projection of the screens of the software used for data acquisition (TimeSens© V2, INRAE, ChemoSens, Dijon). Then, the consumers went into individual booths. The instructions for each temporal method were shown on the screen for a minimum of 30 seconds to ensure that the consumers read the instructions. For TDS and TCATA, consumers had to click a start button to signal when they put the sample in their mouth (starting the software's chronometer measuring the total duration of perception), and they had to click a stop button to signal that they no longer perceived anything (stopping the chronometer). For TDS, consumers were asked to click on the button corresponding to an attribute as soon as they perceived it as dominant (i.e. when the attribute attracted their attention). For TCATA, they were instructed to check the box corresponding to an attribute as soon as they perceived it, and to uncheck it when they no longer perceived it. For both methods, they were informed that some attributes may never be selected, and others may be selected multiple times during the tasting. Detailed instructions are provided by Visalli, Cordelle, et al. (2022).

The consumers evaluated the five samples of guacamole and, after a forced five-minute break, the five ice teas. The duration of the session was approximately 45 minutes. Then, each consumer took samples of chocolates and crisps to their homes. The samples of crisps were stored in individual disposable plastic cups with lids, and the chocolate samples were packed

into aluminium foil. The consumers were informed that they would receive two separate e-mails to connect to the Internet sessions using a web browser. They received their first e-mail the same day of the tasting in the lab, and a second e-mail was sent to them the day after they completed session 2, when they were instructed to evaluate the samples of chocolate.

2.2. Data analysis

The time of selection of an attribute was defined as the time between the click on the start button and the click on the button/checkbox of an attribute indicating that it was perceived as dominant/applicable. The time of unselection of an attribute was defined as the time between the click on the start button and either: (i) the time of the click on the button of the next dominant attribute (TDS), (ii) the time of the click on the checkbox to indicate that the current attribute was no longer perceived as applicable (TCATA), or (iii) the time of the click on the stop button (TDS/TCATA).

For all computations, the times of selection/unselection of attributes not selected by a consumer were considered as unobserved data (NA). As a result, the difference between a time of selection and a time of unselection of an attribute not selected was also considered unobserved data (and not zero as the duration used in usual analyses such as ANOVA, PCA or CVA).

To take into account multiple citations of a same attribute, when an attribute was cited several times non-consecutively by a same consumer, the multiple times of selection/unselection of the attribute were considered as distinct temporal events. In this case, the attribute was suffixed by a number (e.g. "Salty_2" for the second citation of "Salty").

All statistical analyses were performed using R 4.1.0 software (R Core team, 2022) with an alpha risk (significance level) of 5%.

2.2.1. Assessment of the imprecision of TDS and TCATA measurements

First, the total durations of perception (tdp) were estimated per product and method by calculating the means and standard deviations of the times to click on stop. The means and standard deviations of the number of distinct attributes selected and of the number of citations per consumer were also computed by product and method to determine if the total duration of perception had an impact on the behaviour towards attributes.

Second, standard deviations (sd) were calculated per product and method for: (i) the times to first attribute selection; (ii) the times to click on stop; (iii) the times of selection of attributes; (iv) the times of unselection of attributes. Standard deviations were computed between consumers (differences in measurements from different consumers for the same product) and within consumers (differences in measurements from the same consumer for the replicated samples, averaged across consumers). The levels of imprecision of each of the four measurements were expressed as “relative standard deviations” ($100 \times \text{sd}/\text{tdp}$). The rationale was to compute some sort of coefficient of variation, but having tdp as the common denominator instead of the mean of the measurement (which would have been lower/larger for attributes cited at the beginning/end for (iii) and (iv)). This makes the absolute values of the relative standard deviations of the four measurements comparable to each other (the lower the better). To test whether time standardization improves the precision of the measurements, calculations of (iii) and (iv) were performed with and without time standardization (sd being equal to 0 for (i) and (ii) after time standardization).

Third, the Jaccard index (Jaccard, 1901) was used to evaluate the level of similarity between two sets of attributes elicited in two sequences of perception. This index was preferred to the Manhattan distance used by Castura et al. (2016) because the latter grants the same importance to cited and non-cited attributes (i.e. the Manhattan distance uses info from non-selection and the Jaccard index does not).

The Jaccard index was first computed without taking into account the temporal dimension of the data (“CATA transformation”):

(1) $J_{\text{static}} = (\text{number of attributes cited both in sequence 1 and 2} / \text{total number of different attributes cited in sequence 1, 2 or both})$

The Jaccard index was then computed at each discretized time t_i (0.1 s or 1% of the total duration with time standardization), then averaged over all times, as described in Frascolla et al. (2023):

(2) $J_{\text{temporal}} = \text{mean}(\text{number of attributes cited both in sequence 1 and 2 at } t_i / \text{total number of different attributes cited in sequence 1, 2 or both at } t_i)$

This was done to determine the extent to which temporal measurements affect the level of performances of qualitative measurements.

To assess agreement (between consumers), J_{static} and J_{temporal} were computed per product, with sequences corresponding to evaluations of the same product by different consumers. To assess repeatability (within consumers), J_{static} and J_{temporal} were computed per product, with sequences corresponding to evaluations of the replicated samples by the same consumer. One value of J_{static} and J_{temporal} was obtained by consumer (or couple of consumers for the agreement) and product. Thus, to determine agreement and repeatability at the panel level, all J_{static} and J_{temporal} were averaged across consumers. The average values can be between 0 (i.e. total disagreement/non-repeatability) and 1 (i.e. perfect agreement/repeatability), the larger the better.

2.2.2. Tests for difference between median times of selection/unselection of the attributes

The median times of selection/unselection of the attributes and their confidence intervals were computed using bootstrap. Medians were used instead of means because they are unaffected by outliers or non-symmetric distributions that were observed in the data. Bootstrapping was used as an alternative to statistical inference because the assumption of the normality of the distribution was not verified with times of selection/unselection of attributes. The bootstrap approach is described below:

Step 1: Drawing of virtual panels. One thousand bootstrapped panels were drawn. The number of consumers drawn with replacement at random for each panel was chosen as the minimum number of consumers who evaluated the samples in the two methods (e.g. $n=68$ for crisps). This was done to avoid a panel size effect and therefore allow a fair comparison between TDS and TCATA.

Step 2: Computation of medians of citation rates and times of selection/unselection of the attributes. Within each bootstrapped panel, for each “attribute x product” pair, three statistics were computed at the panel level: the median time of selection of the attribute, the median time of unselection of the attribute, and the median citation rate of the attribute.

Step 3: Computation of confidence intervals of median citation rates and times of selection/unselection of the attributes. The three statistics computed in Step 2 were aggregated as medians across the 1000 bootstrapped panels. The confidence intervals of the

median of these medians were determined from the $(\alpha/2)$ and $(1-\alpha/2)$ quantiles of the bootstrapped distribution.

Step 4a: Tests for within-product differences between median times of selection/unselection of the attributes. Within each bootstrapped panel, for each product, the differences between median times of selection/unselection of the attributes were computed (i) for each attribute and (ii) for each pair of attributes. Then, the medians of the differences between median times and their corresponding confidence intervals were computed from the $(\alpha/2)$ and $(1-\alpha/2)$ quantiles of the bootstrapped distribution. A difference between two median times of selection/unselection of attributes was considered significant when its confidence interval did not contain 0.

A significant difference between the median time of selection and the median time of unselection for a same attribute (e.g. “median time of unselection of *Salty* – median time of selection of *Salty*”) means that at the panel level there was relative consensus among consumers in the periods at which they began and finished perceiving this attribute (i.e. the situation where half of the consumers has finished perceiving the attribute before the other half has started perceiving it is not verified).

A significant difference between the median time of selection (unselection) for two distinct attributes A1 and A2 (e.g. “median time of selection (unselection) of *Salty* – median time of selection (unselection) of *Fat*”) means that the median time of begin (end) of perception of A1 for consumers who cited A1 was significantly lower than the median time of begin (end) of perception of A2 for consumers who cited A2 (not necessarily the same consumers as those who cited A1). If overall citation rates were close to 100% for both attributes, this means that for most consumers A1 started (ended) to be perceived as dominant/applicable before A2. Similarly, a significant difference between the median time of selection and the median time of unselection for two distinct attributes A1 and A2 (e.g. “median time of selection of *Salty* – median time of unselection of *Fat*”) means that for most consumers A2 stopped to be perceived before A1 started to be perceived.

Step 4b: Tests for between-product differences between median times of selection/unselection of the attributes. Within each bootstrapped panel, for each pair of products, the differences in median times of selection (unselection) were computed for each

attribute. A significant difference means that for consumers who cited the attribute in the first product, the median time of begin (end) of perception of the attribute was significantly lower than the median time of begin (end) of the attribute in the second product for consumers who cited the attribute in the second product. If overall citation rates of the attribute were close to 100% in both products, it means that for most consumers the attribute started (ended) to be perceived as dominant/applicable earlier in the product 1 than in product 2.

Step 5: The groups of attributes selected/unselected at significantly different periods were determined post-hoc using the compact letter display (CLD) algorithm (Piepho, 2012). Here, numbers have been used instead of letters to better evoke the idea of order of temporal events.

2.2.3. Representations of within and between-product differences

The evolution of citation rates over time and standardized time were represented as TDS curves (Pineau et al., 2009) and TCATA curves (Castura et al., 2016). The curves were smoothed using locally estimated scatterplot smoothing (LOESS, Cleveland, 1979). For TDS, significance was represented by a horizontal line whose ordinate corresponded to the minimum value the dominance rate would have to equal to be considered significantly greater than ($1/\text{number of attributes}$), based on a binomial test.

Differences in citation rates over time were plotted using TDS and TCATA difference curves. The significance of differences was assessed using the McNemar's proportion test (McNemar, 1947) with Yates' continuity correction to prevent overestimation of statistical significance for expected counts smaller than five (Yates, 1934). Significant differences were highlighted as bold lines on the curves.

To demonstrate the importance of taking into account the temporal imprecision of the measurements, the TDS/TCATA curves and the difference curves were compared to "time-variability plots". Time-variability plot are new representations of TDS/TCATA data based on bootstrap test results and enabling to visualize uncertainty in times of selection and times of unselection of the attributes. The attributes were represented as horizontal bands with heights proportional to median citation rates. They were either grouped by product then ordered on the y-axis by decreasing median times of selection (within-product differences, see Step 4a of Section 2.2 and Figures 58.B, 58.D, 59.B and 59.D), or grouped by product

(between-product differences, see Step 4b of Section 2.2 and Figures 60.C-K and 61.C-K). The left coordinate of each band corresponds to the median time of selection of the attribute, and the right coordinate corresponds to the median time of unselection of the attribute. The confidence intervals around median times were represented as black horizontal lines. The post-hoc groups (Step 5 of Section 2.2) were indicated as numbers to the left (group for time of selection) and right (group for time of unselection) of the bands. Multiple groups were summarized by the first and last numbers separated by a hyphen (e.g. “1, 2, 3, 4” was summarized “1-4”). Attributes having median citation rates lower than the difference between their upper and lower confidence level limits were not represented. The rationale was to remove attributes having median citation rates too small in regard to their levels of heterogeneity observed with bootstrapping.

2.2.4. Assessment of the temporal resolution of TDS and TCATA measurements

The bootstrap tests described in Section 2.2.2 made it possible to discriminate between times of selection or unselection of attributes within and between products. The significance of these tests reflects the ability of the temporal methods to capture with confidence the differences in periods of perception of dominant/applicable sensory attributes while accounting for temporal imprecision due to disagreement among consumers. Thus, the temporal resolution of the methods was assessed by product category as the minimum time (in seconds) elapsed between two temporal events (times of selection/unselection of the same/different attributes) so that they are significantly discriminated at the panel level. The calculation was carried out within and between products, for each type of event (selection or unselection of an attribute), for all product categories.

3. Results

3.1. Precision of TDS and TCATA measurements

3.1.1. Differences between product categories and methods in total durations of perception

Table 56. Mean and standard deviation of total duration of perception (in second), number of distinct attributes selected and number of citations per consumer, by method and product category.

	Total duration TDS	Total duration TCATA	Number of attributes TDS	Number of attributes TCATA	Number of citations TDS	Number of citations TCATA
Crisps	28.1 ± 13.4	35.5 ± 17.4	3.8 ± 1.3	4.1 ± 1.2	5.1 ± 3.0	4.4 ± 1.5
Ice teas	30.9 ± 13.5	41.8 ± 24.2	3.8 ± 1.2	3.8 ± 1.2	5.5 ± 2.7	4.3 ± 1.9
Guacamoles	36.7 ± 13.2	43.9 ± 12.8	4.6 ± 1.7	5.0 ± 1.7	6.3 ± 3.2	5.3 ± 2.0
Chocolates	46.2 ± 20.1	55.0 ± 25.5	4.3 ± 1.5	4.6 ± 1.6	6.0 ± 3.0	5.1 ± 2.5

Table 56 shows that the total duration of perception depends on the product category. The longest duration is observed for chocolates, a category of products with a hard-solid matrix that requires more oral processing than other product categories. The ranking of durations of perception is the same with TDS and TCATA. However, durations of perception are systematically longer with TCATA compared to TDS. This result suggests either that consumers did not interpret the TDS and TCATA instructions in the same way, or that the data collection method impacted the measurement of durations.

Table 56 also shows that overall, a few more distinct attributes were selected on average with TCATA compared to TDS. However, the mean number of citations was higher in TDS compared to TCATA, and on average more attributes were selected several times in TDS compared to TCATA. The mean duration of perception was between 6 and 8 s for a dominant attribute with TDS, and between 8 and 10 s for an applicable attribute with TCATA. Overall, the mean durations of perception of the attributes were longer for products having the longest total durations of perception.

3.1.2. Overall levels of imprecision of TDS and TCATA measurements

Table 57. Relative standard deviation (RSD) of temporal measurements (B: between consumers; W: within consumer; std: after time standardization). Values represent average percentages (100 x standard deviation of the value/mean total duration) across the four product categories (minimum and maximum are indicated between brackets). The lower the RSD the lower the imprecision

Measured value	TDS	TCATA
Times to click on stop	B: 42 (36-48) W: 12 (10-13)	B: 45 (29-58) W: 12 (9-14)
Times to first attribute selection	B: 15 (13-19) W: 4 (4-5)	B: 9 (6-11) W: 3 (2-4)
Times of selection of attributes	B: 27 (24-29) B _{std} : 24 (23-25) W: 14 (13-15) W _{std} : 14 (12-15)	B: 22 (19-27) B _{std} : 22 (20-25) W: 11 (8-12) W _{std} : 11 (9-12)
Times of unselection of attributes	B: 34 (31-35) B _{std} : 23 (15-29) W: 12 (9-14) W _{std} : 12 (10-14)	B: 38 (31-41) B _{std} : 27 (26-30) W: 18 (16-20) W _{std} : 16 (14-17)

Table 57 shows the most imprecise measurements are related to the times to click on stop: the relative standard deviations (RSD) vary on average between 42 (TDS) and 45% (TCATA) between consumers (larger for the product categories with the lowest total durations of perception), and around 12% within consumers (roughly the same level for all product categories). The RSD of these measurements are similar between TDS and TCATA, with larger variations with TCATA. The RSD of the times to first attribute selection vary on average between 9 (TCATA) and 15% (TDS) between consumers (larger for the chocolates), and about 4% within consumers (about the same level for all product categories). The RSD of these measurements are slightly better with TCATA.

The RSD of the times of selection of the attributes vary on average between 22 (TCATA) and 27% (TDS) between consumers (larger for chocolates and lower for guacamoles), and between 11 and 14% within consumers (about the same level for all product categories). Time standardization only slightly reduces the RSD between consumers for TDS. The RSD of the times of unselection of the attributes vary on average between 34 (TDS) and 38% (TCATA) between consumers (lower for guacamoles and about the same level for the other product categories), and between 12 (TDS) and 18% (TCATA) within consumers (about the same level for all product categories). Time standardization reduces RSD between consumers for TDS and TCATA, but it has almost no impact on the precision within consumers. No method seems to

be more imprecise than the other. TCATA is slightly less imprecise regarding the measurements of times of selection, and TDS regarding the measurements of times of unselection.

3.1.3. Overall levels of agreement and repeatability of TDS and TCATA measurements

Table 58. Agreement and repeatability of temporal measurements. Values represent the Jaccard index averaged across the four product categories (minimum and maximum are indicated between brackets). The higher the better.

Indicator	TDS	TCATA
Agreement in selection of attributes after CATA transformation	0.34 (0.32-0.41)	0.39 (0.35-0.43)
Agreement in selection/unselection of attributes at each time point (mean over times)	0.09 (0.07-0.10)	0.17 (0.11-0.16)
Agreement in selection/unselection of attributes at each standardized time point (mean over times)	0.17 (0.14-0.20)	0.24 (0.22-0.28)
Repeatability in selection of attributes after CATA transformation	0.48 (0.40-0.58)	0.47 (0.45-0.52)
Repeatability in selection/unselection of attributes at each time point (mean over times)	0.15 (0.11-0.18)	0.25 (0.20-0.28)
Repeatability in selection/unselection of attributes at each standardized time point (mean over times)	0.20 (0.14-0.23)	0.32 (0.24-0.36)

Table 58 shows that after CATA transformation, on average and for a same product two sequences of attributes cited by two different consumers have between 32 and 43% of attributes in common (lower for the product categories having the longest total durations of perception), with a slight advantage for TCATA. Repeatability is comparable between the two method, with less variability among product categories with TCATA. On average, consumers selected approximately 50% common attributes when evaluating replicated samples (higher for ice teas, lower for guacamoles).

When considering selection and unselection of attributes over time, agreement and repeatability decrease by a ratio of 2 (TCATA) to 4 (TDS). This result shows that considering the temporal dimension of the data at a fine grain (to the second/percentage of total duration) drastically reduce the computed values for agreement and repeatability of sensory measurements. The agreement was higher for crisps and ice teas, the repeatability was lower for guacamoles, the two indicators of performances being about the same level for the other product categories. Time standardization improves both indicators of performances, which are better in TCATA compared to TDS.

3.2. Tests and representations of within-product differences between attributes with and without considering temporal precision of the measurements

In this section, the product C1 (Crisp) was chosen as an example because its sequentiality was well marked (see Supplementary material for other examples in each product category).

3.2.1. Comparison of TDS curves and time-variability plots

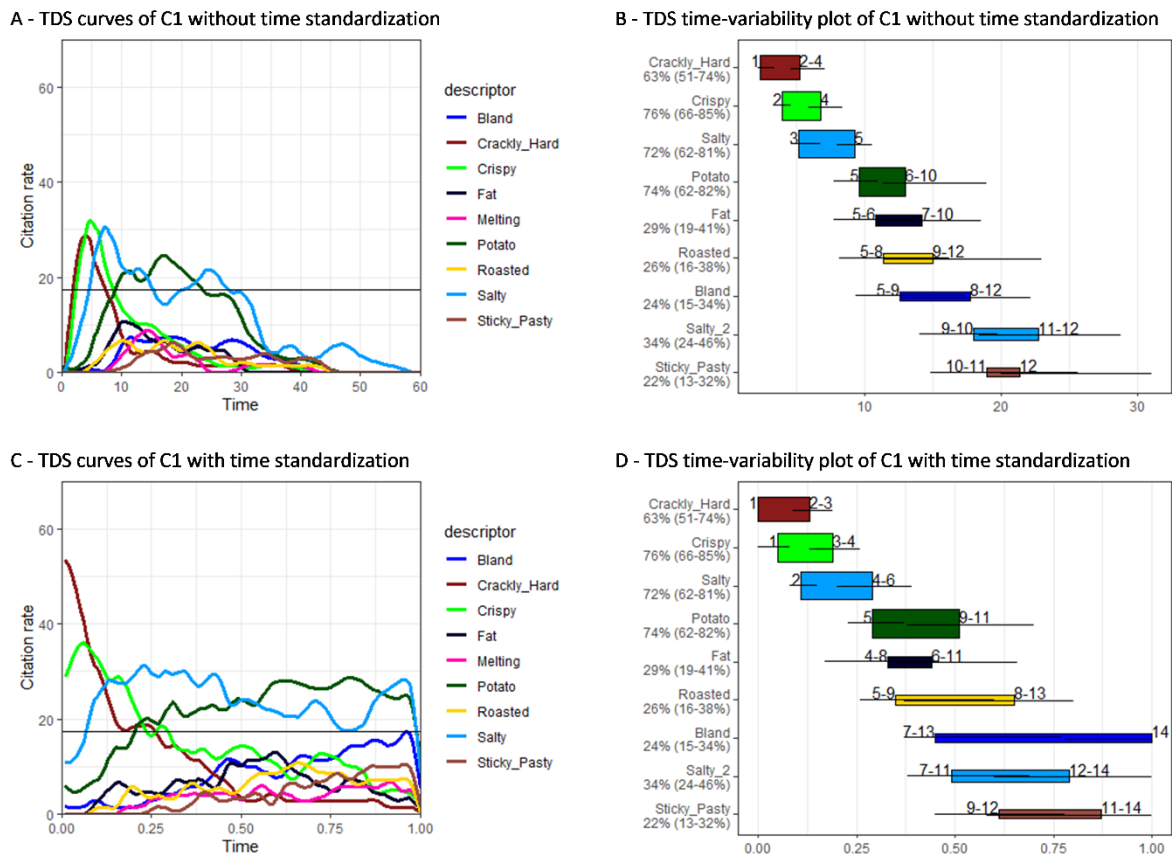


Figure 58. Representation of within-product differences between attributes for Crisp C1 (TDS). On Figures A and C (TDS curves), the horizontal lines correspond to the significance level of the binomial test. On Figures B and D (time-variability plots), the percentages below the name of the attributes correspond to the median citation rates and their confidence intervals. Box heights are proportional to the median citation rates. Left and right box borders correspond to median times of selection and unselection. Horizontal lines crossing the boxes are confidence intervals of median times of selection and unselection. Post-hoc groups of attributes with shared numbers (above confidence intervals) are not significantly different in median times of selection or unselection (see Section 2.2.2).

Figure 58.A (TDS curves) suggests that *Crackly/Hard* (from 2 to 7 s) and *Crispy* (from 2 to 8 s) were the first dominant sensations, followed by *Salty* (from 5 to 15 s), *Potato* (from 10 to 23 s), and *Salty* again (from 20 to 26 s).

Figure 58.B (time-variability plot) shows that *Crackly/Hard* was selected by 63% (confidence interval: 51-74%) of consumers at 2.4 s in median and unselected at 5.3 s in median. In details (only visible in raw data), among consumers who perceived *Crackly/Hard*, half selected the attribute between 1.0 and 2.4 s then unselected it after a median duration of 2.2s, and the other half selected it between 2.5 and 19.0 s then unselected it after a median duration of 3 s. This result shows that the TDS curves (Figure 58.A) of *Crackly/Hard* is not the result of 30% on consumers having consensually perceived *Crackly/Hard* as dominant for 5 s between 2 and 7 s, but of the accumulation of consumers having delayed perceptions of the attribute for a short duration (2.5 s in median).

The same interpretation holds with *Crispy*. Figure 58.B shows that the attribute was selected by 76% (confidence interval: 55-85%) of consumers at 4 s and unselected at 6.8 s in median. In details (only visible in raw data), 38% percent of consumers selected the attribute between 0.7 and 4.0 s then unselected after a median duration of 2.3 s, and 38% selected it between 4.0 and 38.5 s then unselected it after a median duration of 2.6 s.

The *Crackly/Hard* and *Crispy* curves are superimposed on Figure 58.A. Fifty-one percent of the consumers perceived both *Crackly/Hard* and *Crispy* (only visible in raw data). Among these 35 consumers, 72% perceived *Crackly/Hard* before *Crispy*. This behaviour is captured by the time-variability plots (Figure 58.B): the confidence interval ([0.6 ; 2.2]) of the difference between the median times of selection of the attributes does not intersect zero. It means that in median, the consumers (who perceived *Crackly/Hard*, post-hoc group 1) started to perceive *Crackly/Hard* earlier than the consumers (who perceived *Crispy*, post-hoc group 2) started to perceive *Crispy*. On the contrary, while it seems that *Crispy* ended to be perceived as dominant shortly after *Crackly/Hard* on the TDS curves, the time-variability plot shows that the difference between the times of unselection is not large enough to be considered as significant regarding the variability in consumer perception.

Salty was selected by 72% of consumers at 5.2 s and unselected at 9.3 s in median (Figure 58.B). *Potato* was selected by 74% of consumers at 9.6 s and unselected at 13 s in median. Fifty-three percent of the consumers perceived the two attributes (only visible in raw data), among them 76% perceived *Salty* before *Potato*. While in median the two periods of perception of the attributes are separated (as observed on TDS curves), the confidence interval ([-2 ; 1.6]) of the difference between the median time of unselection of *Salty* and the

median time of selection of *Potato* intersects zero. The difference is not large enough to be considered as significant because of too much heterogeneity among consumers regarding the two events.

The time-variability plots show that *Salty* was selected a second time (*Salty_2*) by 34% of consumers at 18 s and unselected at 22.8 s in median (Figure 58.B), which is only suggested by the TDS curves (Figure 58.A). However, due to a relatively low percentage of citation, the confidence intervals of the times of selection and unselection of the attribute are very large, which represents the uncertainty in the second period of perception of *Salty*. Similarly, *Fat*, *Roasted*, *Bland* and *Sticky/Pasty* (Figure 58.B) were selected by a non-negligible percentage of consumers (29%, 26%, 24% and 22%, respectively), but they did not reach significance on TDS curves (Figure 58.A) due to too much heterogeneity in times of selection/unselection of these attributes. Despite their large confidence intervals, a temporal pattern can be observed for these attributes. Indeed, the confidence intervals of the difference between the times of unselection of *Fat*, *Roasted* and *Bland* (post-hoc groups 5 to 9) and the times of selection of *Crackly/Hard* and *Crispy* (post-hoc groups 2 to 4) do not intersect.

For *Bland*, the median times of selection are not significantly different from the median times of unselection. This shows that half of the consumers who perceived *Bland* ended to perceive the attribute as dominant before the other half started to perceive it.

Figure 58.C (time-standardized TDS curves) suggests that time standardization results in a temporal pattern similar to that in Figure 58.A.

Figure 58.D (time-standardized time-variability plot) is relatively similar to Figure 58.B, except for *Crispy* (median time of selection not significantly different from those of *Crackly/Hard*, median time of unselection not different from those of *Salty*) and *Bland* (median time of selection different from those of *Potato*, median time of unselection different from those of *Roasted*, *Fat* and *Potato*). The size of the bars and of the confidence intervals are larger with time standardization, notably for *Potato*, *Roasted*, *Bland*, *Salty_2*, and *Sticky/Pasty*. This result shows that time standardization stretches the periods of perception. With the time-variability plots, this is particularly visible with attributes having low citation rates (e.g. *Roasted*, *Sticky/Pasty*) and/or non-consensual times of selection/unselection (e.g. *Bland*).

3.2.2. Comparison of TCATA curves and time-variability plots

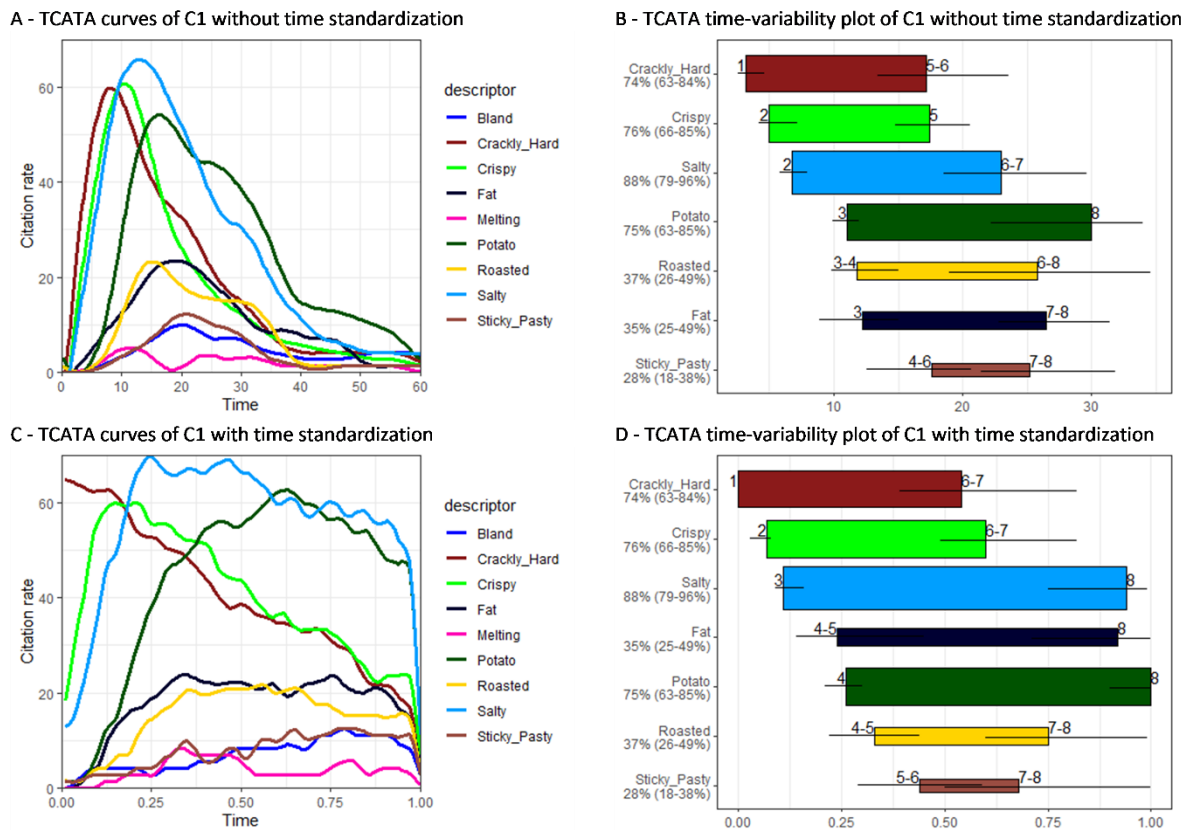


Figure 59. Representation of within-product differences between attributes for Crisp C1 (TCATA). Figures A and C are TCATA curves. On Figures B and D (time-variability plots), the percentages below the name of the attributes correspond to the median citation rates and their confidence intervals. Box heights are proportional to the median citation rates. Left and right box borders correspond to median times of selection and unselection. Horizontal lines crossing the boxes are confidence intervals of median times of selection and unselection. Post-hoc groups of attributes with shared numbers (above confidence intervals) are not significantly different in median times of selection or unselection (see Section 2.2.2).

Figure 59.A (TCATA curves) suggests that *Crackly/Hard*, *Crispy* and *Salty* were the first attributes perceived as applicable with minor delays between the moments of perceptions. *Potato* started to be perceived about 5 s later. *Fat* and *Roasted* were perceived with lower citation rates just before and after *Potato*, but the different slopes suggest a delayed perception among the consumers. A few consumers perceived *Sticky/Pasty* and *Bland* in the middle of the tasting. The curves also suggest that for a majority of consumers, *Crackly/Hard*, *Crispy*, and *Roasted* ceased to be perceived first. *Salty* and *Fat* ceased to be perceived 5 to 10 s later, and *Potato* was perceived until the end.

Figure 59.B (time-variability plot) shows that *Crackly/Hard* was selected by 74% of consumers, at 3.2 s in median (post-hoc group 1). This median time of selection was significantly lower than those of *Crispy* and *Salty* (5.0 and 6.8 s, post-hoc group 2), themselves lower than those of *Potato, Roasted, Fat* and *Sticky_Pasty* (11.0, 11.8, 12.2 and 17.6 s, post-hoc groups 3 to 6). *Crackly/Hard* was unselected at 17.2 s in median (post-hoc groups 5-6), a time significantly lower than those of *Potato, Fat* and *Sticky/Pasty* (30.0, 26.5, and 25.2 s, post-hoc groups 7 to 8), but not different significantly from the median times of unselection of *Crispy, Salty* and *Roasted* (17.4, 23.0, and 25.8 s, post-hoc groups 5 to 8).

Figure 59.C (time-standardized TCATA curves) separate slightly more the curves of *Crispy* and *Salty* at the beginning of the perception, and slightly less the curves of *Salty* and *Potato* at the end.

Figure 59.D (time-standardized time-variability plot) is relatively similar to Figure 59.B, except for a significant difference in times of selection between *Crispy* and *Salty*. The periods of perception of *Salty* are longer after time standardization, suggesting that this attribute was often unselected the last but with at various times.

The interpretation of TDS/TCATA curves and TDS/TCATA time-variability plots do not lead to exactly the same conclusions. These results suggest that interpreting the periods of perception of the attributes can be tricky when the slopes of the curves increase or decrease slowly, when the curves overlap, or when there are several groups of consumers having very different periods of perception for the same attribute. The time-variability plots made it possible to visualize the uncertainty on the periods of perception, and the post-hoc tests to objectify the differences between the median times of selection/deselection of the attributes. They complemented the TDS/TCATA curves and helped to determine whether short-duration signal can be interpreted when accounting for the uncertainty due to the heterogeneity of consumers' temporal perception).

3.3. Tests and representations of between-product differences by attribute

In this section, two pairs of samples were compared: C1 vs. C1_rep, to assess the non-repeatability of the measurements (in this study non-repeatability might be combined with an order effect, see Section 2.1.2); and C1 vs. C4, to assess whether temporal differences exist in the perception of the same attributes between two distinct products. The conclusions drawn

from the difference curves were then compared to those drawn from the time-variability plots by attribute. To limit the number of figures, as the data transformation improves the precision of the measurements, only time-standardized TDS/TCATA curves and time-variability plots are presented (see Supplementary material, Annex 2 for TDS/TCATA curves and time-variability plots without time standardization).

3.3.1. Comparison of TDS difference curves and time-variability plots

Figure 60.A (time-standardized TDS difference curves between C1 and C1_rep) shows that there are significant differences in citation rates between the two replicated samples at the beginning of the tasting. The longer period of difference is observed with *Crackly_Hard* (about 15% of the total durations). Shorter periods of difference are observed with *Salty* and *Crispy* (about 5% of the total duration).

Figure 60.B (time-standardized TDS difference curves between C1 and C4) shows that there are significant differences between C1 and C4. For *Crackly_Hard*, a significant difference in citation rates is observed at the beginning of the tasting during a period of about 25% of the total duration. For *Potato* a significant difference in citation rates is highlighted at the end of the tasting for approximately 35% of the total duration. For *Salty* and *Bland*, significant differences in citation rates are observed during all the tasting.

Figures 60.C to 60.K (time-variability plots between C1, C1_rep and C4) show there are no significant differences in the times of selection/unselection between the replicated samples for the nine attributes when taking into account the temporal imprecision of the measurements. Between C1 and C4, differences were found significant for the times of selection and unselection for *Bland* (Figure 60.C), for the times of selection for *Sticky_Pasty* (Figure 60.K), and for the times of unselection for *Salty* (Figure 60.J), but this difference is not interpretable because *Salty* was cited only by 3% of consumers in C4).

Only one common significant difference between the difference curves and time-variability plots was observed, for *Bland*. While difference curves show that the attribute was consistently more cited in C4 than in C1 (Figure 60.B), the time-variability plot also shows that for consumers who perceived it, the time of selection of *Bland* was significantly delayed in C1 compared to C4 (Figure 60.C).

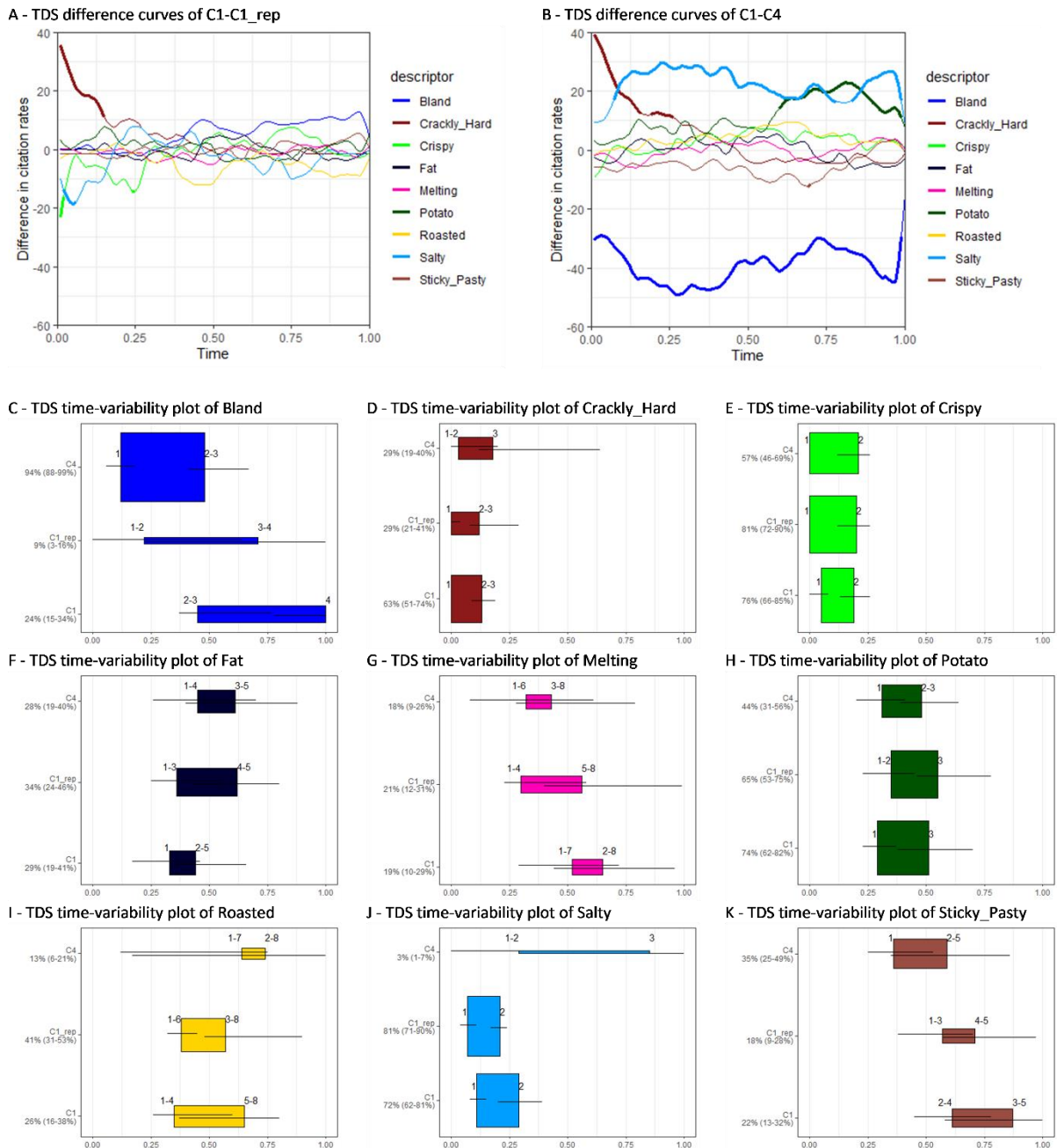


Figure 60. Representation of differences between C1, C1_rep and C4, by attribute (TDS, with time-standardization). On Figures A and B (TDS difference curves of C1 – C1_rep and C1 – C4), bold lines correspond to significant differences in citation rates (proportion test) between C1 and C1_rep (A) and between C1 and C4 (B). On Figures C to K (time-variability plots of C1, C1_rep and C4 by attribute), the percentages below the names of the products correspond to the median citation rates and their confidence intervals. Box heights are proportional to the median citation rates. Left and right box borders correspond to median times of selection and unselection of the attributes. Horizontal lines crossing the boxes are confidence intervals of median times of selection and unselection. Post-hoc groups of attributes with shared numbers (above confidence intervals) are not significantly different in their median times of selection or unselection (see Section 2.2.3).

The differences found significant on either the TDS curves or the time-variability plots can be explained in different ways.

(i) Perception vs. no perception of the attribute regardless of time (e.g. *Salty* was almost not perceived in C4 while it was perceived by more than 20% of consumers at almost any time in C1).

(ii) Disagreement among consumers in the perception of the attribute (i.e. non-intersecting confidence intervals for citation rates) but agreement in the times of selection/unselection (e.g. between standardized times t_0 and t_{12} , *Crackly/Hard* was perceived by 30 to 50% of consumers in C1 and about 15% in C4 at the same period).

(iii) Agreement among consumers in the perception of the attribute but (for at least one product) disagreement among consumers in the times of selection/unselection of the attribute leading to large confidence intervals of the median times of selection/unselection (e.g. *Potato* was perceived in median between standardized times [t23 ; t37] and [t38 ; t70] in C1, and in between standardized times [t20 ; t41] and [t39 ; t64] in C4).

(iv) Agreement among consumers in the perception of the attribute (i.e. intersecting confidence intervals for citation rates) and on temporal differences (e.g. *Sticky_Pasty*).

3.3.2. Comparison of TCATA difference curves and time-variability plots

Figure 61.A (time-standardized TCATA difference curves between C1 and C1_rep) shows that significant differences in citation rates exist between the two replicated samples with TCATA. The longer periods of significant difference are observed at the beginning of the tasting, with *Crispy*, *Salty* and *Crackly_Hard* (for about 5, 15 and 25% of the total duration, respectively).

Figure 61.B (time-standardized TCATA difference curves between C1 and C4) shows that there are significant differences in citation rates between C1 and C4 for *Bland* and *Salty* during all the tasting, but also for *Crackly_Hard*, *Potato*, *Sticky_Pasty* and *Roasted* (50 to 75% of total duration).

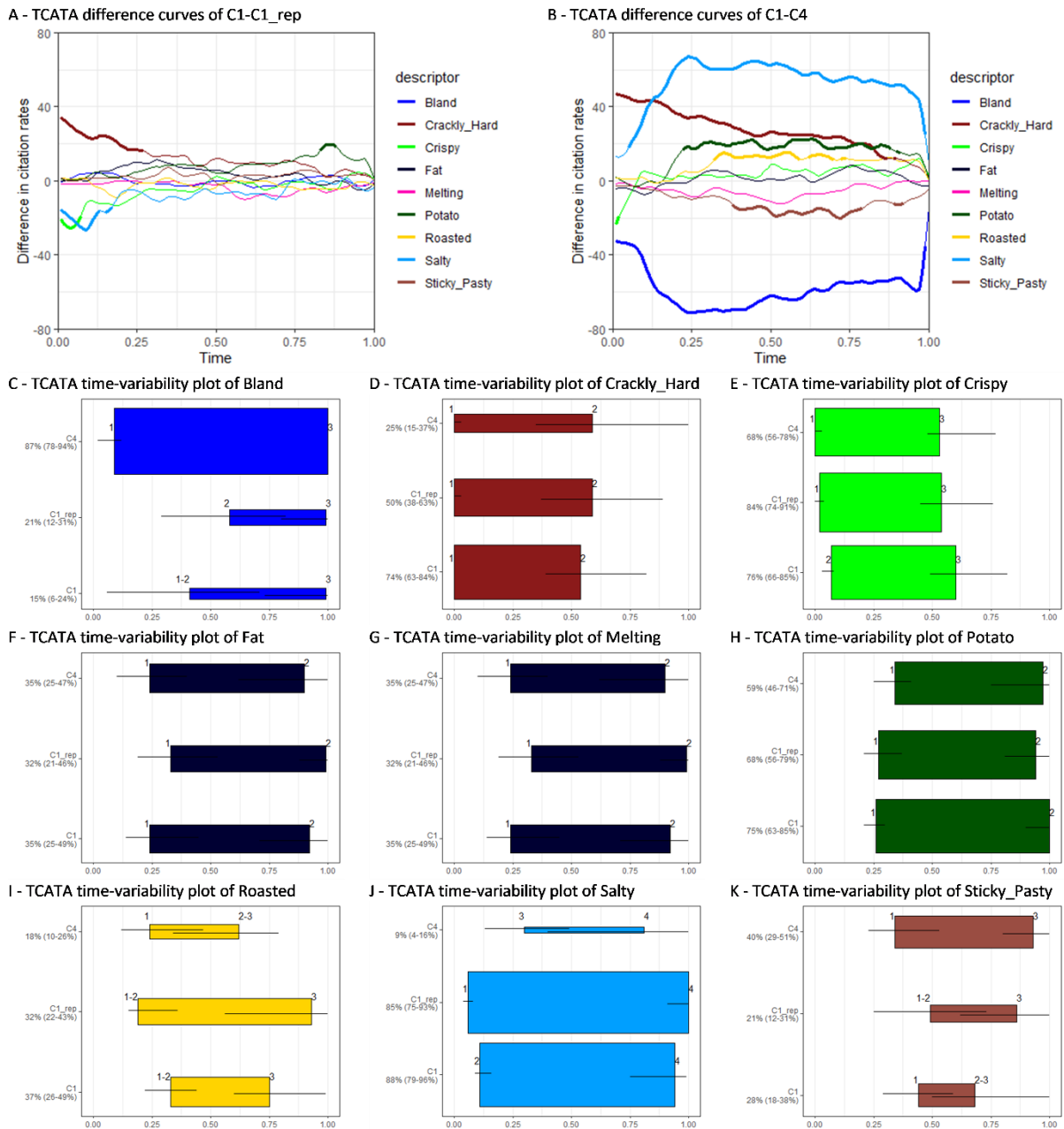


Figure 61. Representation of differences between C1, C1_rep and C4, by attribute (TCATA, with time-standardization). On Figures A and B (TDS difference curves of C1 – C1_rep and C1 – C4), bold lines correspond to significant differences in citation rates (proportion test). On Figures C to K (time-variability plots of C1, C1_rep and C4 by attribute), the percentages below the name of the products correspond to the median citation rates and their confidence intervals of the attributes. Box heights are proportional to the median citation rates. Left and right box borders correspond to median times of selection and unselection. Horizontal lines crossing the boxes are confidence intervals of median times of selection and unselection. Post-hoc groups of attributes with shared numbers (above confidence intervals) are not significantly different in their median times of selection or unselection (see Section 2.2.3).

Figures 61.C to 61.K (time-variability plots between C1, C1_rep and C4) show there are no significant differences in the times of selection/unselection between the two replicated samples except for *Salty* (same explanation as case iv in the previous section). Between C1 and C4, differences were found significant only for the times of selection for *Crispy* (case iv) and *Salty* (case i). The same explanations as in the previous section apply to differences captured with difference curves but not with time-variability plots.

To summarize, Figure 60.A and 61.A show that between-product significant differences between citation rates are not necessarily the same as differences between times of selection/unselection. Indeed, each of the two representations is weak on the concept on which the other representation capitalizes, which makes them complementary. Difference curves make decisions based on “how much” (citation rates) then represent “when” (times), while time-variability plots make decisions based on “when” then represent “how much”. The results show there are much more significant differences related to citation rates than to times of selection/unselection, and that the time-variability plots are less sensitive to repeatability and/or order effects between replicated samples. Some differences are observed between the methods. With difference curves, TCATA highlight significant differences for more attributes and over longer durations compared to TDS (including for replicated samples). With time-variability plots, the level of discrimination of the two methods is comparable. Time-variability plots do not discriminate between the replicated samples with TDS while they do with TCATA. Time-variability plots do capture differences in times of unselection of the attributes with TDS while they do not with TCATA.

3.4. Temporal resolution of TDS and TCATA measurements

The results of sections 3.3 illustrated different cases where TDS/TCATA difference curves and time-variability plots did not lead to the same interpretations on the temporal differences within and between products. This raises questions about the practical significance of (i) the short-durations periods when attributes are above significance line in TDS curves, (ii) the order of perception of different attributes when their curves are separated by short-durations periods, and (iii) the short-durations differences observed on Figures 60.A, 60.B, 61.A and 61.B. Thus, it is proposed to determine the temporal resolution of the methods as the minimum time necessary between two events (times of selection/unselection of same/different attributes) so that they are significantly discriminated.

Table 59. Time necessary between two temporal events so that they are significantly discriminated. Values represent the minimum times observed between two temporal events so that they are significantly discriminated (mean and minimum/maximum between brackets), over all the attributes of the four products (excluding the replicated sample), by product category and method.

Comparison	Product category	TDS - times of selection	TDS – times of unselection	TCATA – times of selection	TCATA – times of unselection
Within-product	Crisp	2.4 (1.3-3)	3.9 (2.2-5.7)	2.2 (1.4-3.0)	6.4 (5.5-7.0)
	Chocolate	5.7 (3.2-7.9)	4.7 (3.8-5.8)	6.1 (4.4-7.0)	10.4 (7.5-16.2)
	Guacamole	3.8 (3.2-4.2)	4.5 (2.8-6.4)	4.2 (2.6-5.2)	5.6 (4.6-6.4)
	Ice tea	3.3 (2.9-4.1)	4.2 (2.8-6.0)	2.9 (2.7-3.4)	6.8 (4.7-8.5)
Between-product	Crisp	4.0 (1.4-6.2)	4.2 (-)	7.8 (1.8-13.6)	7.3 (7.3)
	Chocolate	4.9 (3.2-6.6)	7.8 (4.6-9.6)	10.8 (8.8-17.5)	10.8 (-)
	Guacamole	6.3 (5.0-9.4)	5.4 (3.0-6.8)	5.7 (4.6-7.3)	10.0 (8.0-12.1)
	Ice tea	3.7 (2.0-6.4)	4.8 (-)	3.5 (1.7-7.4)	7.9 (6.4-8.4)

Table 59 shows that within-product, both methods were able to capture some significant differences of the order of the second, but on average the temporal resolution is between 2 and 6 s for TDS, and between 2 and 10 s for TCATA, depending on the attribute and on the total duration of perception of the product category. Transposed as a percentage of the total duration, the temporal resolution is between 9 (Crisps) and 12% (Chocolate) of the total duration of perception for TDS, and between 6 and 18% for TCATA. Both methods performed better to discriminate between times of selection. Between-product, the same general conclusions can be drawn, but overall larger differences are required to conclude on significant differences, in particular with TCATA.

4. Discussion

4.1. The accuracy of the measurements of durations of perception is questionable

The results presented in section 3.1.1 showed that the measurements of durations varied considerably between methods. Even though consumers were given the same instructions to stop the recording, the durations measured with TCATA were 25% longer than those measured with TDS. Some similar differences between TDS and TCATA were previously reported by Meyners (2020). Other authors observed differences between durations measured with different temporal tasks (Rizo et al., 2019), or between durations measured with and without performing a sensory evaluation task (Devezeaux de Lavergne et al., 2015, Visalli et al., 2020). It is possible that due to the conceptual difference between tasks requiring a focus on either dominance or applicability, the notion of absence of perception has been

interpreted differently depending on the method. However, we believe that the difference between the concepts cannot completely explain such large differences. We therefore hypothesized that the differences observed between the methods could be more representative of the difficulty of the task than results of consumers self-assessments of this difficulty conducted in previous studies (Ares et al., 2015). Indeed, it was observed in other fields (Block et al., 2010) that the duration of the task often reflects the cognitive load associated with its complexity. Further studies would be needed to find the real explanation, but in any case, the measurements of total durations of perception are not accurate for at least one of the methods. This confirms that what is being measured with durations is unclear (Visalli, Béno, et al., 2023).

4.2. Heterogeneity in times of selection/unselection of attributes is a major source of imprecision

The results presented in section 3.1.2 showed that heterogeneity of consumer behaviour linked to temporal measurements is a major source of imprecision. This concerned more particularly the temporal events related to the end of perception (times of unselection of attributes and times to click on stop). Even after time standardization, the variability between consumers in times of selection/unselection of attributes remains non-negligible, of the order of 25% of the mean total duration of perception of the product. Converted into seconds, this represents an average difference comprised between 7 and 14 s, depending on the product category.

TCATA is slightly less imprecise regarding the measurements of times of selection, and TDS regarding the measurements of times of unselection. Selection of attributes might be more precise with TCATA because when evaluating applicability there is no need for a constant comparison between attributes to conclude whether it is dominant or not. In turn, regarding deselection, TDS borrows strength from the fact that deselection is triggered by the selection of a new attribute, which is probably an easier task compared to declaring something is no longer applicable.

Imprecision was higher with the most sensory complex product categories. A potential explanation could be that sensory complexity result in a greater number of concurrently perceived sensations, leading to delayed answers due to hesitation (Dietz, Cook, et al., 2022;

Varela et al., 2018) and thus to inconsistency (Meyners & Castura, 2019). However, it is also easier to obtain imprecise temporal measurements with products evaluated over longer durations, and in this study, the product categories perceived as more complex were also those perceived during longer durations. Imprecision also depends on the sensory modality of the attributes. For the Crisp product category, the list of attributes included textures that were probably easy to identify such as Crackly/Hard and Crunchy. This could have led to more consensual perceptions, as suggested by Lesme et al., (2020), and explain why the temporal profiles of crisps were more sequential than those of the other products.

In this study regular TCATA was used, and results could have been different if the fading variant has been used (Ares et al., 2016). Fading TCATA implicitly acknowledges the difficulty of unselecting attributes and measuring the durations of applicability. In this sense, the method suggests the existence of a temporal resolution established at the level of the fading duration. Along the same lines, Derks et al. (2022) discretized TCATA data using buckets whose size varied between 2 and 5 s, based on the expected delay between the sensory stimulus and the panellist's response. They showed that processing TCATA data in this way increased the discrimination ability of the method.

4.3. Agreement and reliability are poor with temporal measurements

Imprecision concerned both measurements within consumers (non-repeatability) and between consumers (disagreement). In absolute value, the repeatability was not much higher than the agreement (see Section 3.1.3). It means that for a given product, there is almost as much differences between two sequences collected by the same consumer than by different consumers. Roughly, for a given product and at any given time, the consumers selected the same attribute 20 and 30% of the times in TDS and TCATA, respectively. The difference between TDS and TCATA may be explained by a higher probability of intersection of two random sequences of n attributes ($n > 1$) compared to the probability of intersection of two random sequences of one attribute.

Considering agreement and reliability at each discretized time might be too demanding because it may give too much importance to slightly delayed perception. When repeatability was assessed without taking into account the temporal dimension of the data (citations only), it was shown that consumers selected the same attributes about 50% of the times. This level

of repeatability becomes comparable to the lowest level of repeatability obtained with a static CATA (Ares, Antúnez, et al., 2014). These results confirm that it is important to have a large number of consumers to limit the effect of imprecision and address the heterogeneous nature of temporal sensory perception, as advised by Castura (2018). As mentioned in introduction, consumers are not expected to be individually repeatable. However, adding at least one replicated sample can help interpret the size of the differences between products, with larger differences expected between different products than between replicated samples.

4.4. The imprecision of temporal measurements can make TDS/TCATA curves and difference curves difficult to interpret

TDS and TCATA ask consumers to report events related the presence or absence of dominant/applicable attributes. For the purpose of plotting TDS and TCATA curves and difference curves, these events are transformed into discrete time series. Then, statistical tests are independently performed for each attribute at each discrete time. This process ignores temporal dependence of time series data and the variability in times of selection/unselection of attributes. Thus, two consumers having perceived an attribute between [times t and $t+10$] or between [times $t+10$ and $t+30$] contribute in the same way to the calculation of the citation rate at time t . This can make the interpretation of the TDS and TCATA curves tricky. When times of selection/unselection of attributes are not reasonably aligned between consumers, attribute curves may be flattened. In this case, it is not possible to determine whether these attributes were perceived by a majority of consumers during short periods but at different moments or by a few consumers for long periods at the same time. Time standardization increased the precision of the measurements of times of selection/unselection of attributes. However, as there is more variability in times to click on stop than in times to click on first attributes, time standardization stretches the sequences more to the right than to the left. This can be an issue as this gives more importance to the attributes cited at first and last rank.

As presented in section 3.3, the differences in citation rates highlighted by the TDS/TCATA difference curves are not necessarily linked to differences in times of selection/unselection of attributes. Indeed, the meaning of a difference in citation rates between two products A and B of about 15% (the threshold for significance of differences observed in Section 3.3) is unclear. Such a difference can be due to an attribute perceived by a few consumers in product

A and not perceived in product B, to an attribute perceived just below the significance line in A and largely above in B (in TDS), or to an attribute perceived by 100% of consumers in A and 85% in B. In these cases, these differences in agreement related to the dominance or applicability of an attribute at a given time are not very informative because they do not necessarily enable to conclude on quantitative changes in the perception of the products. Indeed, several studies have shown that citation rates are not necessarily related to intensities (Kawasaki et al., 2016; Kiumarsi et al., 2021; Pierguidi et al., 2021; Visalli, Béno, et al., 2023).

4.5. Time-variability plots allow testing and visualization of temporal differences within and between products

Bootstrap has already been used by other authors to analyse temporal sensory data based on citation rates (Okamoto, 2021). Here, bootstrap made it possible to test differences in times of selection/unselection of attributes. Surprisingly, this is the first article that compares times of selection/unselection of attributes without transforming the data into time series data.

From a statistical perspective, with the bootstrap approach: (i) no parametric assumption is required; (ii) the number of statistical tests is drastically reduced compared to TDS/TCATA curves and difference curves, limiting the inflation of the alpha risk due to the multiplication of tests (Meyners et al., 2010, Meyners et al., 2019) ; (iii) confidence intervals are included for the times of selection/unselection and for citation rates for each attribute; (iv) within-product tests do not depend on an additional reference such as the average citation proportion over all products with TCATA or the significance line with TDS; (v) within-product tests make it possible to statistically conclude on the temporality of perception without relying on subjective visual interpretation of curves; (vi) between-product tests take into account the imprecision of temporal measurements, which result in more conservative tests.

The proposed framework of analysis presents several practical advantages: (i) it can be applied with TDS and TCATA data; (ii) it can be applied with or without time standardization (but time-variability plots reduce the need to align begin and end of perception between consumers because medians are not sensitive to outlier durations for these two variables); (iii) it makes it possible to draw time-variability plots that are consistent with the nature of TDS and TCATA data and interpretable on the temporal aspects of perception; (iv) multiple citations of the same attributes by the same consumer can be distinguished. However, time-variability plots

do not make it possible to monitor the evolution of the citation rates over time, and they can be too sensitive to temporal differences occurring with attributes few cited. These representations do not replace but complement TDS/TCATA curves and difference curves.

It is to be noted that this approach is based on comparison of times of selection/unselection computed at the panel level. It does not directly evaluate transitions between attributes as in other approaches like dyads (Castura & Li, 2016), graphs (Castura, 2020), semi-markov chains (Lecuelle et al., 2018) or categorical functional data analysis (Peltier et al., 2023).

A selection criterion was proposed to determine the relevant attributes, that is to say the attributes which have been cited sufficiently not to be considered as noise. The difference between the upper and lower confidence level limits of intervals of median citation rates was used to estimate the minimum level a citation rate should reach to be considered “large sufficient” compared to the non-reproducibility. This process enabled the selection of the most consensual attributes, whether or not these attributes present specific temporal patterns. In this study, depending on the product category and method, this minimum level ranges between 13% and 24%, with an average of approximately 20%. This average level is comparable to the selection threshold based on Pareto principle (20%) used with studies involving Just-About-Right measurements (Ares et al., 2014; Plaehn, 2012; Xiong & Meullenet, 2006). The threshold can be set by the experimenter instead of being determined automatically. Higher values would lead to the selection of attributes that would be few in number but perceived in a more consensual manner, and therefore to more conservative conclusions.

4.6. The temporal resolution of the methods can be considered as the minimum size of differences to be interpreted

The results presented in Table 59 suggests that roughly, the temporal resolution of the method is in average about 10% of the total duration of perception. Although this value remains to be confirmed with other types of products and panels, the findings of this study have several practical implications. As a general rule, to draw robust conclusions when the temporal resolution is unknown, we suggest using this threshold of 10% of the total duration of perception of the product as the minimum duration for a temporal signal to be considered for interpretation. Shorter duration differences should be interpreted with caution as they

may be due to imprecision in temporal measurements. These differences may concern: (i) the minimum time interval on the x-axis between the TDS/TCATA curves of two attributes to consider that one attribute is perceived before/after the other; (ii) the minimum duration of successive significant differences in citation rates on TDS/TCATA difference curves to be considered as meaningful; (iii) the minimum duration above significance for an attribute to be considered as dominant on TDS curves.

It seems that the temporal resolution of TDS and TCATA is better (i.e. smaller temporal differences are captured) for times of selection compared to times of unselection, which could be explained by the large heterogeneity observed in measured durations of perception of the attributes. This result echoes the findings on durations in Section 4.2, and concurs with findings of Visalli, Béno, et al. (2023). This suggests that TDS and TCATA are more effective at capturing times when attributes are perceived as dominant/applicable than at capturing durations. In this study, the two methods were overall in agreement regarding the times of selection of the attributes.

The temporal resolution of TCATA appeared to be worse than that of TDS. However, as pointed out in section 4.2, the measured durations are longer with TCATA. When the resolution in seconds is converted to percentages of the total durations, the difference of temporal resolution between the two methods no longer exists. TDS highlighted the more within-product temporal differences. TCATA highlighted fewer but more consensual differences in citation times. This result shows the conceptual difference between dominance (more changing over time) and applicability (more stable over time). TCATA discriminates the end times of perception of the attributes less, which agrees with McMahon et al. (2017), who suggested a relative grouping of attributes (earlier vs. delayed perception). This result can be explained by the difficulty for consumers to uncheck the applicable attributes, as previously observed (Reyes et al., 2017; Vidal et al., 2017).

4.6. Limitations of this study

Several limitations should be reported. First, the specific design order of the study (fixed product order) limits the generalizability of conclusions related to the repeatability of temporal measurements. Indeed, certain differences observed between replicated samples could also be due to a product order effect. However, if the TDS and TCATA measurements

were affected by this potential bias, they were affected in the same way. Second, the context of the evaluations was different for guacamoles and ice teas (in the laboratory) compared to crisps and chocolates (at home). Comparing the two contexts was beyond the scope of this study. Internet connection speed was not an issue for recording times, since sessions were downloaded to consumers' devices before being run. Third, it is possible that some consumers used touchscreen devices (or devices with different screen sizes) at home when they had used a mouse in the laboratory. This could have affected the measurements of times of selection/unselection of attributes which could be shorter with touchscreen devices (Visalli et al., 2016). Potential previous experience of the consumers with the temporal methods could also have affected the performances. As “device” and “experience” factors were not controlled across panels, this could limit the generalizability of the conclusions related to the difference between the methods. Fourth, the absence of definitions for the attributes of the products evaluated at home could have decrease the performances of the consumers. Finally, conclusions of this study cannot be extended to data collected with trained panels, for which temporal precision and resolution remain to investigate.

5. Conclusion

TDS and TCATA are qualitative temporal sensory evaluation methods now considered as rapid methods usable by consumers. Both methods have been proven effective in capturing key perceptual changes. These methods are sometimes used to discriminate between products based on “subtle” temporal differences. However, it should not be forgotten that temporal perception is a complex phenomenon, and it should not be expected more from the methods than they can offer. Heterogeneity in temporal perception of consumers can be a problem when TDS/TCATA or difference curves are used as confirmatory data analyses rather than descriptive statistics. Indeed, the higher the imprecision in temporal measurements, the weaker the method is at detecting significant short-duration temporal signals and differences within or between stimuli. Thus, to draw robust conclusions related to temporal aspects of perception, it is recommended to consider that a minimum time is needed between two temporal events (i.e. the begin/end of perception of sensory attributes) to conclude they are significantly discriminated (i.e. advanced or delayed). As a general rule, it is suggested to consider that within and between-product temporal differences lasting for less than 10% of the total duration of perception of the product are uncertain and can be spurious. Finally,

beyond statistical aspects, the interpretation of a short duration signal must be done taking into account its practical interest in relation to the objective of the temporal sensory measurement (e.g. product formulation or fundamental knowledge on perception).

Supplementary material

Annex 1 – Results from ESN survey

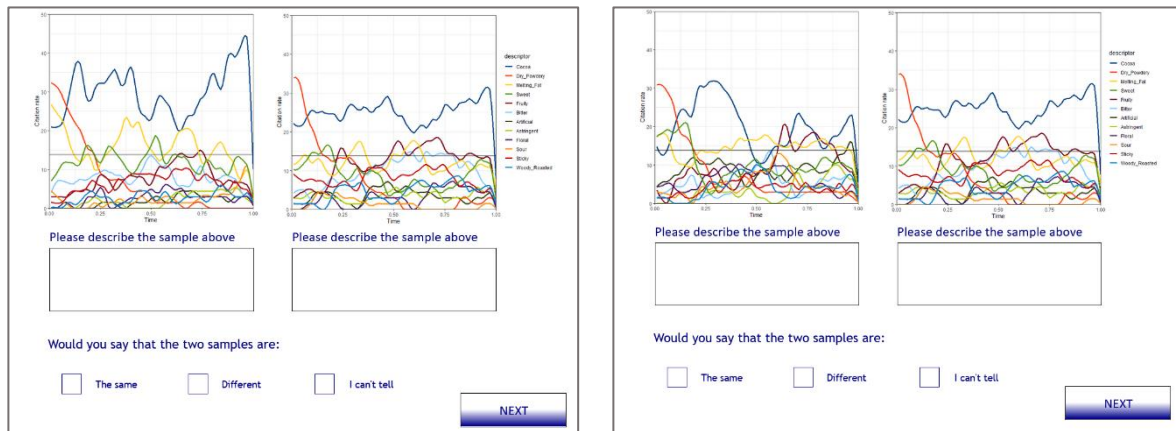


Figure 62. Screenshots of the survey administered to members of the European Sensory Network, asking them to compare the chocolates A and B, and C to B (C being a replication of B).

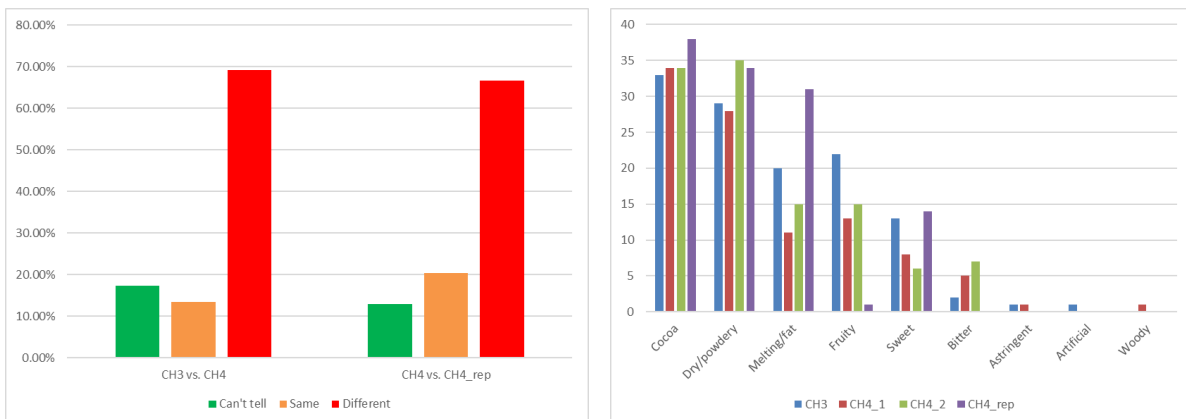
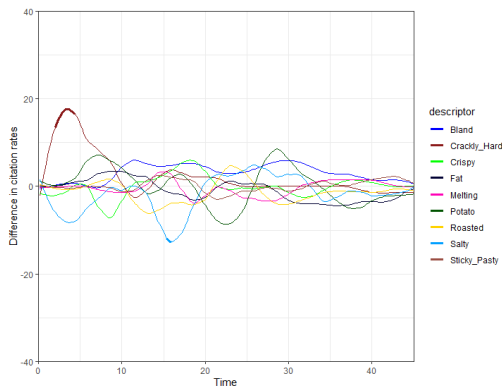


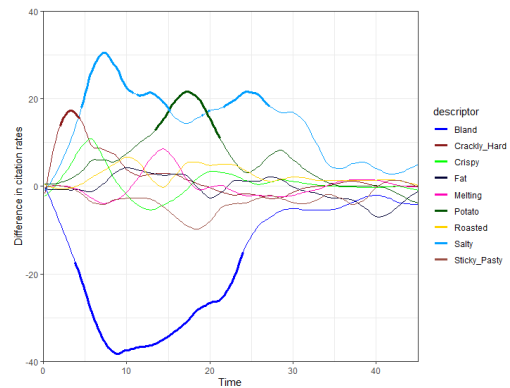
Figure 63. Results of the ESN survey. Left: Percentage of answers for the comparison of samples. Right: Descriptors used to describe the sample. Seven participants interpreted the citation rates as intensities, 6 concluded that there was no temporality in the TDS signal, and 1 thought the curves were drawn randomly.

Annex 2 – Additional results for Crisps

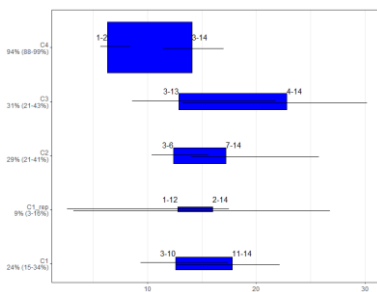
A - TDS difference curves of C1-C1_rep without time standardization



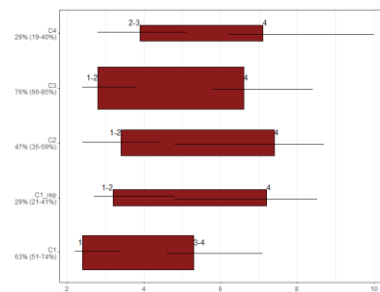
B - TDS difference curves of C1-C4 without time standardization



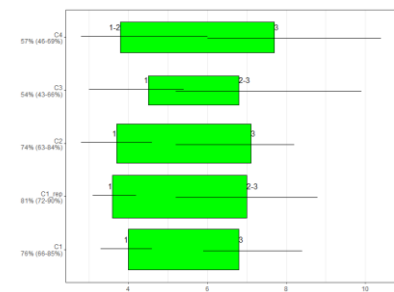
C - TDS time-variability plot of Bland



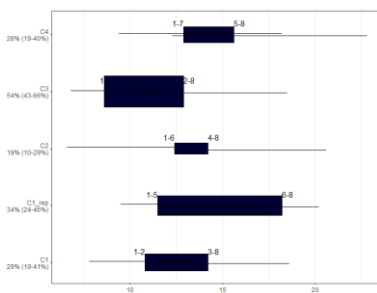
D - TDS time-variability plot of Crackly_Hard



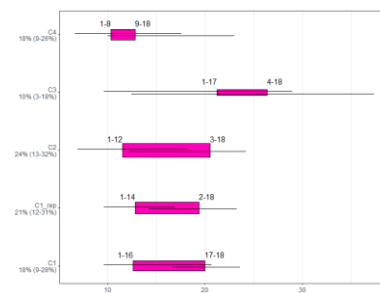
E - TDS time-variability plot of Crispy



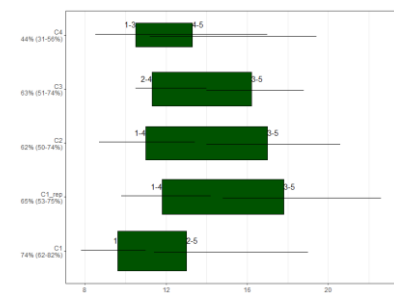
F - TDS time-variability plot of Fat



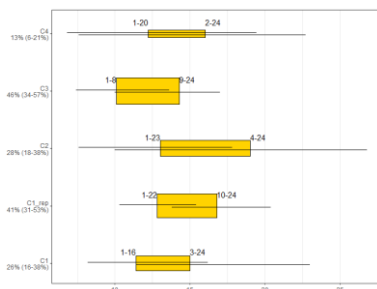
G - TDS time-variability plot of Melting



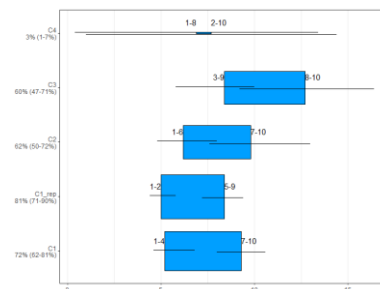
H - TDS time-variability plot of Potato



I - TDS time-variability plot of Roasted



J - TDS time-variability plot of Salty



K - TDS time-variability plot of Sticky_Pasty

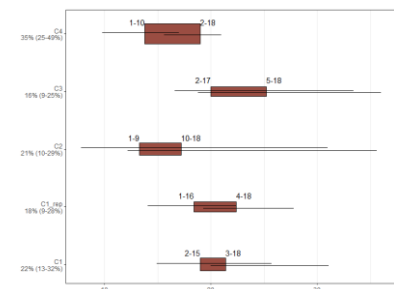
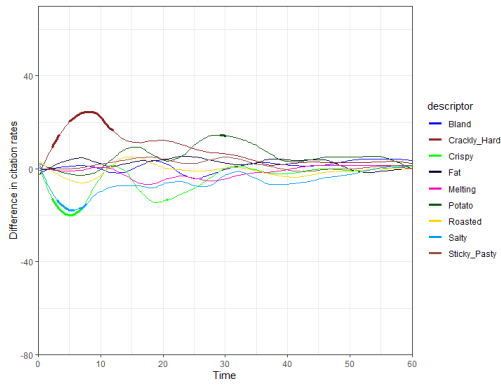
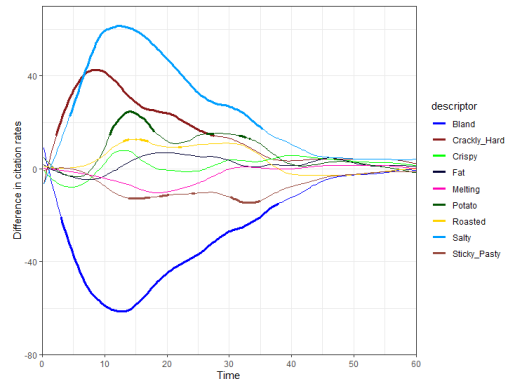


Figure 64. Representation of differences between C1, C1_rep and C4, by attribute (TDS, without time-standardization).

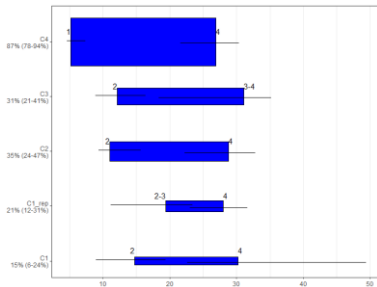
A - TCATA difference curves of C1-C1_rep without time standardization



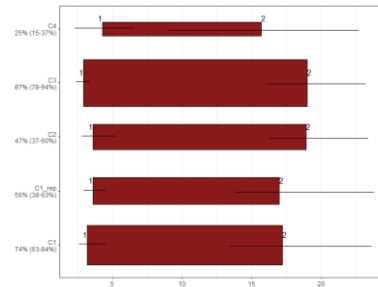
B - TCATA difference curves of C1-C4 without time standardization



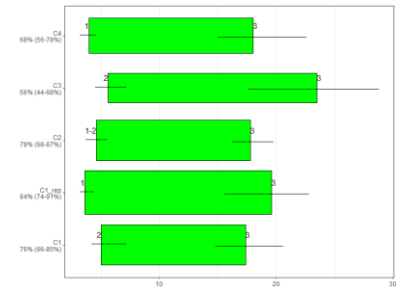
C - TCATA time-variability plot of Bland



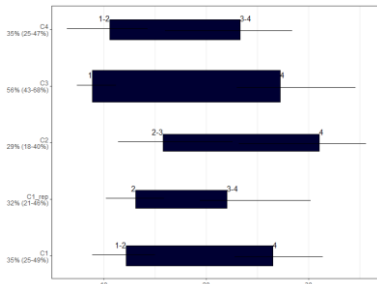
D - TCATA time-variability plot of Crackly_Hard



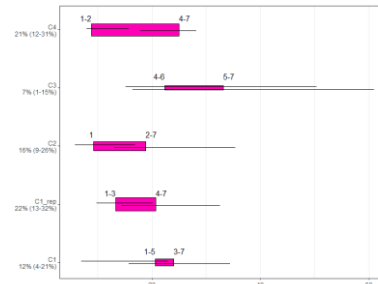
E - TCATA time-variability plot of Crispy



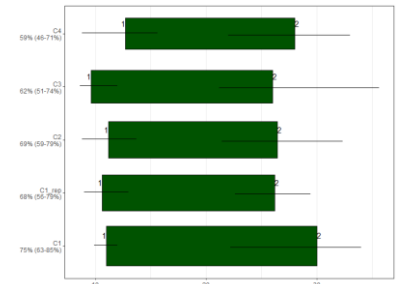
F - TCATA time-variability plot of Fat



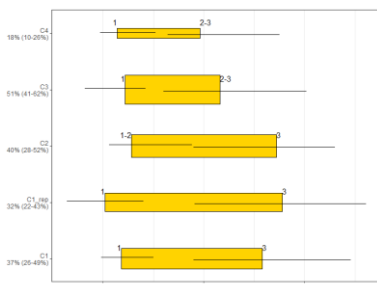
G - TCATA time-variability plot of Melting



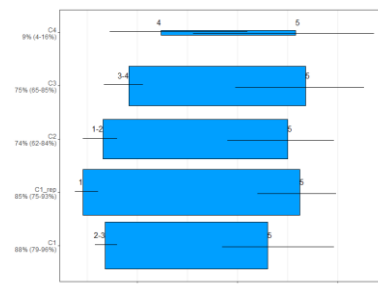
H - TCATA time-variability plot of Potato



I - TCATA time-variability plot of Roasted



J - TCATA time-variability plot of Salty



K - TCATA time-variability plot of Sticky_Pasty

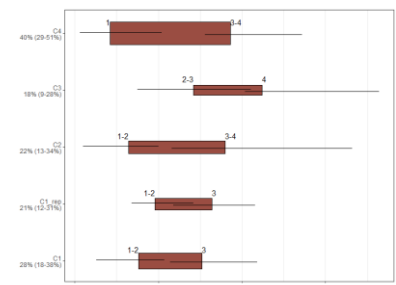
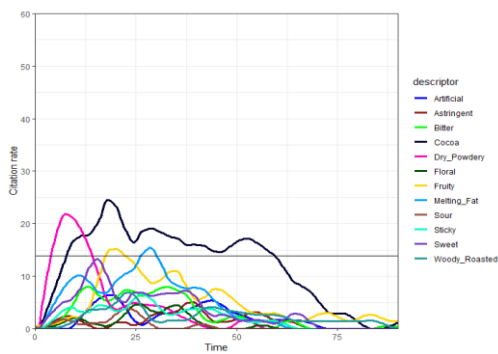


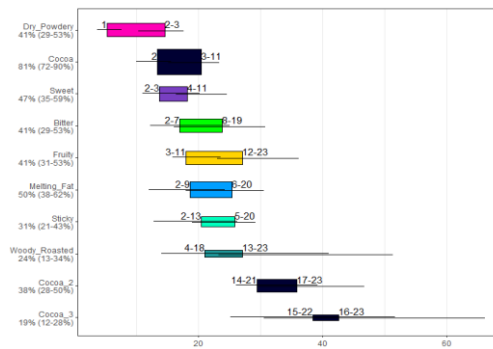
Figure 65. Representation of differences between C1, C1_rep and C4, by attribute (TCATA, without time-standardization).

Annex 3 – Additional results for Chocolates

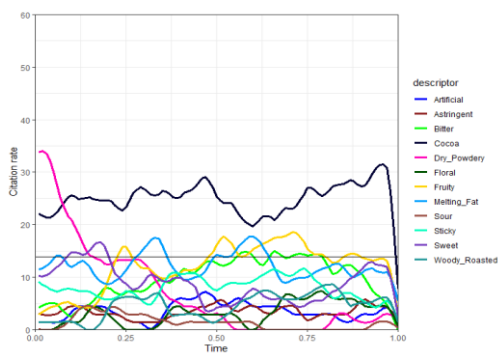
A - TDS curves of CH4 without time standardization



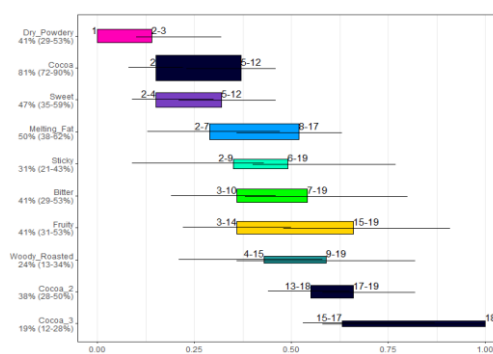
B - TDS time-variability plot of CH4 without time standardization



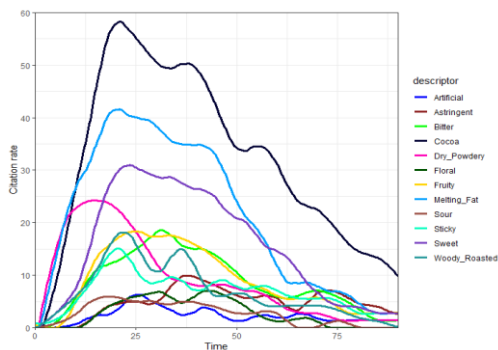
C - TDS curves of CH4 with time standardization



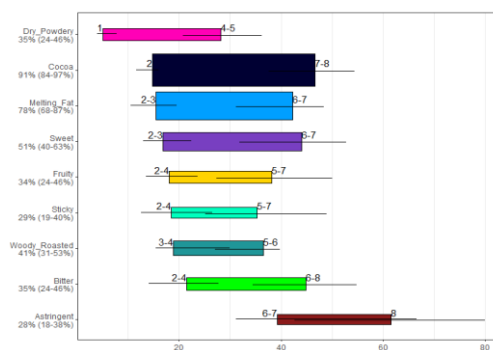
D - TDS time-variability plot of CH4 with time standardization



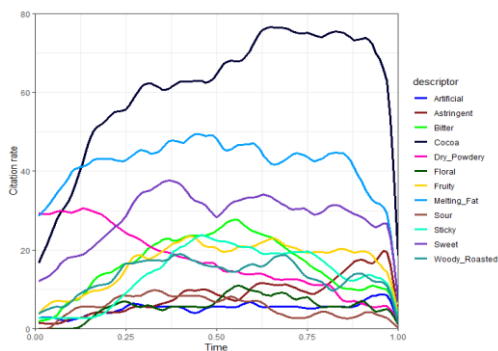
E - TCATA curves of CH4 without time standardization



F - TCATA time-variability plot of CH4 without time standardization



H - TCATA curves of CH4 with time standardization



H - TCATA time-variability plot of CH4 with time standardization

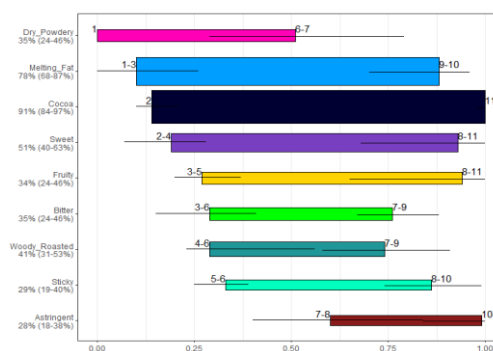
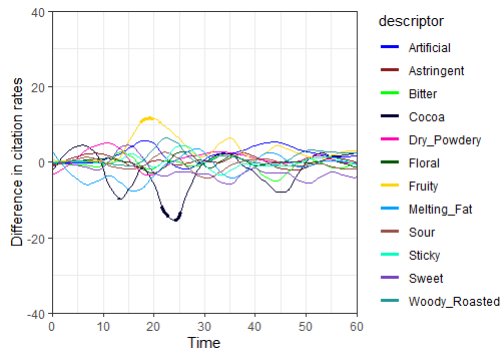
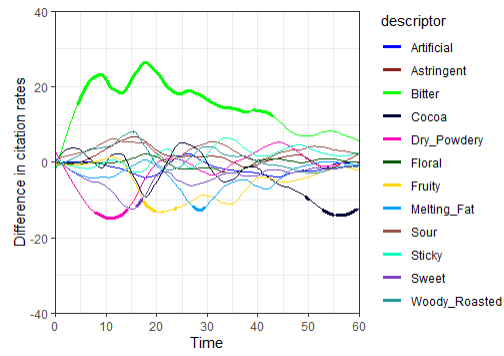


Figure 66. Representation of within-product differences between attributes for Chocolate CH4.

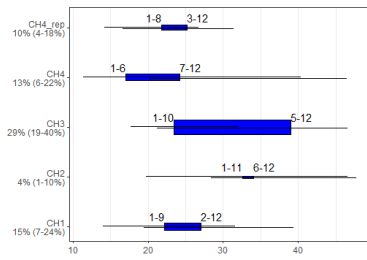
A - TDS difference curves of CH4-CH4_rep without time standardization



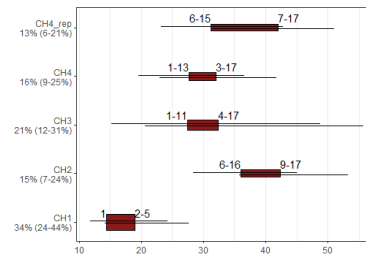
B - TDS difference curves of CH1-CH4 without time standardization



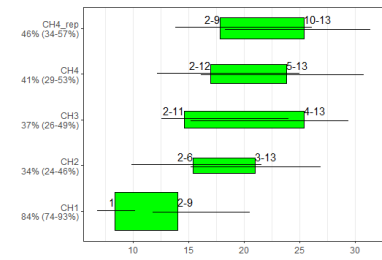
C - TDS time-variability plot of Artificial



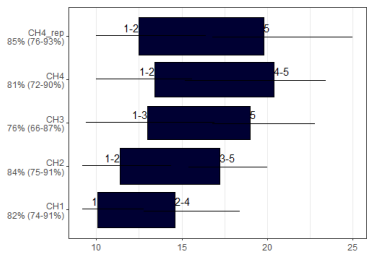
D - TDS time-variability plot of Astringent



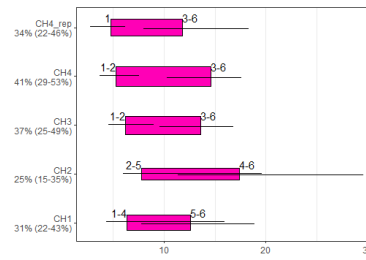
E - TDS time-variability plot of Bitter



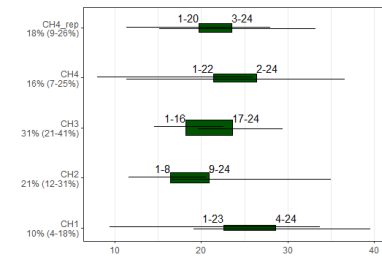
F - TDS time-variability plot of Cocoa



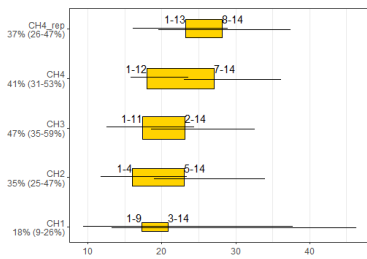
G - TDS time-variability plot of Dry/Powdery



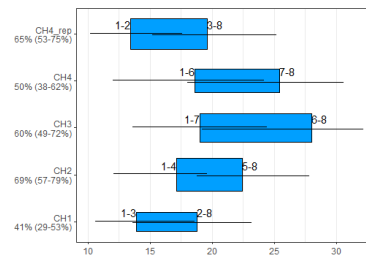
H - TDS time-variability plot of Floral



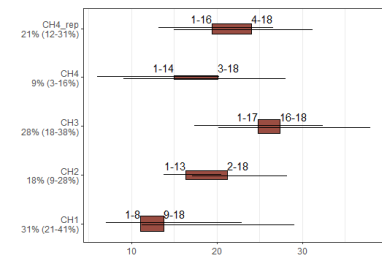
I - TDS time-variability plot of Fruity



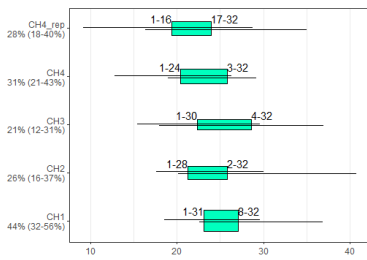
J - TDS time-variability plot of Melting



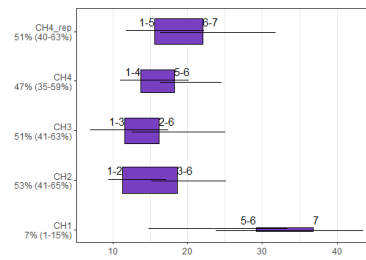
K - TDS time-variability plot of Sour



L - TDS time-variability plot of Sticky



M - TDS time-variability plot of Sweet



N - TDS time-variability plot of Woody/Roasted

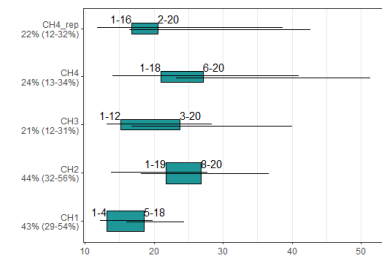
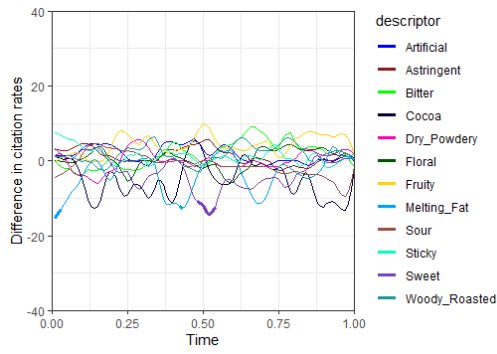
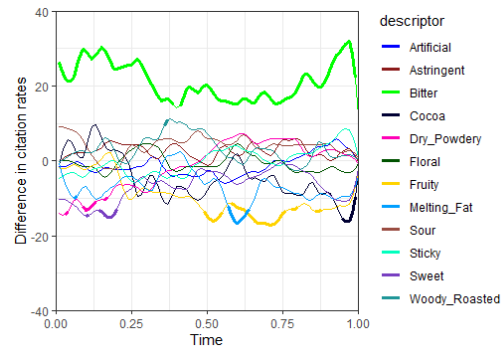


Figure 67. Representation of differences between CH1, CH4 and CH4_rep, by attribute (TDS, without time-standardization).

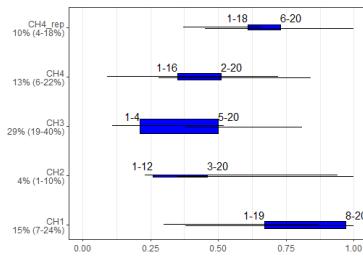
A - TDS difference curves of CH4-CH4_rep with time standardization



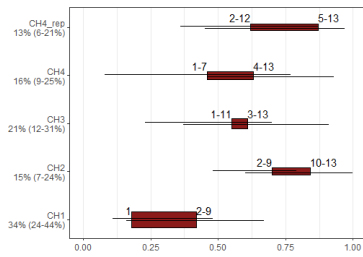
B - TDS difference curves of CH1-CH4 with time standardization



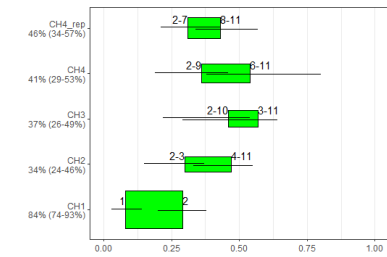
C - TDS time-variability plot of Artificial



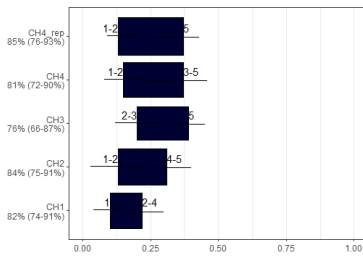
D - TDS time-variability plot of Astringent



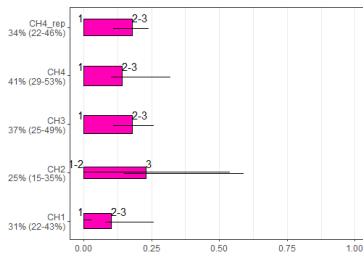
E - TDS time-variability plot of Bitter



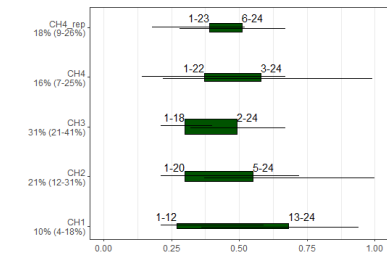
F - TDS time-variability plot of Cocoa



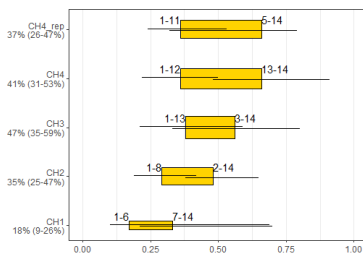
G - TDS time-variability plot of Dry/Powdery



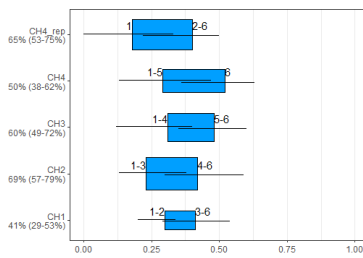
H - TDS time-variability plot of Floral



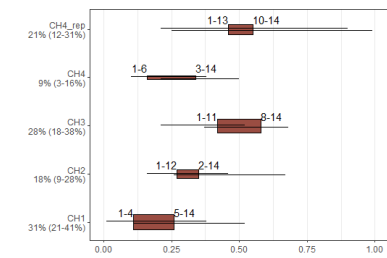
I - TDS time-variability plot of Fruity



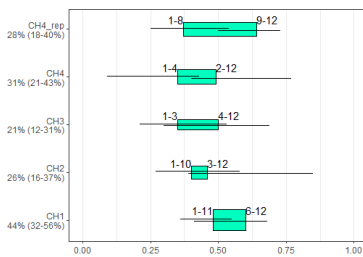
J - TDS time-variability plot of Melting



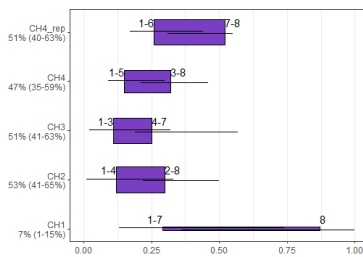
K - TDS time-variability plot of Sour



L - TDS time-variability plot of Sticky



M - TDS time-variability plot of Sweet



N - TDS time-variability plot of Woody/Roasted

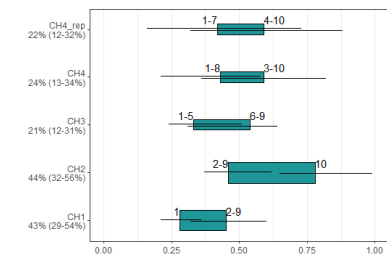
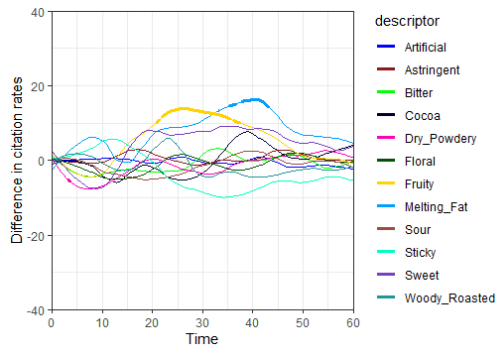
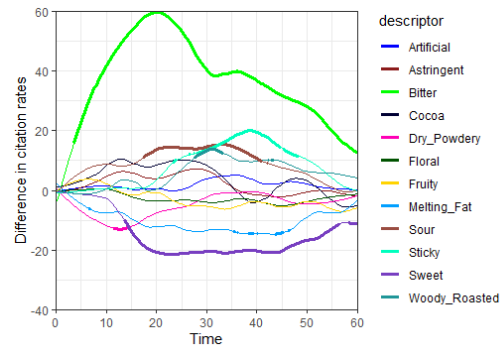


Figure 68. Representation of differences between CH1, CH4 and CH4_rep, by attribute (TDS, with time-standardization).

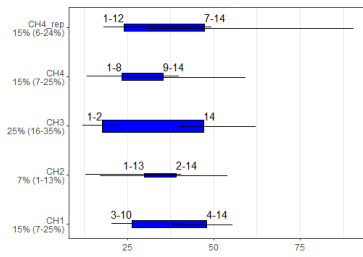
A - TCATA difference curves of CH4-CH4_rep without time standardization



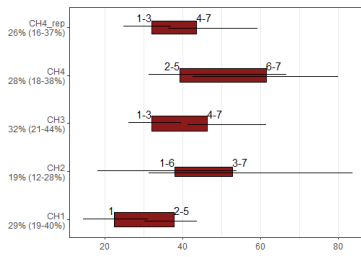
B - TCATA difference curves of CH1-CH4 without time standardization



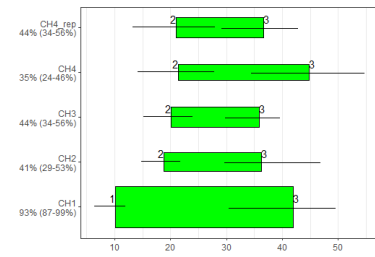
C - TCATA time-variability plot of Artificial



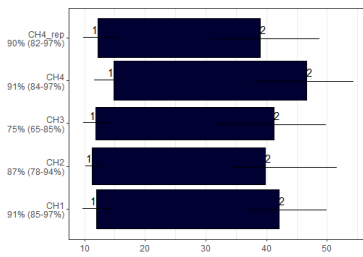
D - TCATA time-variability plot of Astringent



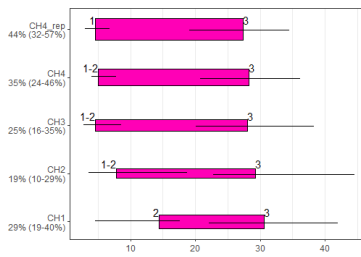
E - TCATA time-variability plot of Bitter



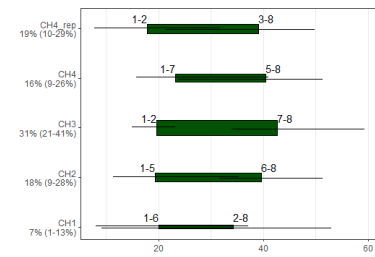
F - TCATA time-variability plot of Cocoa



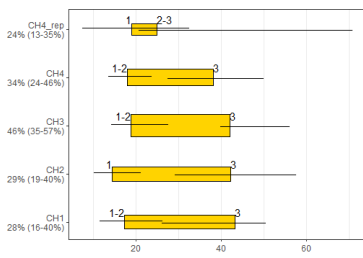
G - TCATA time-variability plot of Dry/Powdery



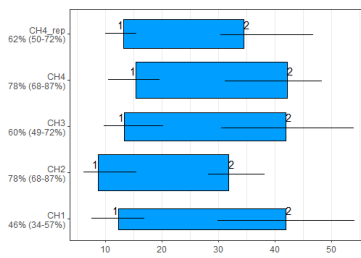
H - TCATA time-variability plot of Floral



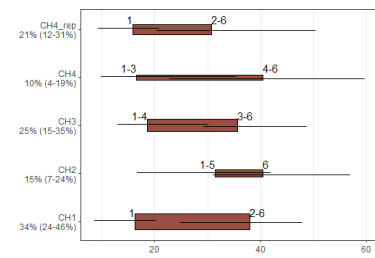
I - TCATA time-variability plot of Fruity



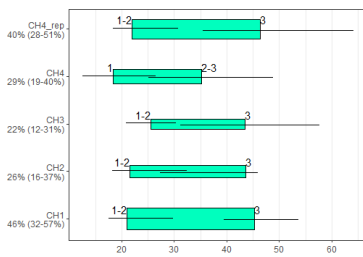
J - TCATA time-variability plot of Melting



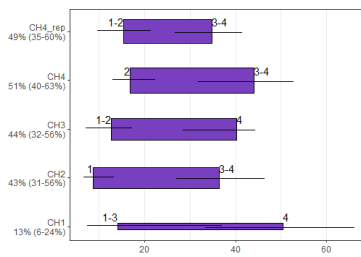
K - TCATA time-variability plot of Sour



L - TCATA time-variability plot of Sticky



M - TCATA time-variability plot of Sweet



N - TCATA time-variability plot of Woody/Roasted

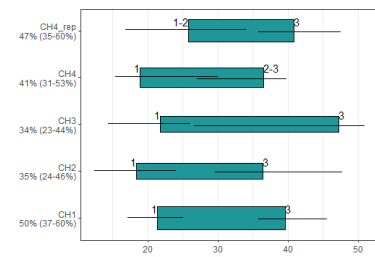
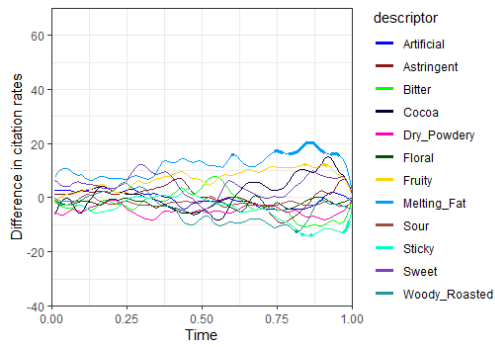
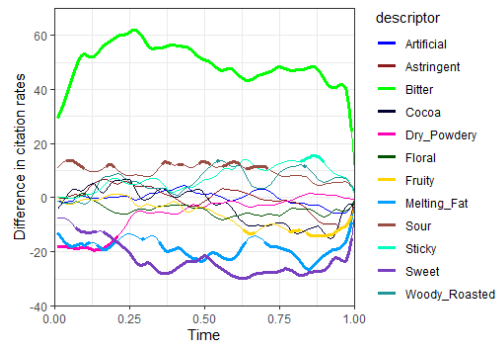


Figure 69. Representation of differences between CH1, CH4 and CH4_rep, by attribute (TCATA, without time-standardization).

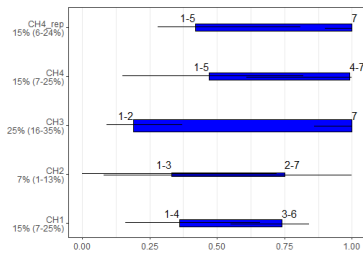
A - TCATA difference curves of CH4-CH4_rep with time standardization



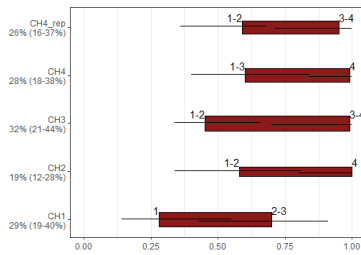
B - TCATA difference curves of CH1-CH4 with time standardization



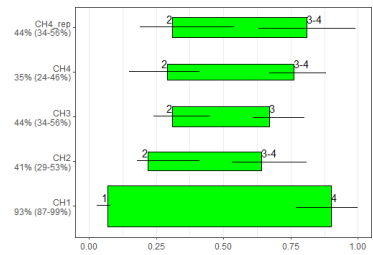
C - TCATA time-variability plot of Artificial



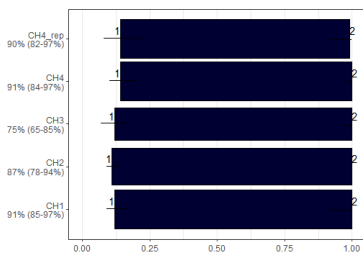
D - TCATA time-variability plot of Astringent



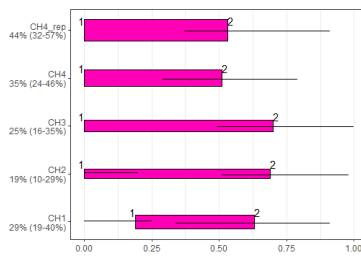
E - TCATA time-variability plot of Bitter



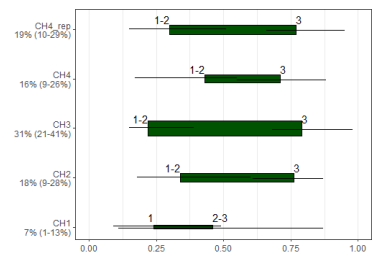
F - TCATA time-variability plot of Cocoa



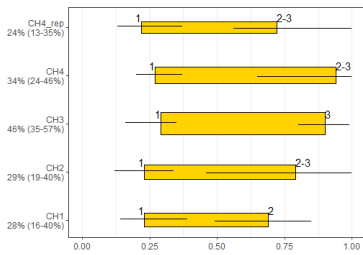
G - TCATA time-variability plot of Dry/Powdery



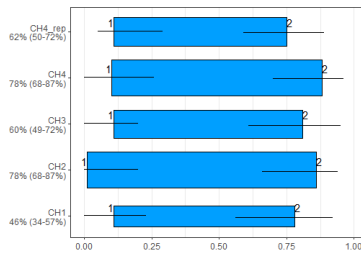
H - TCATA time-variability plot of Floral



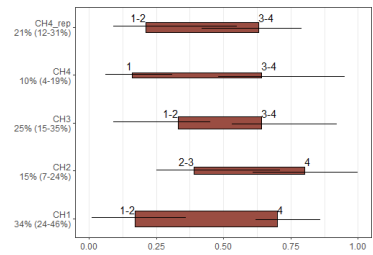
I - TCATA time-variability plot of Fruity



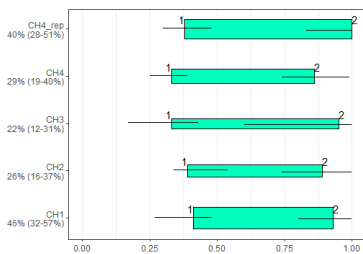
J - TCATA time-variability plot of Melting



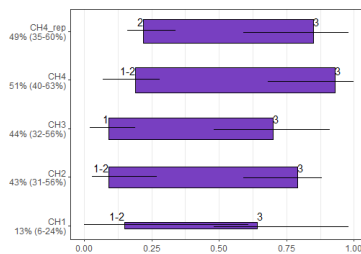
K - TCATA time-variability plot of Sour



L - TCATA time-variability plot of Sticky



M - TCATA time-variability plot of Sweet



N - TCATA time-variability plot of Woody/Roasted

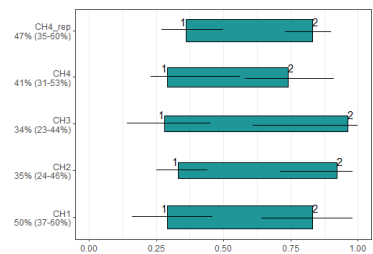
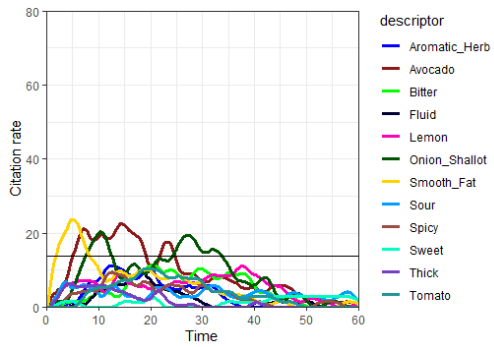


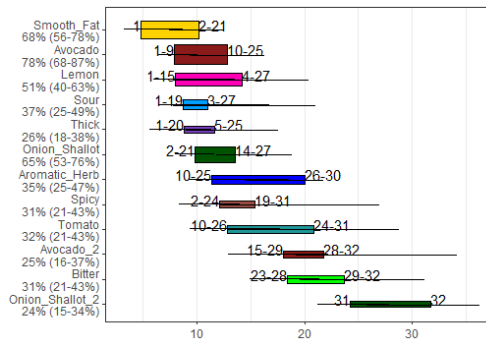
Figure 70. Representation of differences between CH1, CH4 and CH4_rep, by attribute (TCATA, with time-standardization).

Annex 4 – Additional results for Guacamoles

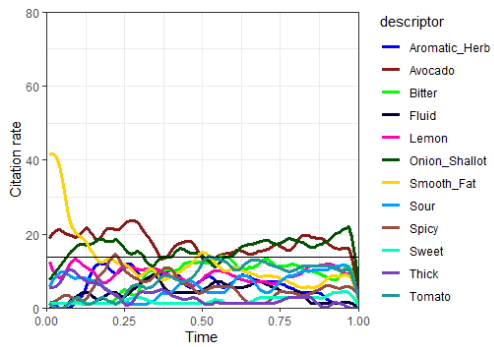
A - TDS curves of G1 without time standardization



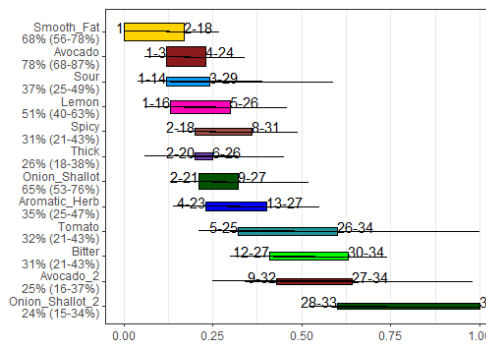
B - TDS time-variability plot of G1 without time standardization



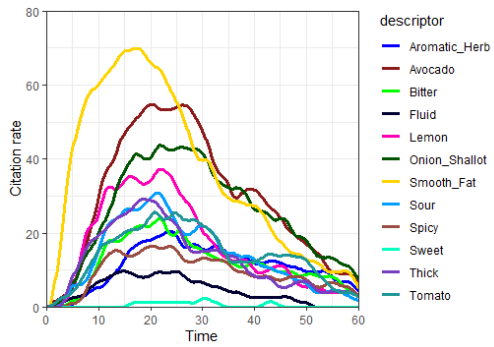
C - TDS curves of G1 with time standardization



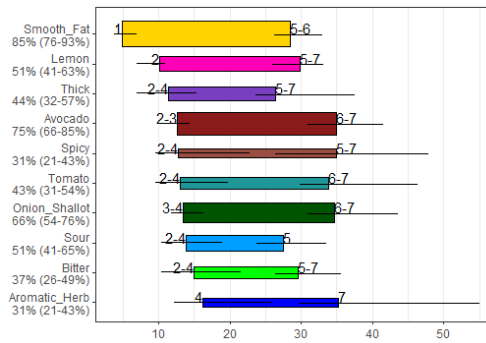
D - TDS time-variability plot of G1 with time standardization



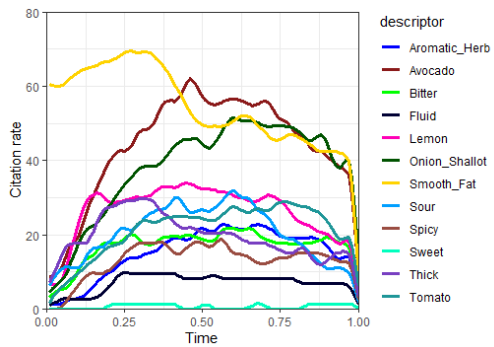
E - TCATA curves of G1 without time standardization



F - TCATA time-variability plot of G1 without time standardization



G - TCATA curves of G1 with time standardization



H - TCATA time-variability plot of G1 with time standardization

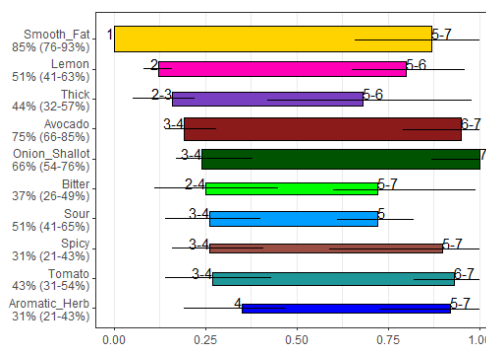
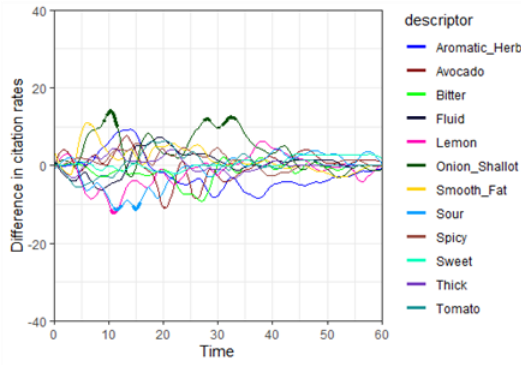
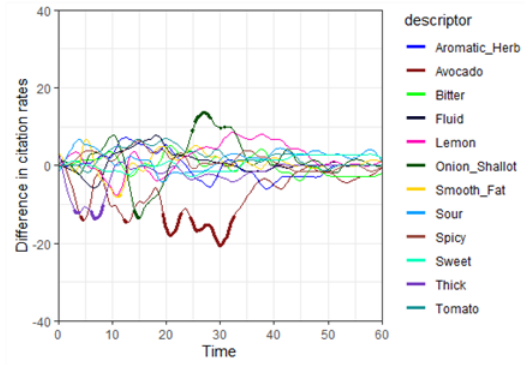


Figure 71. Representation of within-product differences between attributes for Guacamole G1.

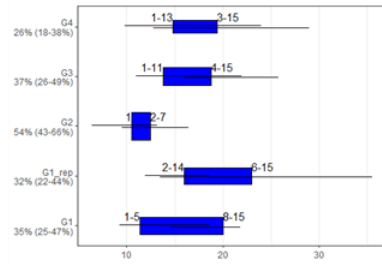
A - TDS difference curves of G1-G1_rep without time standardization



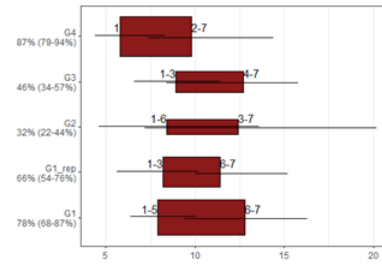
B - TDS difference curves of G1-G4 without time standardization



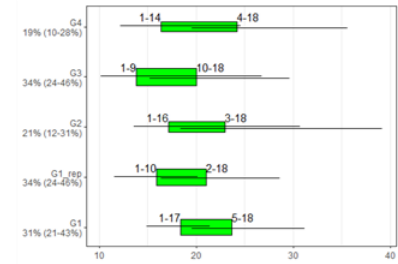
C - TDS time-variability plot of Aromatic herbs



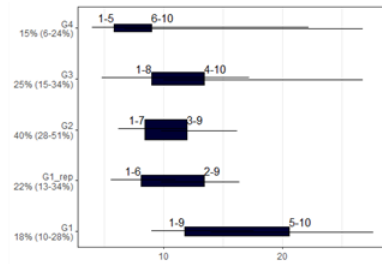
D - TDS time-variability plot of Avocado



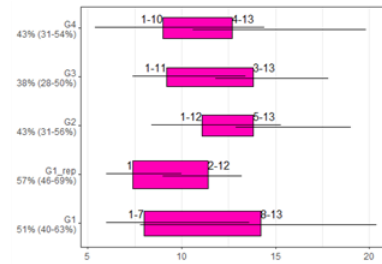
E - TDS time-variability plot of Bitter



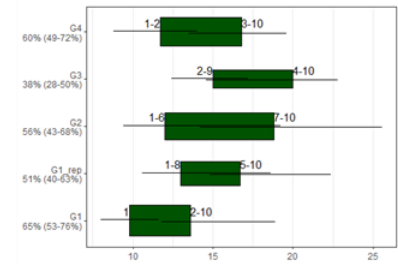
F - TDS time-variability plot of Fluid



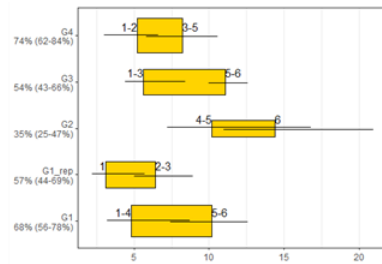
G - TDS time-variability plot of Lemon



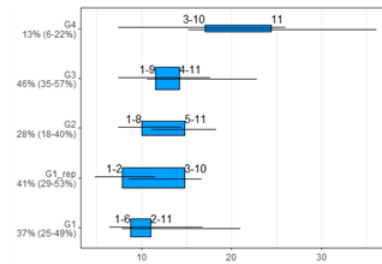
H - TDS time-variability plot of Onion/Shallot



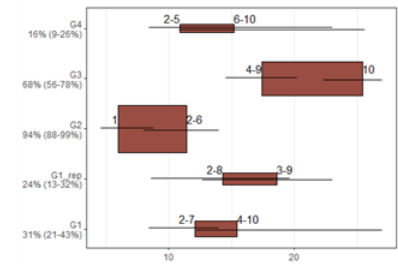
I - TDS time-variability plot of Smooth/Fat



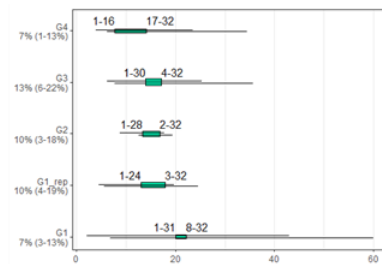
J - TDS time-variability plot of Sour



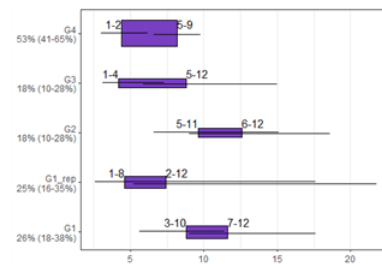
K - TDS time-variability plot of Spicy



L - TDS time-variability plot of Sweet



M - TDS time-variability plot of Thick



N - TDS time-variability plot of Tomato

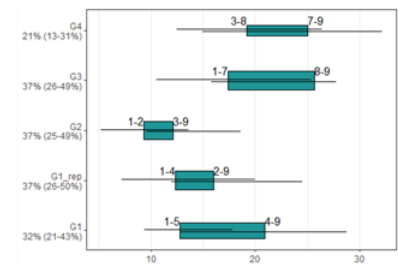
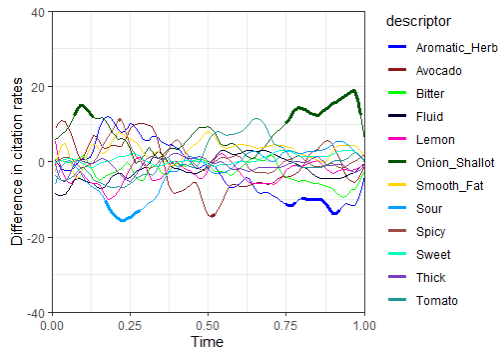
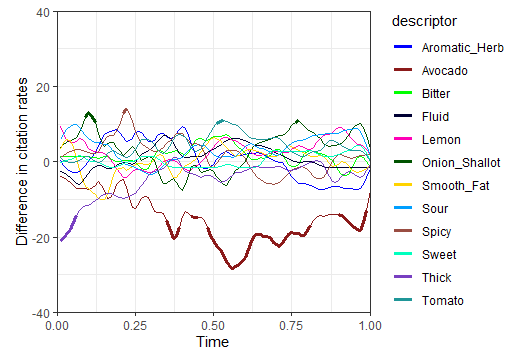


Figure 72. Representation of differences between G1, G1_rep and G4, by attribute (TDS, without time-standardization).

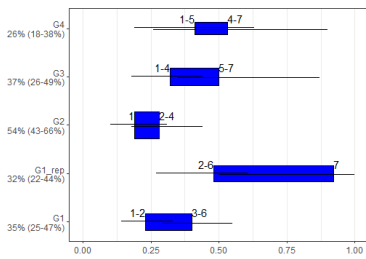
A - TDS difference curves of G1-G1_rep with time standardization



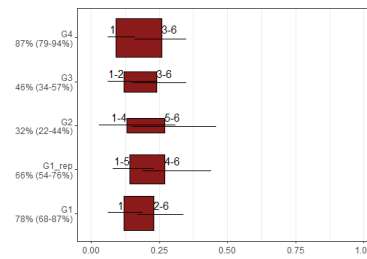
B - TDS difference curves of G1-G4 with time standardization



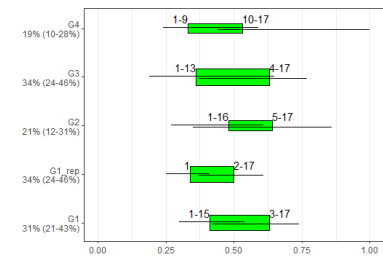
C - TDS time-variability plot of Aromatic herbs



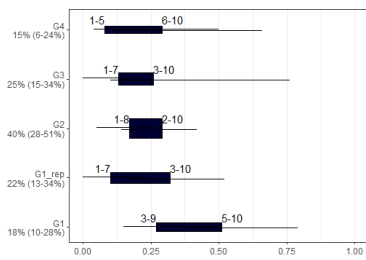
D - TDS time-variability plot of Avocado



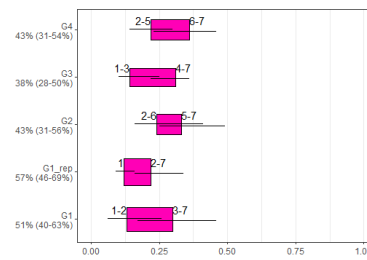
E - TDS time-variability plot of Bitter



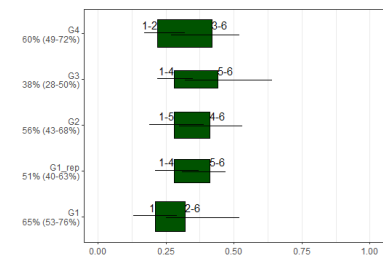
F - TDS time-variability plot of Fluid



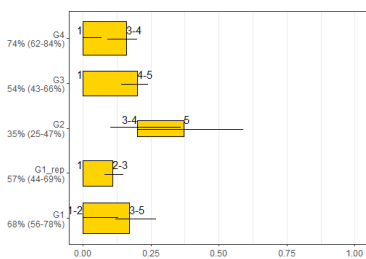
G - TDS time-variability plot of Lemon



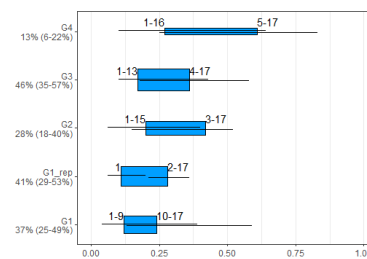
H - TDS time-variability plot of Onion/Shallot



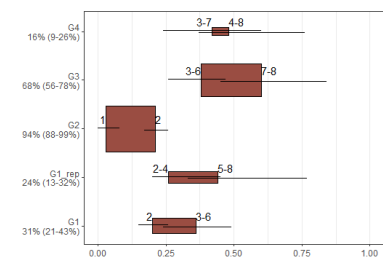
I - TDS time-variability plot of Smooth/Fat



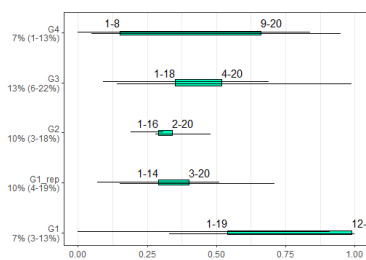
J - TDS time-variability plot of Sour



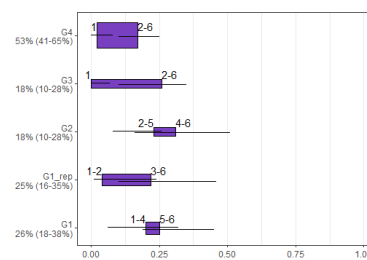
K - TDS time-variability plot of Spicy



L - TDS time-variability plot of Sweet



M - TDS time-variability plot of Thick



N - TDS time-variability plot of Tomato

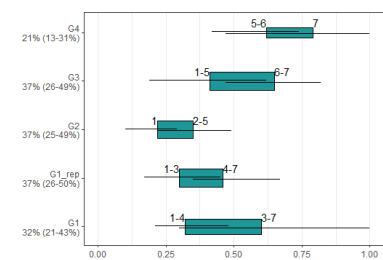
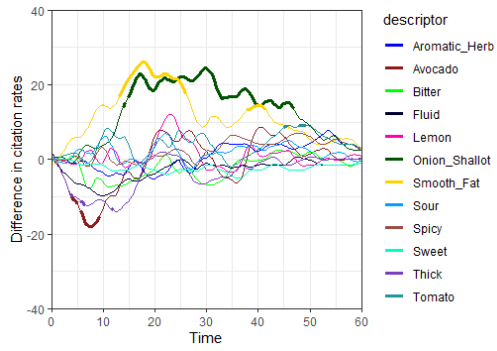
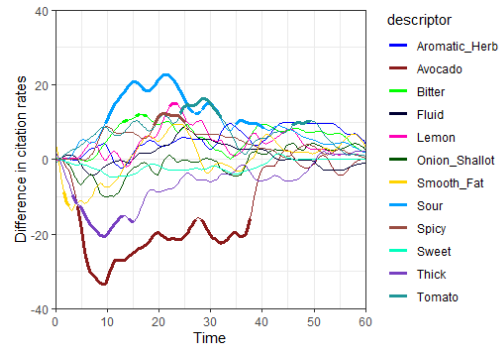


Figure 73. Representation of differences between G1, G1_rep and G4, by attribute (TDS, with time-standardization).

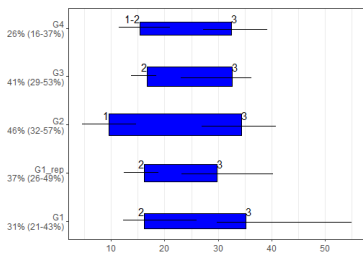
A - TCATA difference curves of G1-G1_rep without time standardization



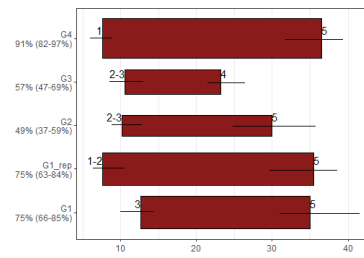
B - TCATA difference curves of G1-G4 without time standardization



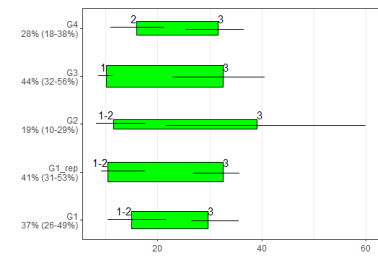
C - TCATA time-variability plot of Aromatic herbs



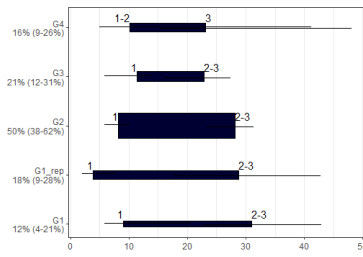
D - TCATA time-variability plot of Avocado



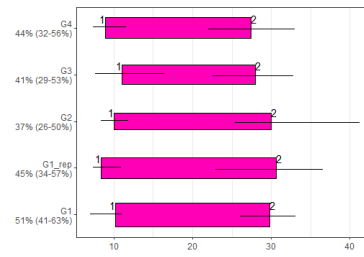
E - TCATA time-variability plot of Bitter



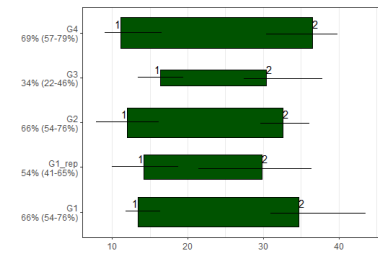
F - TCATA time-variability plot of Fluid



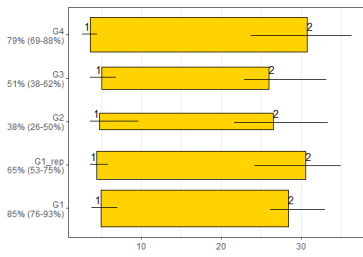
G - TCATA time-variability plot of Lemon



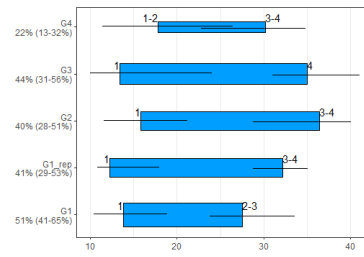
H - TCATA time-variability plot of Onion/Shallot



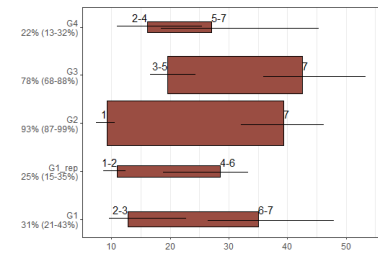
I - TCATA time-variability plot of Smooth/Fat



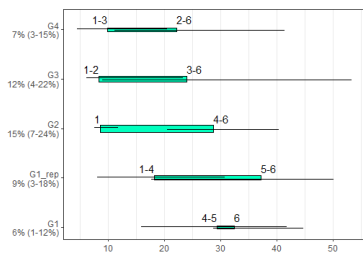
J - TCATA time-variability plot of Sour



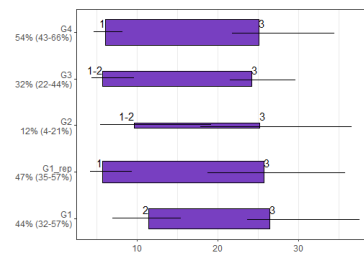
K - TCATA time-variability plot of Spicy



L - TCATA time-variability plot of Sweet



M - TCATA time-variability plot of Thick



N - TCATA time-variability plot of Tomato

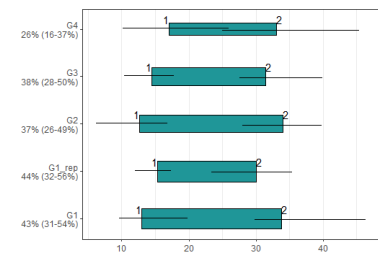
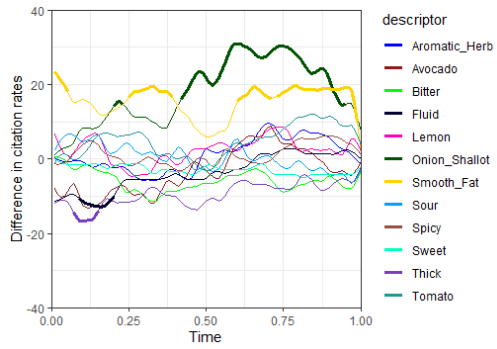
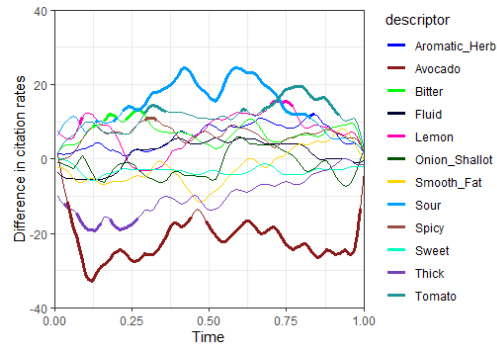


Figure 74. Representation of differences between G1, G1_rep and G4, by attribute (TCATA, without time-standardization).

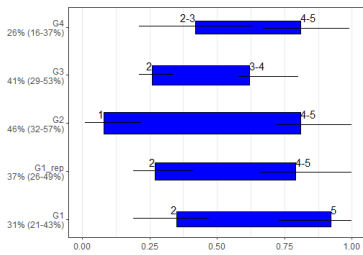
A - TCATA difference curves of G1-G1_rep with time standardization



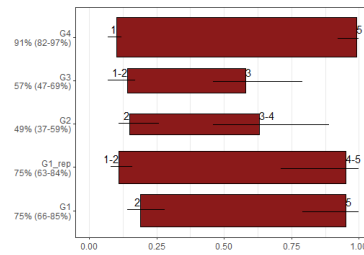
B - TCATA difference curves of G1-G4 with time standardization



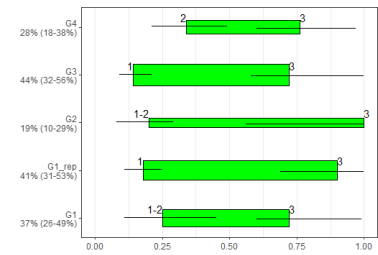
C - TCATA time-variability plot of Aromatic herbs



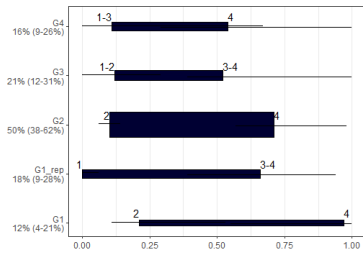
D - TCATA time-variability plot of Avocado



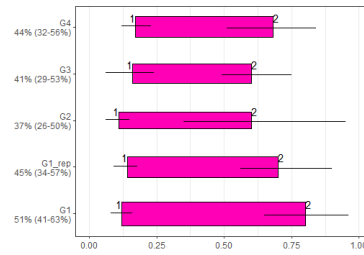
E - TCATA time-variability plot of Bitter



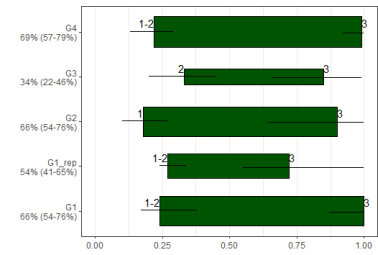
F - TCATA time-variability plot of Fluid



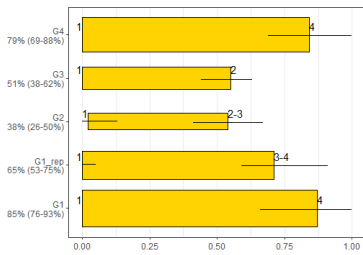
G - TCATA time-variability plot of Lemon



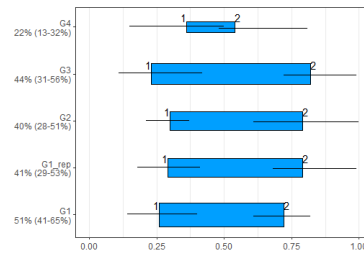
H - TCATA time-variability plot of Onion/Shallot



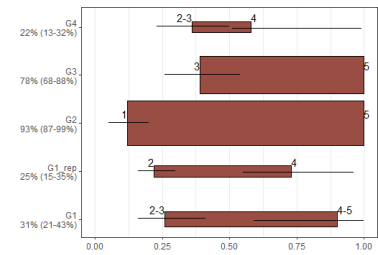
I - TCATA time-variability plot of Smooth/Fat



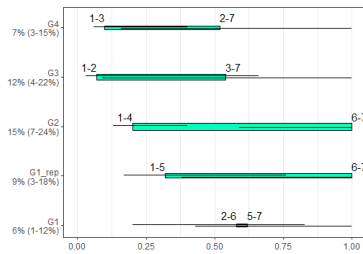
J - TCATA time-variability plot of Sour



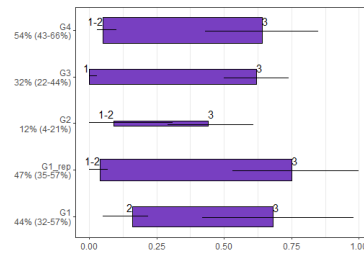
K - TCATA time-variability plot of Spicy



L - TCATA time-variability plot of Sweet



M - TCATA time-variability plot of Thick



N - TCATA time-variability plot of Tomato

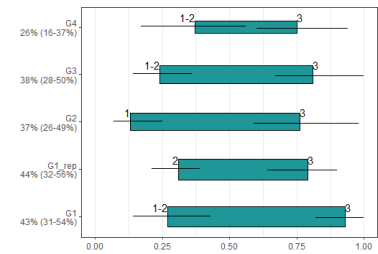
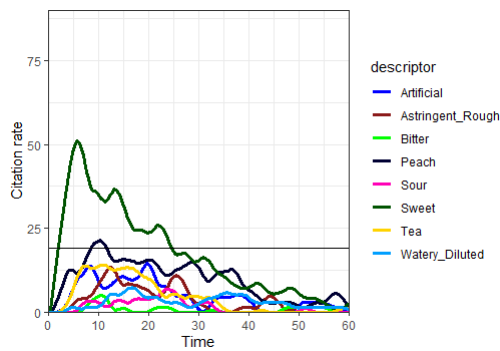


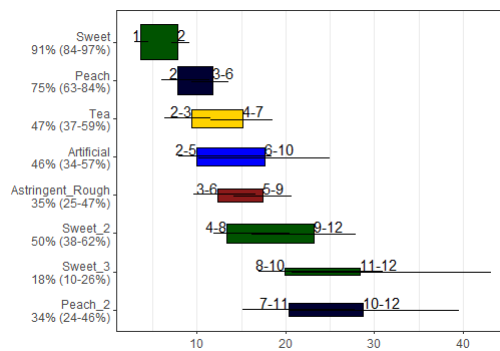
Figure 75. Representation of differences between G1, G1_rep and G4, by attribute (TCATA, with time-standardization).

Annex 5 – Additional results for Ice teas

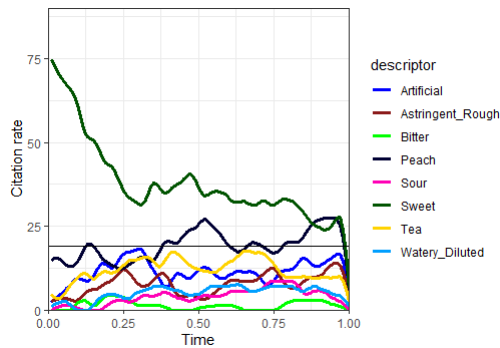
A - TDS curves of IT3 without time standardization



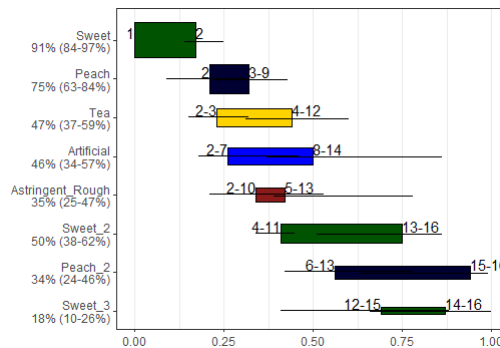
B - TDS time-variability plot of IT3 without time standardization



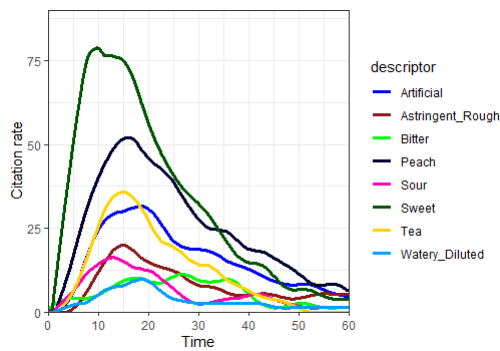
C - TDS curves of IT3 with time standardization



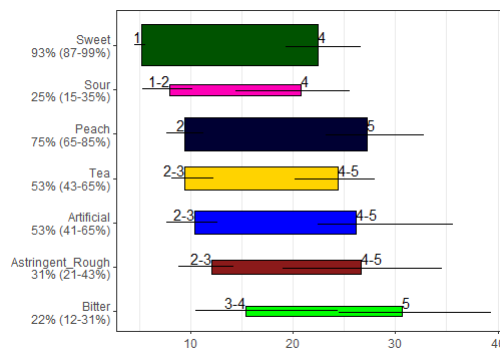
D - TDS time-variability plot of IT3 with time standardization



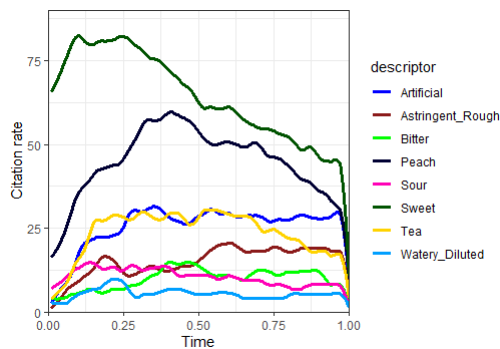
E - TCATA curves of IT3 without time standardization



F - TCATA time-variability plot of IT3 without time standardization



G - TCATA curves of IT3 with time standardization



H - TCATA time-variability plot of IT3 with time standardization

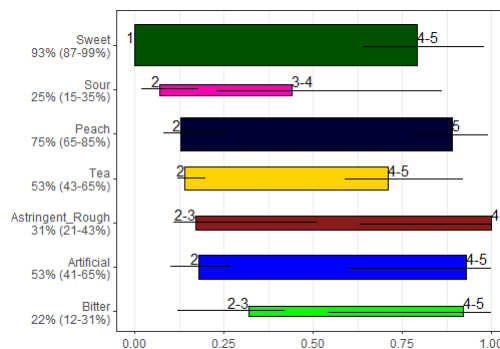
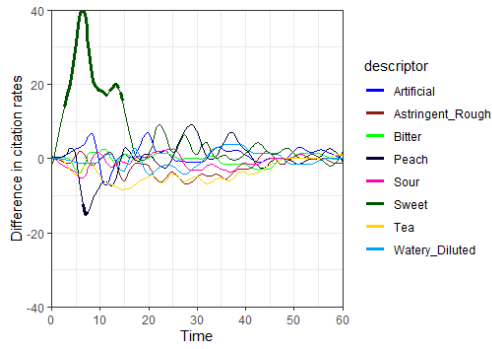
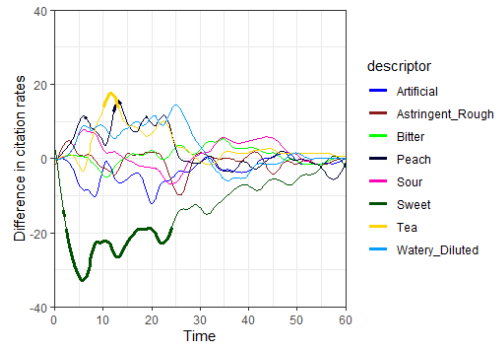


Figure 76. Representation of within-product differences between attributes for Ice tea IT3.

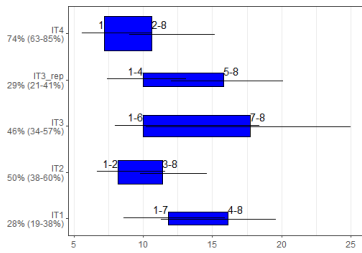
A - TDS difference curves of G3-G3_rep without time standardization



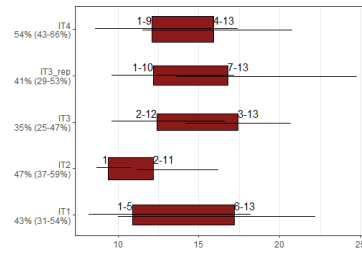
B - TDS difference curves of G1-G3 without time standardization



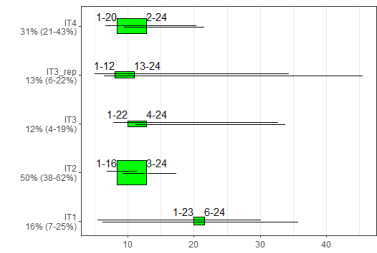
C - TDS time-variability plot of Artificial



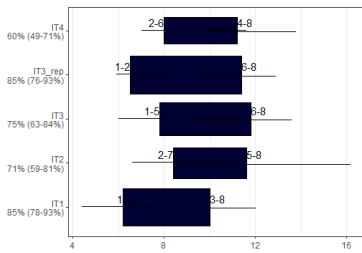
D - TDS time-variability plot of Astringent/Rough



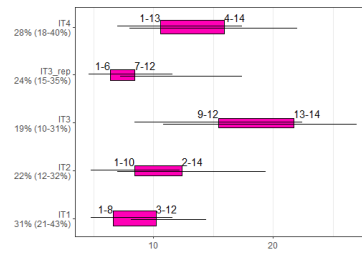
E - TDS time-variability plot of Bitter



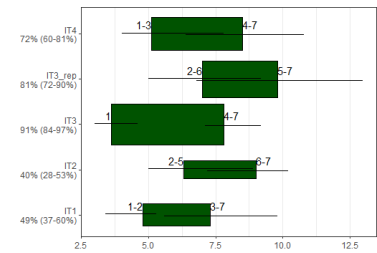
F - TDS time-variability plot of Peach



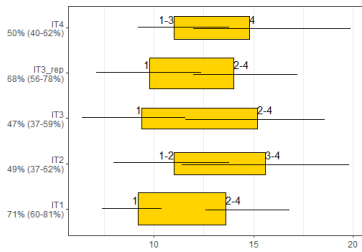
G - TDS time-variability plot of Sour



H - TDS time-variability plot of Sweet



I - TDS time-variability plot of Tea



J - TDS time-variability plot of Watery/Diluted

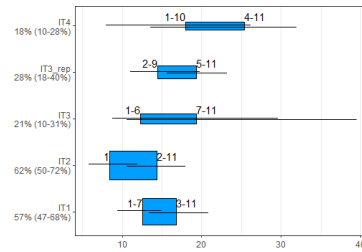
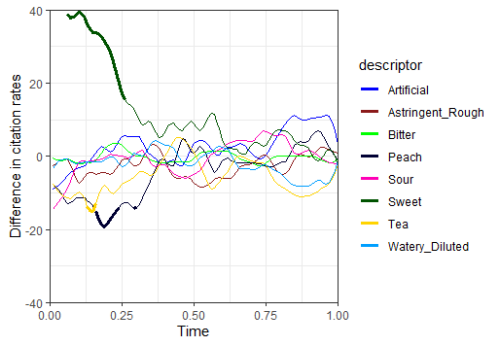
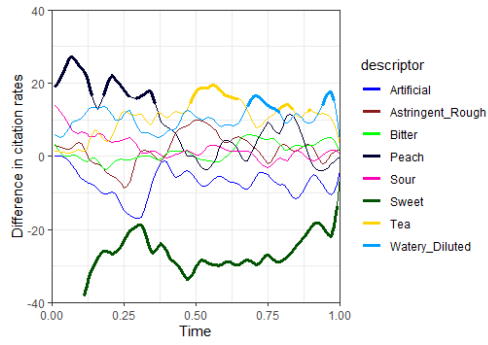


Figure 77. Representation of differences between G1, G3 and G3_rep, by attribute (TDS, without time-standardization).

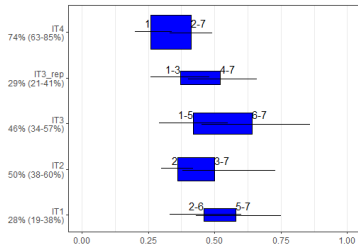
A - TDS difference curves of G3-G3_rep with time standardization



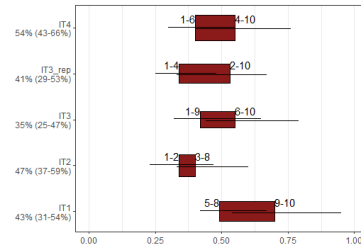
B - TDS difference curves of G1-G3 with time standardization



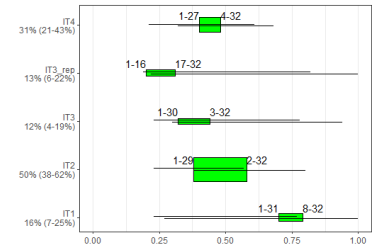
C - TDS time-variability plot of Artificial



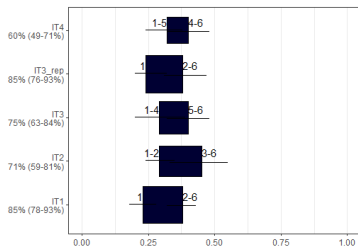
D - TDS time-variability plot of Astringent/Rough



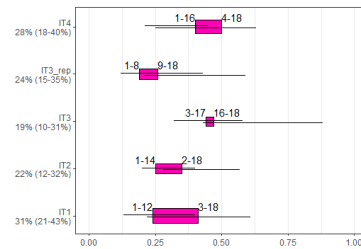
E - TDS time-variability plot of Bitter



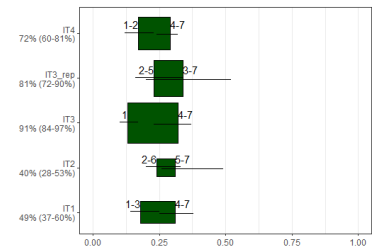
F - TDS time-variability plot of Peach



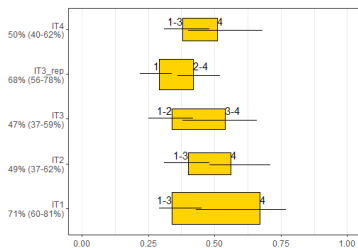
G - TDS time-variability plot of Sour



H - TDS time-variability plot of Sweet



I - TDS time-variability plot of Tea



J - TDS time-variability plot of Watery/Diluted

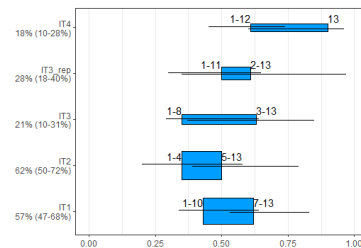
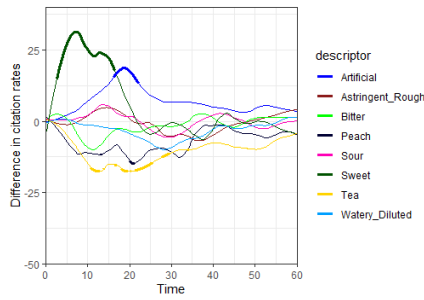
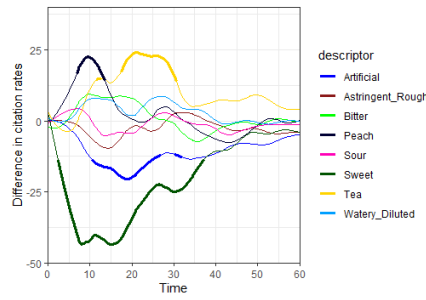


Figure 78. Representation of differences between G1, G3 and G3_rep, by attribute (TDS, with time-standardization).

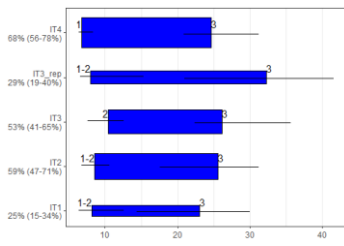
A - TCATA difference curves of G3-G3_rep without time standardization



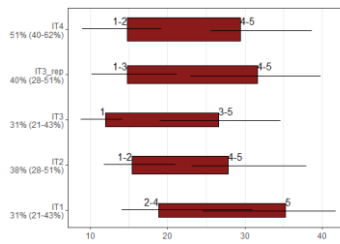
B - TCATA difference curves of G1-G3 without time standardization



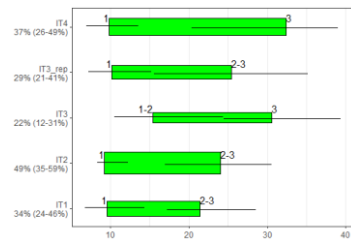
C - TCATA time-variability plot of Artificial



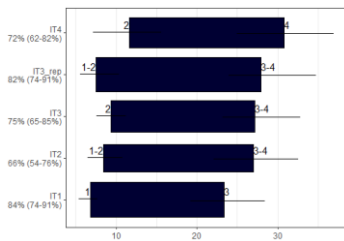
D - TCATA time-variability plot of Astringent/Rough



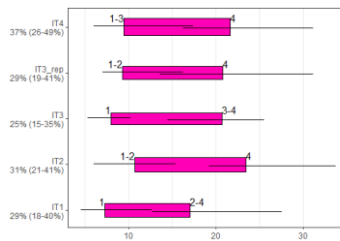
E - TCATA time-variability plot of Bitter



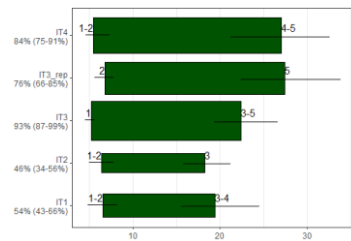
F - TCATA time-variability plot of Peach



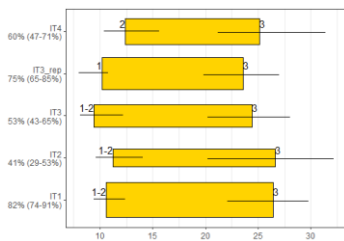
G - TCATA time-variability plot of Sour



H - TCATA time-variability plot of Sweet



I - TCATA time-variability plot of Tea



J - TCATA time-variability plot of Watery/Diluted

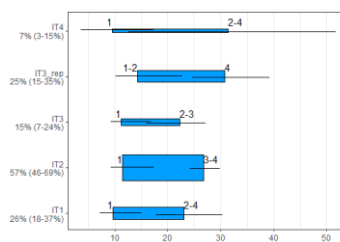
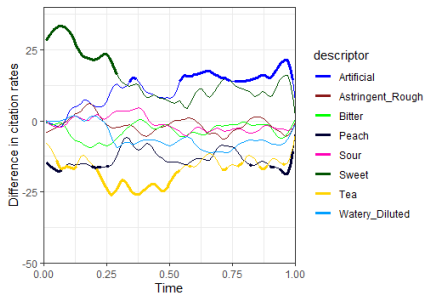
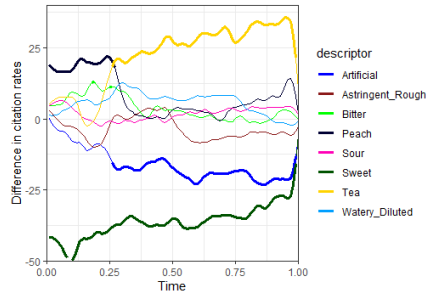


Figure 79. Representation of differences between G1, G3 and G3_rep, by attribute (TCATA, without time-standardization).

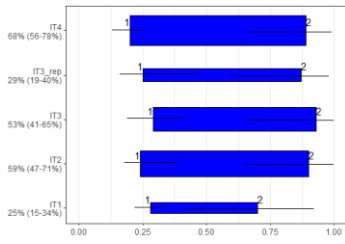
A - TCATA difference curves of G3-G3_rep with time standardization



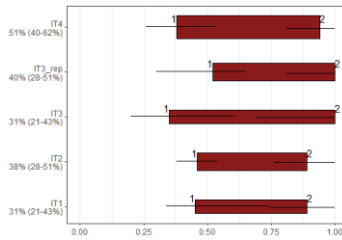
B - TCATA difference curves of G1-G3 with time standardization



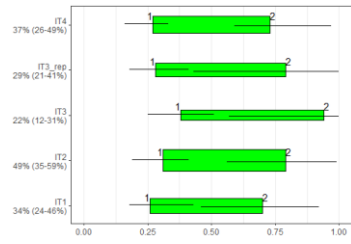
C - TCATA time-variability plot of Artificial



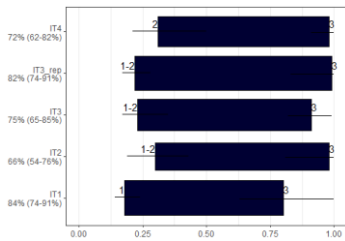
D - TCATA time-variability plot of Astringent/Rough



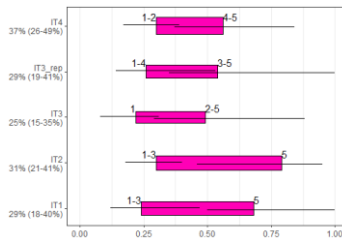
E - TCATA time-variability plot of Bitter



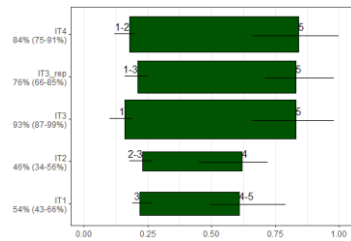
F - TCATA time-variability plot of Peach



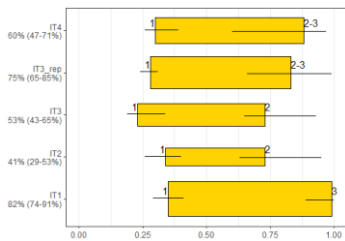
G - TCATA time-variability plot of Sour



H - TCATA time-variability plot of Sweet



I - TCATA time-variability plot of Tea



J - TCATA time-variability plot of Watery/Diluted

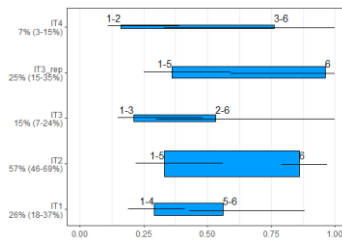


Figure 80. Representation of differences between G1, G3 and G3_rep, by attribute (TCATA, with time-standardization).

III.2.4. Product discrimination ability of temporal sensory evaluation methods used with consumers

This article was published in Food Quality and Preference:
<https://doi.org/10.1016/j.foodqual.2024.105123>

N. Béno: Investigation, Writing - Review & Editing. S. Cordelle: Investigation, Writing - Review & Editing. P. Schlich: Conceptualization, Methodology, Supervision, Writing - Review & Editing. M. Visalli: Conceptualization, Methodology, Software, Formal analysis, Writing – Original draft.

The objective of this study was to determine which temporal sensory evaluation methods (TDS, TCATA, AEF-D, AEF-A), variables (citations, durations, times of citation) and statistical analyses are the most discriminative when measures are collected from consumers. This study is based on data collected on model stimuli delivered by a gustometer (controlled temporal differences) and on four categories of commercial food products (crisps, guacamoles, chocolates, ice teas) presenting different levels of sensory complexity.

With controlled stimuli, analyses of citations and durations both revealed the same differences, suggesting that product discrimination is mainly brought by presence/absence of sensory attributes. Compared to TCATA, TDS highlighted slightly more significant differences (including expected and unexpected differences). Short differences in durations were not captured. AEF-A was less discriminative, but presented no validity nor repeatability issues.

With commercial food products, analyses of citations were more discriminative than analyses of durations and times of citations. However, the combined analysis of citations and time-related variables resulted in a small gain of discrimination. Applicability-based methods (TCATA and AEF-A) were overall more discriminative compared to dominance-based methods (TDS and AEF-D), but had more repeatability issue.

These results suggest that when data are collected from consumers, in most situations analysing durations and times of citations would not bring much more information compared to a time-independent analysis of citations. Performances of TCATA and TDS are almost similar when used with simple stimuli presenting clear temporal patterns. With complex food products, TCATA is the most discriminative, but also the less reliable, and AEF-A presents the better compromise between discrimination and repeatability. This study showed that there is probably no ultimate method, and that every gain in temporal resolution or discrimination is assumably at the cost of a loss of reliability.

1. Introduction

To measure the temporal sensory perception of panellists during the tasting of food products, essentially two methods are used in practice: Temporal Dominance of Sensations (TDS, Pineau et al., 2009) and Temporal Check-All-That-Apply (TCATA, Castura et al., 2016). TDS and TCATA are qualitative evaluation methods that enable panellists to record temporal sequences of dominant or applicable attributes while they are tasting products. Dominant attributes are most often defined as those that trigger the panellist's attention at a given time (only one attribute can be chosen at any time), while applicable attributes are all those that are perceived at a given time (one or several attributes can be chosen at any time). Except regarding their operational definitions, the two evaluation tasks are very similar as consumers have to click on buttons (TDS) or checkboxes (TCATA) corresponding to dominant/applicable attributes. In TCATA, the consumers must also uncheck attributes they no longer perceive, except when using the fading option (Ares et al., 2016).

With trained panels, TDS (Bord et al., 2019; Braghieri et al., 2016; Bruzzone et al., 2013; Charles et al., 2017; Devezeaux de Lavergne et al., 2015; Etaio et al., 2016; Frost et al., 2018; Labbe et al., 2009; Meillon et al., 2009; Ng et al., 2012; Paulsen et al., 2014) and TCATA (McMahon et al., 2017; Varela et al., 2021) have proven their ability to capture differences between products not measured by static descriptive evaluation methods. If the added value of temporal measures seems established when they are collected from trained panels, there is no such clear evidence with consumer studies. Indeed, some authors have questioned the complementarity of static and temporal measures collected from consumers. After having compared results obtained with CATA (Adams et al., 2007) and TCATA, Vidal et al. (2019) and Bommel et al. (2020) concluded that there is no guarantee that temporal measures bring additional information compared to static measures. Others authors are more nuanced, observing that additional insights can be obtained for some specific attributes (Alcaire et al., 2017).

Most of these conclusions were obtained based on the study of the discriminative ability of the evaluation methods. With trained or semi-trained panels, discrimination was found to the advantage of TCATA (Ares et al., 2015; Berget et al., 2020; Delompré et al., 2020; Meyners, 2020; Nguyen et al., 2018) or similar between the two methods (Sharma & Duizer, 2019; Wu et al., 2019). With consumer panels, discrimination was found to the advantage of TCATA (Ares

et al., 2015; Gonzalez-Estanol et al., 2022), of TDS (Nguyen & Wismer, 2022), or similar between the two methods (Esmerino et al., 2017; Paglarini et al., 2020; Velázquez et al., 2020b). However, the statistical tools used to study discrimination vary as several variables can be derived from TDS and TCATA data. Among all the studies mentioned above, most conclusions relied on evaluations of the evolution of citation rates over time using TDS (Pineau et al., 2009) or TCATA (Castura et al., 2016) curves or difference curves. Others relied on inferential statistics, including among others univariate and multivariate analyses of durations (Galmarini et al., 2017) or citations (Lenfant et al., 2009). Thus, it is difficult to know which variables better explain the discrimination between products and the differences between methods.

Recently, Attack-Evolution-Finish (AEF) has been introduced as an alternative temporal evaluation method, in between static and dynamic methods. AEF enables consumers to report periods of perception of dominant (AEF-D, Visalli et al., 2020) or applicable (AEF-A, Visalli et al., 2022) attributes retrospectively after the tasting. Unlike concurrent measures collected with TDS and TCATA (continuous data collection), AEF measures are of low temporal resolution: only three periods corresponding to the beginning, middle and end of perception are evaluated. Despite that, AEF-D was shown to discriminate chocolates in the same way than TDS (Visalli et al., 2020), and on lemon-flavoured carbonated alcoholic drinks AEF-A was shown to be discriminative compared to a concurrent evaluation method (Visalli et al., 2022). These two studies have revived the debate on what is really measured with temporal methods depending on whether consumers evaluate products concurrently or retrospectively to the tasting. Thus, additional works have been conducted to document the temporal resolution (Visalli, Mahieu, et al., 2023) and the validity (Visalli et al., 2023) of temporal measures collected from consumers with TDS, TCATA and AEF methods. These works showed that the temporal resolution of TDS and TCATA was below most expectations, and that repeatability and heterogeneity between consumers could be an issue threatening the validity of the measures.

The objective of this study was to determine which temporal sensory evaluation methods (TDS, TCATA, AEF-D, AEF-A), variables (citations, durations, times of citation) and statistical analyses are the most discriminative while ensuring valid differences when measures are collected from consumers. For this, an in-depth analysis was performed on a published dataset

including measures carried out on controlled temporal taste stimuli delivered by a gustometer (Béno et al., 2023). Then an attempt at generalization was made based on an analysis of another published dataset including four categories of commercial food products (crisps, guacamoles, chocolates, ice teas) presenting different levels of sensory complexity (Visalli, Cordelle, et al., 2022).

2. Materials and methods

2.1. Datasets

Table 60 summarizes the main characteristics of the datasets. All consumers were preselected from a population registered in the ChemoSens Platform's PanelSens database (declared to the relevant authority, authorization number 1148039). They had to fill out an informed consent form, and were compensated for their participation with vouchers worth 10 euros per session. They were randomly assigned to one panel among A, B and C (Béno et al., 2023) or among D, E, F or G (Visalli, Cordelle, et al., 2022), respectively, with a constraint of balance in gender and age between panels. Each panel used a different temporal method to evaluate all the samples. Since the objective was to compare the methods, the order of presentation of the samples was the same for all the consumers.

Liquid solutions were delivered using a gustometer model Burghart GU002. The gustometer is an instrument that enables to program the delivery of predetermined volume and concentration of liquid compounds to the panellist's tongue over a specified period of time.

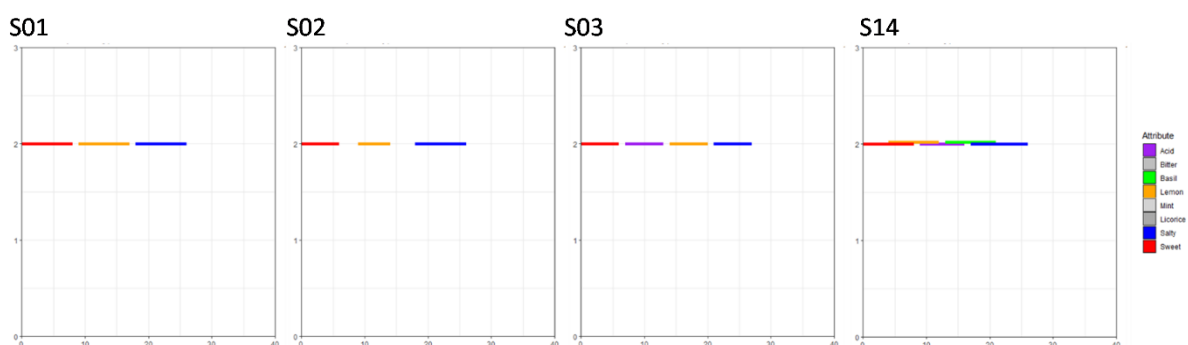


Figure 81. Controlled temporal stimuli delivered using a gustometer. X-axis: time, y-axis: concentration of the compound. From left to right: stimuli S01, S02, S03, S14.

Figure 81 describes the controlled temporal stimuli delivered using the gustometer. They include three to five compounds, either aromatic (*Basil*, *Lemon*) or sapid (*Acid*, *Salty*, *Sweet*), delivered at constant concentration and iso-intensity. Three distractors (*Bitter*, *Mint*,

Liquorice) were added to these five attributes to make a list of eight attributes proposed to panellists.

Table 61 details the times of begin and end of delivery of the compounds constituting the controlled temporal stimuli. Depending on the chemical composition of the stimuli, the differences in sensory perception between pairs of stimuli can be of three different natures: differences in attribute citations (corresponding to one attribute present in the first stimulus but not in the second, e.g. between S01 and S03 for *Acid*), differences in times of first citations (corresponding to different times of begin of delivery, e.g. between S02 and S03 for *Lemon*), and differences in durations of dominance/applicability (corresponding to different durations of delivery, e.g. between S01 and S02 for *Sweet*).

The chosen products were commercial products accepted by a majority of consumers, easy to prepare for the experimenters, and varying in texture. Crisps varied in fat and salt; ice teas varied in sugar/sweetener, flavour and variety of tea; guacamoles varied in avocado content and fat; and dark chocolates varied in cocoa origin and content. The levels of complexity and sensory differences of the product spaces were reported by the consumers using a questionnaire (Visalli, Cordelle, et al., 2022).

Table 60. Characterization of datasets.

Type of product (data paper)	Number of different samples (number of replicated samples)	Number of attributes	of Settings	Complexity	Differences	Number of consumers (panel name)			
						TDS	TCATA	AEF-A/AEF-RATA	AEF-D
Model solutions (Béno et al., 2023)	4 (2)	8	Lab	-	-	50 (A)	50 (B)	49 (C)	-
Guacamoles (Visalli, Cordelle, et al., 2022)	4 (1)	12	Lab	High	High	70 (D)	73 (E)	74 (F)	75 (G)
Ice teas (Visalli, Cordelle, et al., 2022)	4 (1)	8	Lab	Medium	High	70 (D)	73 (E)	74 (F)	75 (G)
Crisps (Visalli, Cordelle, et al., 2022)	4 (1)	9	Home	Low	Medium	70 (D)	72 (E)	74 (F)	75 (G)
Chocolates (Visalli, Cordelle, et al., 2022)	4 (1)	12	Home	High	Medium	68 (D)	72 (E)	71 (F)	75 (G)

Letters A-G between parentheses after panel sizes denote different panels.

Table 61. Times of begin and end of delivery of the controlled temporal stimuli (in seconds)

	S01	S02/S02_2	S03	S14/S14_2
Acid			7 – 13	9 – 16
Basil				13 – 21
Lemon	9 – 17	9 – 14	14 – 20	4 – 12
Salty	18 – 26	18 – 26	21 – 27	18 – 26
Sweet	0 – 8	0 – 6	0 – 6	0 – 8

2.2. Data analysis

Statistical analyses were performed using R 4.1.0 software (R Core team, 2020). All tests were performed at the 5% alpha-risk level.

For all the methods, to test whether discrimination was related to presence/absence of attributes, an attribute was set to 1 if it had been cited as dominant/applicable by a consumer at least once during the full perception, 0 otherwise. It remains to transform the data into CATA-type data.

To assess overall differences between samples, two multivariate analyses were run. Differences in presence/absence were assessed using the multiple response Correspondence Analyses (mrCA, Mahieu et al., 2021) framework. Differences in durations of dominant/applicable attributes were assessed using Canonical Variate Analysis (CVA, Galmarini et al., 2017). Product maps were represented using biplots, and pairwise product comparisons were performed on the significant dimensions of mrCAs or CVAs. The two analyses were performed with and without including the replicated samples to quantify discrimination and non-repeatability, respectively.

To characterize differences between pairs of samples at the attribute level, the differences in citation rates at each time were plotted (referred to as "difference plots" in the rest of the manuscript). For TDS and TCATA, data were first discretized every 0.1 s and left standardized (as if the chronometer started with the first citation of an attribute) to reduce heterogeneity between consumers without distorting the sequences. The curves were smoothed using locally estimated scatterplot smoothing (LOESS, Cleveland, 1979). Significant differences were established based on proportion tests with Yates' continuity correction to prevent overestimation of statistical significance for expected count smaller than five (Yates, 1934). Significant differences were represented as bold lines (TDS and TCATA) or points (AEF-RATA, AEF-D, AEF-A) on the plots. Total durations of significant differences were computed and considered "significant" when they lasted more than 5% of the total time (no statistical test performed). Absolute values of differences in citation rates were averaged over times/periods and attributes to quantify the overall discrimination for each pair of samples. Then, in the same idea as the F-ratio of an ANOVA, discriminating ratios were computed between the overall discrimination value of pairs of different samples (expected "high") and the overall discrimination value of pairs of replicated samples (expected "low"). Discriminating ratios higher than 1 were considered "significant" (no statistical test performed).

To characterize the nature of the differences between pairs of samples at the attribute level, three univariate analyses were run. For the assessment of differences in applicability, type-III ANOVA tables were computed by attribute after having adjusted the following generalized linear model with binomial error distribution: $citation (0 \text{ or } 1) = sample + consumer + error$ (Agresti, 2013). For the assessment of differences in durations, type-III ANOVA tables were computed by attribute after having adjusted either a linear model (for TDS and TCATA) or a generalized linear model with Poisson error distribution (for AEF): $duration = sample + consumer + error$. Pairwise differences between samples were estimated using least square means, and expressed as log of odd ratios (Sroka & Nagaraja, 2018) or difference in durations (Galmarini et al., 2017). Finally, as no statistical model fits for the analysis of attribute time of first citations in TDS and TCATA (because of unbalanced and non-normal data), the differences in median attribute times of first citations were assessed based on a bootstrap approach as described in Visalli, Mahieu, et al. (2023), with a threshold of 25% for the overall citation rate for the selection of attributes.

3. Results

3.1. Model solutions

Figure 82 shows very similar product configurations regardless of the evaluation method and the analysed variable (citation or duration), but different interpretations regarding product characterization. The mrCAs (Figure 82.A, Figure 82.C and Figure 82.E) point out that the main differences observed (length of the arrows) were due to *Basil*, more cited for S14, and *Acid* more cited for S03 and S14. The arrows representing *Sweet*, *Lemon* and *Salty* (the three attributes corresponding to compounds present in all the stimuli) are very short, sometimes shorter than the arrows of the distractors. S01 and S02 (which differed only in durations) are not discriminated, nor are the replicated stimuli. These observations appear congruent with the chemical differences related to presence/absence of compounds in the stimuli (Figure 81). For the three evaluation methods (without including the replicated stimuli, not represented on Figure 82), the mrCAs highlight three significant dimensions, and the mrChi² statistic (representative of the overall size of the differences in citations) is greater with TDS, then with AEF-RATA, then with TCATA.

Basil and *Acid* are also discriminative attributes with CVAs of durations (Figure 82.B, Figure 82.D and Figure 82.F). However, the sizes of the arrows do not correspond to the chemical differences in durations of delivery of compounds, and notably for the arrows representing the distractors. The replicated stimuli are not discriminated except for S14 with TDS (on axis 1-3, not represented on Figure 82, see pairwise differences on Figure 84). The small

differences between S01 and S02 in durations of delivery of the compounds (*Sweet* and *Lemon*) are no more captured than with mrCA. Three axes were tested significant for TDS and TCATA, and two for AEF-RATA. The F-statistic (representative of the overall size of the differences in durations) is greater with TDS, then with TCATA, then with AEF-RATA.

These results show that on controlled stimuli, multivariate analyses of citations (mrCA) and durations (CVA) both capture differences in presence/absence. However, short differences in durations between stimuli are not necessarily captured by CVA.

Figure 83 shows the evolution of differences in attribute citations rates over time. Unexpected differences between the replicated samples are observed with TCATA (for *Acid*, Figure 83.A) and TDS (for *Basil* and *Acid*, Figure 83.B).

Controlled stimuli S01 and S02 differed on the durations of delivery for *Lemon* and *Sweet* (Table 61), but only TDS highlights significant differences for *Lemon* (Figure 83.C).

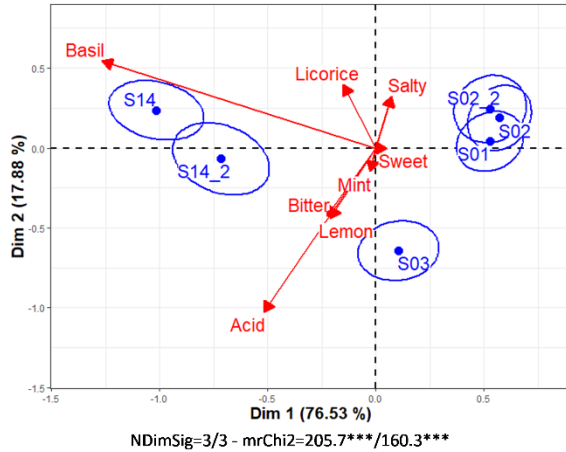
Stimuli S01 and S02 differed from S03 on the times and durations of delivery for *Lemon*, *Salty*, on the durations of delivery for *Sweet* (S01 vs. S03 only) and on the applicability for *Acid* (Table 61). TDS highlights significant differences for all these attributes, while TCATA does not capture difference in citation rates for *Lemon*, and AEF-RATA only captures difference in *Acid* (Figure 83.D). The duration of significant differences observed for *Sweet* is far longer than the difference in the duration of delivery (2 s). Similar conclusions can be drawn from Figure 83.F, showing that the difference in citation rates for *Sweet* is not due to the difference in the duration of delivery.

Stimuli S01 and S02 differed from S14 on the applicability for *Acid* and *Basil*, on the time of delivery for *Lemon* and on the durations of delivery for *Sweet* (S02 vs. S14 only, Table 61). All the methods agree and highlight significant differences for *Acid*, *Basil* and *Lemon* between S01/S02 and S14, but with TDS and TCATA also for *Sweet* (Figure 83.E) and *Salty* (Figure 83.G).

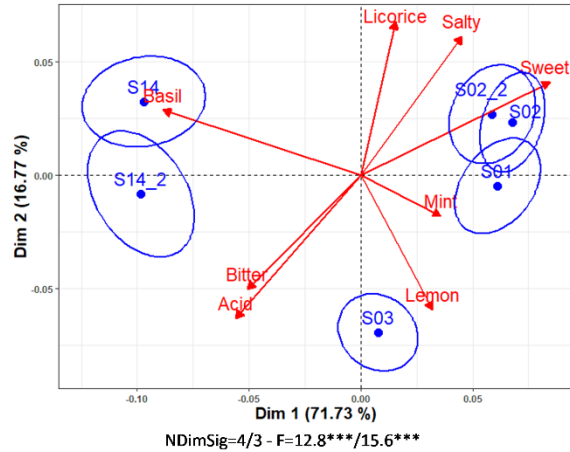
Stimuli S03 and S14 differed on the applicability for *Basil*, on the times durations of delivery for *Acid*, *Lemon* and *Salty*, and on the duration of delivery for *Sweet* (Table 61). This is the only pair of samples for which conclusions are opposed between TDS and TCATA, for *Sweet* (higher citation rates in S03 with TDS and in S14 with TCATA, Figure 83.H). The two methods agreed on other differences while AEF-RATA only show significant differences for *Basil* and *Lemon*.

These results suggest that, on controlled stimuli, citation rates capture qualitative changes of applicability over time, but with a relatively low precision. Short differences in durations of delivery in the stimuli are not systematically highlighted on the temporal curves. TDS highlights more significant differences than TCATA, and AEF-RATA less. However, whatever the method, some differences are not congruent with the chemical composition of the stimuli.

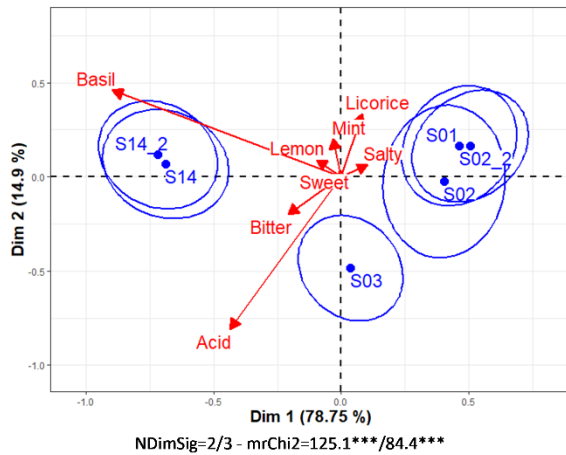
A - mrCA of citations (TDS)



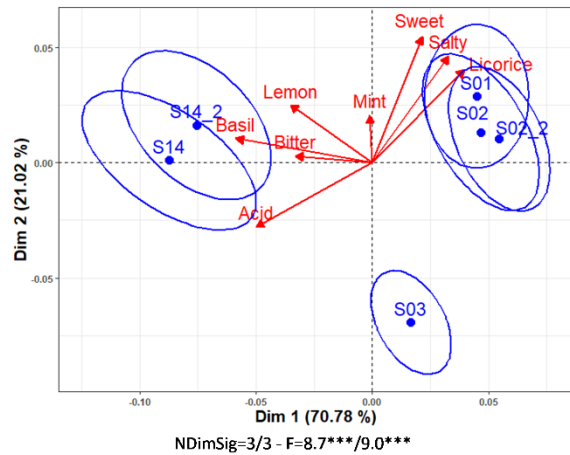
B - CVA of durations (TDS)



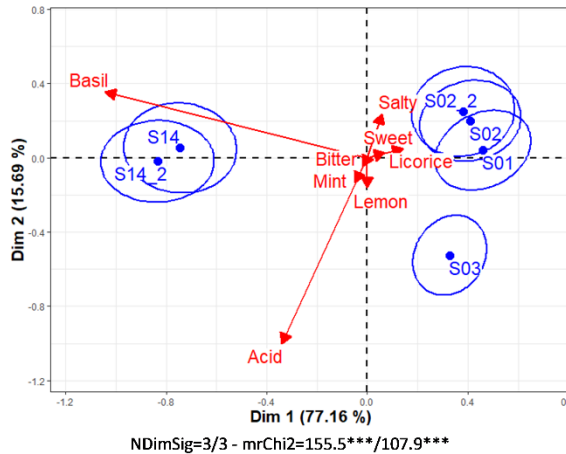
C - mrCA of citations (TCATA)



D - CVA of durations (TCATA)



E - mrCA of citations (AEF-RATA)



F - CVA of durations (AEF-RATA)

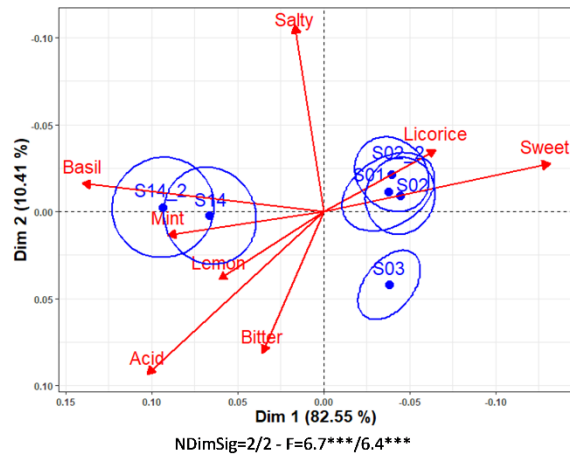


Figure 82. Biplots of multivariate analyses (axes 1-2), with 95% confidence ellipses. First row: TDS, second row: TCATA, third row: AEF-RATA. First column: mrCA of citations, second column: CVA of durations. NDimSig: number of significant dimensions (all stimuli/without replicates). mrChi2: mrChi2 statistic (all stimuli/without replicates). F: MANOVA statistic (all stimuli/without replicates). ***p-value<0.001.

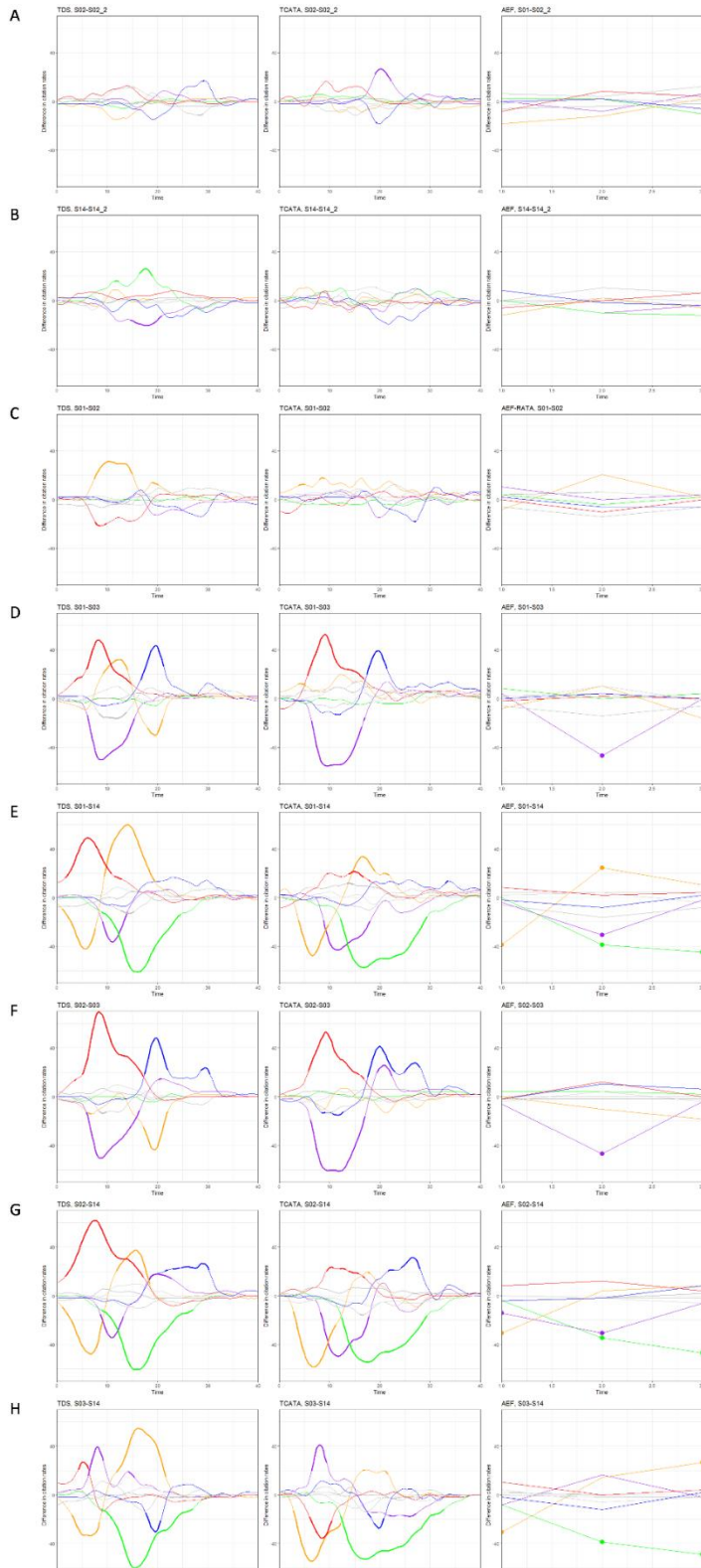


Figure 83. Pairwise differences in citation rates (Y-axis) plotted over time (X-axis). A-H rows of plots each refers to one pair comparison: A=S02-S02_2, B=S14-S14_2, C=S01-S02, D=S01-S03, E=S01-S14, F=S02-S03, G=S02-S14, H=S03-S14. Within each row, first column refers to TDS, second to TCATA and third to AEF-RATA. Bold lines (for TDS and TCATA) and points (for AEF-RATA) correspond to significant differences in citation rates ($\alpha=0.05$).

Controlled stimuli		S02_S02_2			S14-S14_2			S01-S02			S01-S03			S01-S14			S02-S03			S02-S14			S03-S14		
		TDS	TCATA	AEF-RATA	TDS	TCATA	AEF-RATA	TDS	TCATA	AEF-RATA	TDS	TCATA	AEF-RATA	TDS	TCATA	AEF-RATA	TDS	TCATA	AEF-RATA	TDS	TCATA	AEF-RATA	TDS	TCATA	AEF-RATA
Acid	Citation	0.6	2.1	0.1	-0.7	-0.2	-0.5	-0.8	-1.4	0.4	-4.0	-4.5	-2.3	-2.6	-3.2	-1.4	-3.3	-3.1	-2.7	-1.8	-1.9	-1.8	1.4	1.2	0.9
	Duration	0.6	1.5	-0.1	-2.8	-0.2	-0.2	-0.9	0.4	0.4	-3.5	-3.6	-0.7	-1.1	-3.7	0.6	-2.7	-4.0	-1.1	-0.3	-4.1	-1.0	2.4	-0.1	0.1
	Time	0.0	0.4	NT	0.2	-0.4	NT	-1.2	-1.1	NT	9.2	9.1	NT	6.8	7.1	NT	10.5	10.4	NT	8.0	8.2	NT	-2.4	-2.0	NT
	Citation rate over time	0.0	3.6	0.0	6.8	0.0	0.0	0.0	0.0	0.0	9.6	10.6	1.0	5.2	9.4	1.0	11.6	14.2	1.0	9.6	9.8	2.0	4.8	7.4	0.0
Basil	Citation	-0.9	0.3	-0.5	2.5	-0.3	-0.6	0.5	-0.3	-0.6	-0.7	-0.8	0.8	-7.8	-4.2	-6.0	-1.2	-0.5	1.5	-8.3	-3.9	-5.4	-7.0	-3.4	-6.8
	Duration	-0.3	0.7	-0.3	2.2	1.2	-0.2	0.1	-1.0	0.1	-0.2	-0.3	0.8	-6.5	-8.7	-1.6	-0.3	0.6	0.7	-6.6	-7.7	-1.6	-6.3	-8.3	-2.3
	Time	NT	NT	NT	-0.2	-0.4	NT	NT	NT	NT	2.9	5.8	NT	2.9	5.8	NT	NT	NT	NT	2.0	-1.8	NT	-0.2	1.8	NT
	Citation rate over time	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16.0	17.4	2.0	0.0	0.0	0.0	18.4	17.2	2.0	18.4	20.8	2.0
Lemon	Citation	-1.0	-0.5	-1.4	0.4	0.0	-0.2	2.9	1.7	2.1	-0.7	0.8	0.2	-0.2	0.0	0.9	-3.5	-0.9	-1.9	-3.1	-1.7	-1.2	0.5	-0.8	0.7
	Duration	-1.1	-1.3	-0.3	0.5	0.1	-0.1	3.3	3.8	0.1	0.4	2.9	-0.1	2.5	0.3	0.0	-2.9	-0.9	-0.3	-0.8	-3.4	-0.2	2.1	-2.5	0.1
	Time	0.4	-0.3	NT	0.2	-0.7	NT	-1.0	-0.8	NT	-1.2	-0.8	NT	4.6	4.4	NT	0.0	0.0	NT	5.4	5.3	NT	5.6	5.2	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	0.0	9.4	2.6	0.0	8.4	1.5	0.0	17.0	12.4	2.0	6.0	0.0	0.0	15.0	9.4	1.0	15.8	11.4	2.0
Salty	Citation	0.2	-0.6	-0.4	0.1	-1.2	0.0	-0.8	-0.2	-0.4	0.9	0.4	0.7	0.1	1.5	0.4	1.7	0.7	1.1	1.0	1.7	0.7	-0.7	1.0	-0.3
	Duration	-0.3	-0.7	0.0	-1.4	-1.7	0.0	-0.8	0.2	-0.1	2.5	2.7	0.0	1.6	2.4	-0.1	3.3	2.5	0.1	2.3	2.2	0.0	-0.9	-0.3	-0.1
	Time	0.8	0.2	NT	1.0	0.0	NT	-0.4	-0.4	NT	-2.8	-2.8	NT	-1.0	0.2	NT	-2.4	-2.4	NT	-0.5	0.3	NT	1.9	2.7	NT
	Citation rate over time	0.4	0.2	0.0	0.0	0.2	0.0	0.0	0.6	0.0	4.8	6.0	0.0	0.2	0.8	0.0	6.8	12.6	0.0	8.4	5.6	0.0	3.2	3.0	0.0
Sweet	Citation	1.1	0.0	-0.4	1.0	-1.0	-1.3	0.9	-1.7	0	1.5	0.0	0.3	2.0	-0.7	1.6	0.6	1.7	0.3	1.1	1.0	1.6	0.5	-0.7	1.3
	Duration	1.0	1.2	0	1.3	-0.7	0.0	-1.4	0.2	-0.1	4.1	4.6	0	4.1	2.6	0.1	5.5	4.3	0.1	5.5	2.3	0.2	0.0	-2.0	0.1
	Time	0.0	0.0	NT	0.0	0.0	NT	0.0	0.0	NT	0.0	0.0	NT	0.0	0.0	NT	0.0	0.0	NT	0.0	0.0	NT	0.0	0.0	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	11.0	9.5	0.0	11.0	5.8	0.0	13.0	11.0	0.0	15.8	6.0	0.0	2.4	4.6	0.0
Bitter*	Citation	-0.2	-1.5	0.8	-0.2	0.5	1.1	0.7	1.5	-1.5	-0.9	-0.9	-1.2	-0.6	-1.0	-1.5	-1.6	-2.3	0.3	-1.3	-2.5	0.0	0.3	-0.2	-0.3
	Duration	-0.4	-1.2	0.5	0.0	-0.3	0.3	1.3	1.2	-0.6	0.0	0.2	-0.6	0.1	-0.1	-0.6	-1.2	-1.0	0.0	-1.1	-1.3	-0.1	0.1	-0.3	-0.1
	Time	NT	NT	NT	1.4	1.2	NT	1.0	NT	NT	7.4	3.2	NT	3.3	1.6	NT	6.5	5.1	NT	2.6	3.1	NT	-4.4	-1.8	NT
	Citation rate over time	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.6	0.0	0.0	3.0	0.0	0.0	1.8	0.6	0.0	0.0	1.0	0.0	0.0	0.0	0.0
Licorice*	Citation	0.3	0.2	-0.9	0.8	-0.3	0.0	-1.4	0.5	0.9	0.0	1.6	0.6	-2.0	1.0	1.2	1.4	1.1	-0.3	-0.6	0.5	0.3	-2.0	-0.6	0.6
	Duration	-0.6	0.4	-0.6	0.4	-0.6	-0.3	-0.5	0.6	0.6	0.0	1.9	0.3	-0.6	2.0	0.6	0.5	1.3	-0.4	-0.1	1.4	0.0	-0.6	0.0	0.4
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Mint*	Citation	1.2	-0.5	0.0	-0.6	0.2	0.3	0.3	0.7	1.3	0.3	0.7	0.6	0.6	0.0	0.3	0.0	0.0	-0.7	0.3	-0.7	-1.0	0.3	-0.7	-0.3
	Duration	0.9	-1.0	0.0	-0.4	1.6	-0.1	-0.1	2.0	0.7	0.4	1.7	0.4	0.7	0.3	0.1	0.5	-0.3	-0.3	0.8	-1.6	-0.6	0.3	-1.4	-0.3
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	1.8	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.0	0.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Multivariate	mrChi2	3.2	6.0	2.4	2.9	1.0	2.1	8.5	5.6	6.5	18.7	15.5	10.2	46.0	29.8	32.2	24.0	11.5	14.4	48.0	28.6	28.3	33.0	16.4	30.7
	F-MANOVA	0.9	2.0	0.8	2.2	1.2	0.6	2.7	1.8	1.5	6.9	5.7	2.9	18.5	9.3	5.5	9.5	7.2	3.2	18.0	9.0	7.6	17.5	4.9	10.6
	Discriminating ratio	0.9	0.9	1.1	1.1	1.1	0.9	1.1	1.3	1.1	1.8	2.3	1.3	2.5	2.7	2.3	2.0	2.0	1.2	2.7	2.6	2.0	2.1	2.3	2.1

Figure 84. Summary of pairwise differences between controlled stimuli, by attribute (rows) and method (columns). Attributes suffixed by an asterisk are distractors. Citation: log of the odd ratio (ANOVA citation = sample + consumer + error). Duration: mean difference in duration (ANOVA duration = sample + consumer + error). Time: median time of citation (bootstrap test). Citation rate over time: total duration of significant differences in citation rates (difference plot). mrChi2: mrChi2 statistic restricted to the pair of samples (mrCA of citations). F-MANOVA: F statistic of the MANOVA restricted to the pair of samples (CVA of durations). Discriminating ratio: average difference between the pair of samples divided by average difference between the replicated samples (citation rates over time). Red and orange cells: significant unexpected differences between replicated samples (red cells: differences not related to citations). Yellow cells: significant incongruent differences (compared to the chemical composition of the stimuli). Green cells: significant expected differences (compared to the chemical composition of the stimuli). Blue cells: non-significant differences while expected (compared to the chemical composition of the stimuli). NT: discrimination not tested.

Figure 84. Summary of pairwise differences between controlled stimuli, by attribute (rows) and method (columns). shows that all attributes confounded (last three rows), all pairs of samples are discriminated (green cells) by the analyses of citations (with mrChi2), durations (with MANOVA) and discriminating ratio (with difference plots), except for the pair S01/S02. This latter is only discriminated by the discriminating ratio or the analysis of durations with TDS. This result shows that on controlled stimuli CVA of durations and difference plots enable to discriminate more samples compared to mrCA of citations. However, unexpected discrimination (red cells) is also observed between replicated samples with CVA (S14 with TDS) and discriminating ratio (S02 with AEF-RATA and S14 with TDS and TCATA).

The univariate analyses by attribute enable a more detailed investigation of the differences between the pairs of samples. Most of the times, the methods agree on the differences, but AEF-RATA is clearly less discriminative than TDS and TCATA. S02 and S14 are unexpectedly discriminated from their replicates: for *Acid* with TCATA (citations and durations of significant differences), for *Acid* (durations and durations of significant differences) and *Basil* with TDS (citations, durations and durations of significant differences). This result shows that some repeatability issues can be observed when consumers perform temporal measures. Significant differences incongruent with the chemical composition of the stimuli are also observed for some attributes. These differences are mostly observed with TDS and TCATA, some being common to both methods (*Acid* for S01/S03, S01/S14, S02/S03; *Salty* for S01/S14 and S02/S14; *Sweet* for S01/S14, S02/S03; S02/S14; S03/S14).

Table 62 shows that all existing differences due to presence/absence of compounds are captured by the analyses of citations, except one for AEF-RATA. In return, AEF-RATA has no repeatability issue, conversely to TDS and TCATA, which also point out more incongruent differences. Only half of the existing differences in attribute durations are captured by TDS and TCATA, and a quarter by AEF-RATA. Notably, no method captures the short-duration differences for *Lemon* between S01 and S03, for *Sweet* between S01 and S02, and for *Salty* between S03 and S14. TDS point outs more incongruent differences and unexpected differences compared to TCATA, while AEF shows none due to its low discriminative ability for durations. TDS and TCATA capture eighty percent of difference in times of first citations, with no unexpected difference but still more incongruent differences with TDS. The same pattern can be observed with durations of significant differences: TDS is more discriminative (100% of

expected differences captured), but shows more incongruent and unexpected differences compared to TCATA. AEF-RATA is less discriminative, but presents no potential validity nor repeatability issue. These results show that contrarily to multivariate analyses (that mostly conclude similarly with citations and durations), unidimensional analyses of durations and times of citations (time-dependent variables) are complementary to unidimensional analyses of citations (time-independent variable).

Table 62. Count of discriminated pairs of samples.

		TDS	TCATA	AEF-RATA
Citation	Expected	7/7	7/7	6/7
	Incongruent	4	2	1
	Unexpected	1	1	0
Duration	Expected	10/20	11/20	5/20
	Incongruent	5	2	0
	Unexpected	2	0	0
Time	Expected	8/10	8/10	Not tested
	Incongruent	5	3	Not tested
	Unexpected	0	0	Not tested
Citation rate over time	Expected	19/19	17/19	10/19
	Incongruent	8	7	0
	Unexpected	2	1	0

Expected: observed/expected differences corresponding to the chemical composition of the stimuli. Incongruent: differences not corresponding to the chemical composition of the stimuli. Unexpected: differences between the replicated stimuli (non-repeatability).

3.2. Food products

Multivariate analyses (CVA and mrCA maps), univariate analyses and difference plots are available in Supplementary material. Conversely to controlled stimuli, no hypotheses were formulated for expected and incongruent differences for commercial food products.

All attributes confounded, all pairs of different guacamoles (six) and ice teas (six) are discriminated based on the analysis of citations by the three methods. Only TCATA discriminates the six pairs of crisps, and only TDS the six pairs of chocolates, the other methods discriminating five pairs of samples. With respect to pairs of replicated samples, the guacamoles are unexpectedly discriminated by AEF-D and AEF-A; the ice teas by TCATA, AEF-D and AEF-A; the crisps by TDS and AEF-A; and the chocolates by no method.

The analysis of durations generally leads to similar conclusions, except for TCATA (ice teas IT3/IT4 and crisps C1/C2 not discriminated by durations) and TDS (crisps C1/C3 and chocolates

CH2/CH3 not discriminated by durations). The issues of repeatability of citations with TDS and AEF-A are not observed with crisps, but with guacamoles for TCATA. The discriminating ratios are most of the time greater than one, except for pairs of guacamoles G1-G2 (TDS), ice teas IT3-IT4 (AEF-D), crisps C1-C2 (all methods), and chocolates CH2-CH4 (TCATA). These results show that whatever the method, with a few exceptions, the conclusions about differences between pairs of samples are similar based on multivariate analysis of citations or durations. Regarding the unidimensional analyses, the most consensual differences (between methods and variables) are observed for guacamoles with *Avocado*, *Onion/Shallot*, *Smooth/Fat* and *Spicy*; for ice teas with *Artificial*, *Sweet* and *Watery/Diluted*; for crisps with *Bland*, *Crackly/Hard* and *Salty*; and for chocolates with *Bitter*, *Fruity* and *Sweet*. These attributes structure the product configurations that lead to consensual multivariate differences.

Most differences are captured by several variables. However, some analyses of pairwise differences involved attributes few cited (by less than 25% of consumers, as denoted by “NT” with TDS and TCATA). This can be observed with all product categories: for chocolates with *Fluid*, for ice teas with *Bitter* and *Watery/Diluted*; for crisps with *Bland*, *Melting*, and *Sticky/Pasty*; for chocolates with *Artificial*, *Floral* and *Sour*.

Some differences are method-dependent while others are variable-dependent. AEF-D and AEF-A detect more differences of *Lemon* and *Sweet*; *Cocoa* and *Astringent*. TDS and TCATA capture more differences of *Artificial*. TDS never points out differences of *Bitter*. Differences between C1 and C2 are only detected by TCATA and AEF-D. Differences of *Tomato* are mainly shown with difference plots, differences of *Peach* or *Salty* with analysis of time-dependent variables.

Table 63 summarizes the number of significant differences observed on the food products with the unidimensional analyses of the different variables for each method. It shows that except for AEF-A, more pairs of samples are discriminated by citations than by durations. This means that analysing only durations results in a loss of discrimination, with only a marginal reduction of error rate. The single analysis of times of first citations leads to a number of discriminated pairs of samples four to five times lower than with the analysis of citations, with an error rate above 5%. However, the combined analysis of citation and durations (except for AEF-D) results in a gain of discrimination between 7 and 15%. The gain is roughly the same with the combined analysis of citations and times of first citations of TDS and TCATA, but the

error rate is higher. The only inspection of difference plots also results in a loss of discrimination for dominance-based methods, with a higher error rate. For applicability-based methods, this analysis is the most discriminative. However, the combined analysis of citations and difference plots leads to a gain of discrimination between 15 and 42% (larger for applicability-based methods). Applicability-based methods are overall more discriminative compared to dominance-based methods, but they have also higher error rates, notably with TCATA. For this method, whatever the variable analysed, the error rate is always larger than 5%.

Table 63. Count of discriminated pairs of samples, by variable, all categories combined.

		TDS	TCATA	AEF-D	AEF-A
Citation	Discrimination	88	87	98	84
	Non-repeatability	2 (2.2%)	5 (5.4%)	2 (2%)	3 (3.4%)
Duration	Discrimination	55	78	67	86
	Non-repeatability	1 (1.8%)	4 (4.9%)	1 (1.5%)	2 (2.3%)
Time	Discrimination	18	22	NT	NT
	Non-repeatability	1 (5.3%)	3 (12%)	NT	NT
Citation rate over time	Discrimination	75	119	98	115
	Non-repeatability	3 (3.8%)	10 (7.7%)	5 (4.8%)	5 (4.2%)
Citation + Duration	Discrimination	94 (+6.8%)	100 (+14.9%)	99 (+0.1%)	95 (+13.1%)
	Non-repeatability	3 (3.1%)	6 (6%)	2 (2%)	3 (3.1%)
Citation + Time	Discrimination	93 (+5.7%)	101 (+16.1%)	NT	NT
	Non-repeatability	3 (3.1%)	8 (7.3%)	NT	NT
Citation + Citation rate over time	Discrimination	104 (+18.2%)	124 (+42.5%)	113 (+15.3%)	117 (+39.3%)
	Non-repeatability	3 (2.8%)	10 (8.1%)	4 (3.4%)	5 (4.1%)

Discrimination: number of significant differences between pairs of different samples; between brackets: gain of discrimination when combined with analysis of citations. Non-repeatability: number of significant differences between pairs of replicated samples; between brackets: error rate = non-repeatability / (discrimination + non-repeatability). NT: not tested.

4. Discussion

Qualitative temporal sensory evaluation methods record sequences of citations (0/1, i.e. perception/no perception) and their occurrence times. Citations can be analysed with regard to temporal aspects (i.e. the sensory attribute is perceived from time t_1 to time t_2 , t_1 being the time of citation and (t_2-t_1) the duration of dominance/perception), or not (i.e. whether the sensory attribute is perceived or not, at any time). In this study, the multivariate analyses of citations (with mrCAs) and durations (with CVAs) lead to very similar product discrimination whatever the sensory evaluation method. This result was expected because the two variables are strongly correlated (Visalli, Béno, et al., 2023). Thus, the differences captured by multivariate analyses are carried by the main characteristics of the samples which are common according to citation and durations. This makes the multivariate analyses of

durations uninformative about the nature of the differences (durations or presence/absence) between the samples. Moreover, conversely to mrCA, CVA gives too much weight to attributes few cited, which can lead to misleading interpretation of these differences, as it can be seen on Figure 82 (with the distractors). This suggests that it might make more sense to use mrCA to multidimensionally summarize differences due to the presence/absence of attributes, then to complement with other analyses better accounting for the temporal aspects of the data.

Univariate analyses (ANOVAs) of citations and durations show that more product pairs are discriminated with citations than with durations, demonstrating that discrimination is more related to the presence/absence of sensory attributes than to their durations. This result concurs with the findings of Bommel et al. (2020) and Vidal et al. (2019) based on the comparison of CATA and TCATA. In this study, regarding product characterization, if most often the differences captured by citations include differences captured by durations or times of citations, not all the differences intersect, as observed by Alcaire et al. (2017). The results presented in Table 63 show that the combined analyses of citations with durations enable to characterize more precisely the nature of the differences between the samples compared to the sole analyses of citations or durations. This is even clearer on Figure 84, with the controlled stimuli designed to present temporal differences: TDS and TCATA enable samples to be discriminated based on differences in times of citations and durations for some attributes perceived in both samples.

In order to be measurable, univariate temporal differences between samples must be “large enough”, i.e. larger than two seconds with simple controlled stimuli. This value (the temporal resolution of TDS and TCATA) corresponds to the minimum bucket size used by Derks et al. (2022) to discretize temporal data. With complex food products, the minimum measurable temporal differences are likely to be larger. Indeed, even with controlled stimuli, some differences observed in Table 62 are not congruent with the chemical composition of the solutions. Some of these incongruent differences are common to TDS and TCATA and thus can possibly be explained by sensory interactions between attributes. However, some are likely “false positive” in view of the number of unexpected differences also observed between replicated stimuli. Thus, the gain in discrimination capacity obtained with temporal measures and notably with concurrent evaluation methods is made to the detriment of the reliability and potentially to the validity of the conclusions obtained.

Reliability issues should be considered when interpreting the results, and notably subtle differences. This is especially the case with difference plots, which represent the evolution

and the significance of differences in citation rates between pairs of samples. This analysis integrates the three temporal variables: citations, times of citations and durations. This is overall the most discriminative analysis, but also the one presenting the most incongruent and unexpected differences (Table 62 and Table 61), as well as the lowest resolution (Figure 83). This is explained by the fact that conversely to univariate analyses performed once and at individual level, proportion tests are made at the panel level and thus assume the homogeneity in temporal perception between consumers at each time point (Schlich, 2017), which has been proven to be a very strong assumption (Visalli, Mahieu, et al., 2023). To limit overinterpretation of “false positive” differences, a minimum duration corresponding to 5% of the total duration of the stimulus has been chosen in this study. However, the difference in gain of discrimination still seems overrated compared to other analyses. This could be due to large differences in times of citations or to specific perception related to clusters of consumers captured only by difference. Indeed, a difference of 10% can be significant with a proportion test when an attribute is not cited in a sample while it is cited in the other only by a small percentage of consumers. This situation reflects a low confidence in the real perception of the attribute. To avoid overinterpretation, a minimum overall citation rate can be required before testing differences in citation rates (similar to the threshold required before testing the differences in times of citations based on bootstrap tests, set to 25% in this study). This threshold is difficult to establish (Visalli, Mahieu, et al., 2023). A candidate could be the significance level of the TDS curves, however this threshold is contested (Meyners & Castura, 2019). Indeed, as observed, some attributes can be either specific to some samples or never used, meaning that for some products the threshold used in TDS curves could be underestimated.

The issue with low citation rates observed is not specific to TDS or dominance-based methods but common to all static and temporal qualitative analyses using citation rates to determine whether an attribute has been perceived or not. This topic deserves further investigation, especially since no recommendation exists about the appropriate number of consumers to include in studies involving temporal methods. Conversely, with applicability-based methods and notably TCATA with which citation rates are higher, it is difficult to separate what falls under the agreement and the acquiescence bias observed with CATA (Mahieu et al., 2020a) and CATA-like protocols (Callegaro et al., 2015; Kim et al., 2017). Other works tend to confirm that applicability-based methods overestimate citation rates (Visalli, Béno, et al., 2023). This could make the analyses more sensitive and increase the number of false positives with complex food products as seen in this article (Table 61). The question of the reliability and

validity of conclusions based on citation rates is not new and has been studied with CATA (Ares, Antúnez, et al., 2014; Jaeger et al., 2013), but it remains actual with temporal methods (Velázquez et al., 2020; Visalli, Béno, Nicolle, & Schlich, 2023).

The size of the difference among samples has previously been reported to impact reproducibility of the result (Ares, Antúnez, et al., 2014). In this study, the opposite effect is observed as the product category presenting the less differences and the most complexity (Table 60) is also the one for which the less reliability issues are observed. However, the two characteristics were subjectively evaluated by the consumers, thus it would be preferable to estimate the size of differences based on collected data. Adding replicated samples allows to quantify and compare the size of differences between replicated and different pairs of samples. This is the idea underlying the discriminating ratio computed for the difference plots. A ratio lower or equal to one clearly indicates potential reliability issues due to too much similarity between samples.

It is interesting to notice that on controlled stimuli (Table 62), with a low number of attributes to follow, TCATA has less reliability issues compared to TDS, while the opposite is observed on real food products (Table 61). This suggests that the reliability of TCATA decreases with the number of attributes available to consumers or the complexity of the product, probably due to the increased cognitive load caused by the need to focus and the difficulty to react on multiple events temporally close. Surprisingly, AEF-A does not suffer the same issue. If the method is clearly not appropriate for measuring small temporal differences, it presents a fairly good compromise between discrimination (better than TDS but worse than TCATA) and reliability (better than TCATA but worse than TDS) on real food products. This is probably explained by the fact that the more complex is the product space, the more the consumers focus on differences of applicability rather than on temporal differences. In this sense, AEF-A, as a retrospective method, probably collects data very comparable to those that would have been collected with CATA.

The temporal differences captured with AEF-A with durations or difference plots but not with citations rarely align with the temporal differences captured by TDS or TCATA. This result can either suggest that these differences are reliable but non-valid on real food products (no such issue is observed on controlled stimuli), or confirm that retrospective measures result from a different evaluation process than concurrent measures (read the discussion about Kahneman's theory in Visalli, Wakihira, et al., 2022). More than discrimination, this question is probably the most interesting to investigate as it is relative to the ecological validity of the

measures which is sought after with consumer studies (Galiñanes Plaza et al., 2019). According to this criterion, TCATA and probably even more TDS have an undeniable advantage as they are closest to natural immediate perception.

5. Conclusion

As demonstrated with controlled temporal stimuli, when temporal differences are established, concurrent evaluation methods (TDS and TCATA) undoubtedly perform better for product characterization than retrospective evaluation methods (AEF-D and AEF-A). These latter can be used by consumers to determine the main differences between samples, but due to their low temporal resolution they are obviously not appropriate to capture small temporal differences. However, with the real food products tested presenting unknown temporal differences, AEF-A, in practice very similar to CATA, discriminated samples as well as TDS and TCATA. Indeed, the main differences between samples could be explained by the presence or absence of sensory attributes regardless of temporal aspects. This result questions about the real interest of temporal measures with consumers for answering most product development related questions. When discrimination is sought, using static or low temporal resolution measures is unlikely to make any difference. Yet, in some cases, analysing time-dependent variables (times of citations and durations) in complement of citations can improve discrimination and bring complementary insights about the nature of the differences. As there is no possibility to anticipate if temporal differences between samples exist before having testing for them, it could therefore be tempting to systematically use temporal methods to avoid missing potential additional information. Moreover, as perception is a dynamic phenomenon, concurrent evaluation methods are by nature more ecologically valid than retrospective ones, more cognitive. Nonetheless, temporal sensory data are complex to analyse and interpret due to the richness of information, but also of the noise, that they contain. There is no ultimate method, and every gain in temporal resolution or discrimination is assumably at the cost of a loss of reliability. This study shows how difficult it is to generalize conclusions relating to sensory evaluation methods. In the regrettable absence of large open databases of sensory studies that would allow meta-analyses to be conducted, the determination of the “best” method is likely to remain an endless debate.

Supplementary material

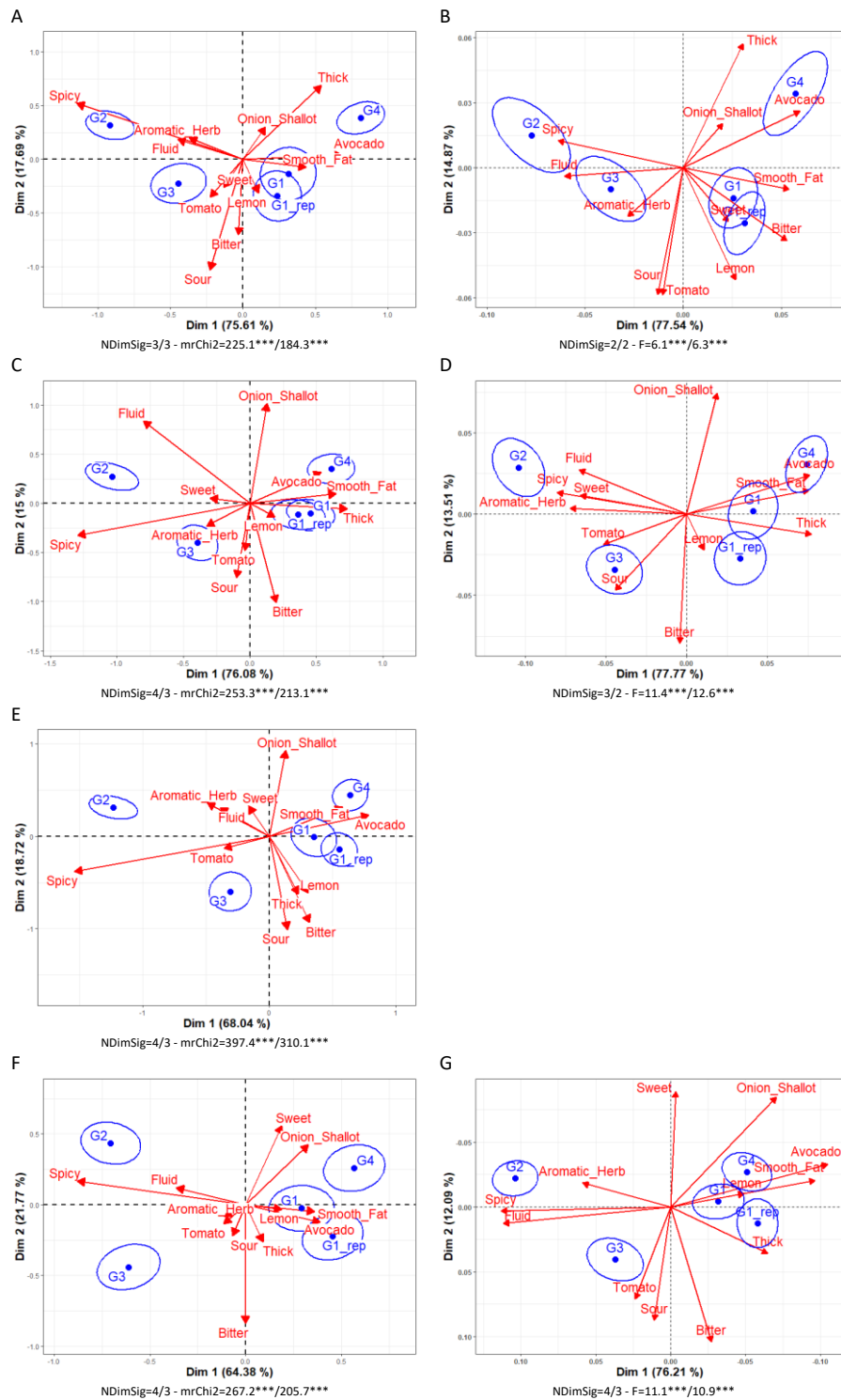


Figure 85. Guacamoles - Biplots of multivariate analyses (axes 1-2), with 95% confidence ellipses. First row: TDS, second row: TCATA, third row: AEF-D, fourth row: AEF-A. First column: mrCA of citations, second column: CVA of durations. NDimSig: number of significant dimensions (all stimuli/without replicates). mrChi2: mrChi2 statistic (all stimuli/without replicates). F: MANOVA statistic (all stimuli/without replicates). ***p-value<0.001. CVAs of durations for AEF-D were not produced due to too much collinearity in durations.

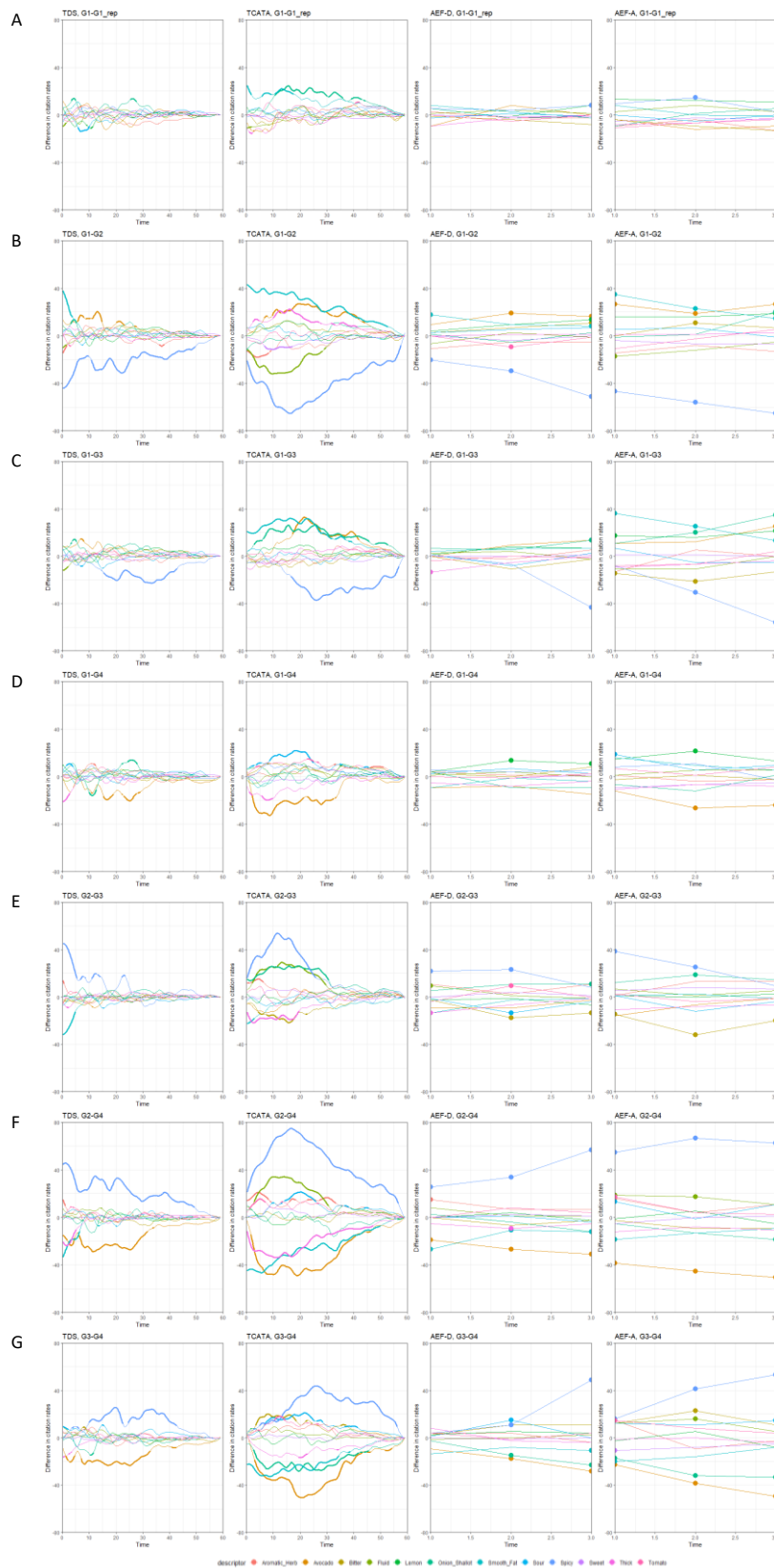


Figure 86. Guacamoles - Pairwise differences in citation rates (Y-axis) plotted over time (X-axis). Rows of plots each refers to one pair comparison. Within each row, first column refers to TDS, second to TCATA, third to AEF-D, fourth to AEF-A. Bold lines (for TDS and TCATA) and points (for AEF-D and AEF-A) correspond to significant differences in citation rates ($\alpha=0.05$).

A - Guacamoles		G1-G1_rep				G1-G2				G1-G3				G1-G4				G2-G3				G2-G4				G3-G4							
		TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A
Aromatic_Herbs	Citation	0.2	-0.4	0.1	-0.7	-1.3	-0.8	-1.3	-0.7	-0.1	-0.6	0.4	-0.7	0.7	0.4	0.7	-0.1	1.2	0.2	1.6	0.0	2.0	1.2	2.0	0.6	0.8	1.0	0.3	0.6	0.8	1.0	0.3	0.6
	Duration	-1.1	0.6	-0.1	-0.4	-0.9	-3.2	-0.8	-0.5	-0.3	0.1	0.2	-0.2	0.0	3.1	0.5	-0.1	0.6	3.2	1.0	0.4	1.0	6.3	1.3	0.4	0.3	3.1	0.3	0.0	0.3	3.1	0.3	0.0
	Time	-2.0	0.2	NT	NT	3.1	8.4	NT	NT	-1.2	0.8	NT	NT	-1.8	1	NT	NT	-4.2	-7.3	NT	NT	-4.6	-7.4	NT	NT	-0.4	0.2	NT	NT	-0.4	0.2	NT	NT
	Citation rate over time	1.4	0.0	0.0	0.0	4.0	8.8	0.0	0.0	0.0	0.0	0.0	0.0	1.6	6.2	0.0	0.0	0.8	5.8	1.0	0.0	1.0	22.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Avocado	Citation	1.0	0.1	0.2	-0.4	3.5	1.9	2.7	2.4	2.5	1.2	1.2	1.3	-1.3	-1.5	-0.6	-1.1	-1.0	-0.7	-1.5	-1.1	-4.8	-3.3	-3.4	-3.5	-3.8	-2.7	-1.9	-2.4	-3.8	-2.7	-1.9	-2.4
	Duration	0.0	-0.4	0.0	-0.1	4.0	8.6	1.4	0.6	3.0	7.4	0.5	0.4	-4.7	-8.2	-0.4	-0.3	-0.9	-1.2	-0.9	-0.2	-8.7	-16.8	-1.8	-0.9	-7.8	-15.6	-0.9	-0.7	-7.8	-15.6	-0.9	-0.7
	Time	-0.6	2.9	NT	NT	0.2	-0.6	NT	NT	0.3	2.2	NT	NT	0.9	3.4	NT	NT	0.0	2.8	NT	NT	0.8	4.2	NT	NT	0.6	1.4	NT	NT	0.6	1.4	NT	NT
	Citation rate over time	0.2	1.2	0.0	0.0	14.4	34.6	2.0	3.0	1.4	24.8	1.0	1.0	11.6	30.4	0.0	2.0	0.4	0.0	0.0	0.0	31.2	44.0	3.0	3.0	31.4	45.6	2.0	3.0	31.4	45.6	2.0	3.0
Bitter	Citation	-0.2	-0.3	-0.4	-1.0	0.8	1.3	2.0	1.4	-0.2	-0.4	-0.4	-1.7	1.0	0.6	0.8	0.1	-1.0	-1.7	-2.4	-3.1	0.3	-0.6	-1.2	-1.3	1.2	1.1	1.2	1.9	1.2	1.1	1.2	1.9
	Duration	-0.8	-1.3	-0.4	-0.6	1.3	2.7	1.6	0.8	0.4	-2.0	-0.4	-0.9	0.1	3.6	0.5	-0.1	-0.9	-4.7	-2.0	-1.7	-1.3	0.9	-1.0	-0.9	-0.4	5.6	1.0	0.8	-0.4	5.6	1.0	0.8
	Time	2.4	2.8	NT	NT	NT	NT	NT	NT	5.2	3.6	NT	NT	0.0	-2.3	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.2	0.4	1.0	1.0	0.0	0.0	0.0	2.0	0.0	0.6	0.0	0.0	0.0	14.8	2.0	3.0	0.0	0.0	0.0	0.0	1.4	19.2	0.0	1.0	1.4	19.2	0.0	1.0
Fluid	Citation	-0.3	-0.6	1.8	0.8	-1.5	-3.2	-0.9	-1.2	-0.4	-0.8	0.9	-0.5	0.4	-0.4	1.3	1.0	1.1	2.4	1.8	0.6	1.9	2.8	2.2	2.1	0.8	0.4	0.4	1.5	0.8	0.4	0.4	1.5
	Duration	0.4	-1.7	1.3	0.4	-0.5	-7.4	-0.5	-0.6	-0.4	-0.3	0.6	-0.4	0.5	0.7	0.8	0.3	0.0	7.1	1.0	0.2	1.0	8.1	1.3	0.9	0.9	1.0	0.3	0.7	0.9	1.0	0.3	0.7
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	1.6	0.0	0.0	0.0	1.4	29.2	0.0	1.0	2.2	0.0	0.0	0.0	0.0	0.6	0.0	0.0	0.0	27.6	1.0	0.0	0.6	28.0	0.0	2.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	1.0
Lemon	Citation	-0.4	0.4	0.4	0.3	0.5	1.0	2.4	1.1	0.8	0.7	1.2	1.3	0.5	0.5	2.4	0.8	0.3	-0.3	-1.1	0.2	0.0	-0.4	0.0	-0.3	-0.3	-0.2	1.1	-0.5	-0.3	-0.2	1.1	-0.5
	Duration	-0.8	0.6	0.2	0.4	1.1	1.6	1.1	0.6	1.0	2.2	0.5	0.6	1.1	2.7	1.2	0.6	0.1	0.6	-0.6	0.1	0.1	1.1	0.1	0.0	0.1	0.5	0.7	-0.1	0.1	0.5	0.7	-0.1
	Time	-0.4	0.0	NT	NT	-4.2	-0.5	NT	NT	-0.8	-0.1	NT	NT	-2.4	-1.2	NT	NT	-3.4	-0.3	NT	NT	1.7	-1.3	NT	NT	-1.4	-1.0	NT	NT	-1.4	-1.0	NT	NT
	Citation rate over time	1.4	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0	0.0	2.0	0.4	1.2	2.0	1.0	0.0	0.0	0.0	0.0	0.2	0.4	0.0	0.0	0.0	0.6	0.0	0.0	0.0	0.6	0.0	0.0
Onion_Shallot	Citation	0.8	0.7	0.2	0.2	0.5	0.0	-0.1	0.5	1.6	1.9	1.5	1.8	0.3	-0.2	-0.5	-0.8	1.0	1.9	1.6	1.2	-0.3	-0.2	-0.4	-1.4	-1.3	-2.1	-2.0	-2.6	-1.3	-2.1	-2.0	-2.6
	Duration	2.4	7.5	0.0	0.0	1.4	0.9	0.0	0.2	2.6	9.6	1.0	0.7	1.0	-0.5	-0.3	-0.1	1.2	8.7	1.0	0.5	-0.4	-1.4	-0.3	-0.3	-1.7	-10.1	-1.3	-0.8	-1.7	-10.1	-1.3	-0.8
	Time	-2.4	0.6	NT	NT	-1.9	0.8	NT	NT	-3.8	-2.2	NT	NT	-1.2	1.5	NT	NT	-2	-3.0	NT	NT	0.9	0.4	NT	NT	2.5	3.8	NT	NT	2.5	3.8	NT	NT
	Citation rate over time	4.2	30.4	0.0	0.0	2.2	0.4	0.0	1.0	2.2	47.8	1.0	2.0	7.6	0.0	0.0	0.0	0.0	30.6	1.0	1.0	0.2	0.2	0.0	1.0	1.8	42.8	2.0	3.0	1.8	42.8	2.0	3.0
Smooth_Fat	Citation	0.6	1.7	0.9	0.5	1.9	3.6	2.2	2.2	0.8	2.7	1.0	2.4	-0.4	0.6	-0.6	1.1	-1.1	-0.9	-1.2	0.2	-2.3	-3.1	-2.8	-1.0	-1.2	-2.1	-1.6	-1.2	-1.2	-2.1	-1.6	-1.2
	Duration	1.0	7.0	0.2	0.0	2.2	14.0	1.1	0.7	1.0	10.5	0.4	0.7	0.4	0.3	-0.2	0.2	-1.2	-3.5	-0.7	0.0	-1.8	-13.6	-1.3	-0.5	-0.6	-10.2	-0.7	-0.5	-0.6	-10.2	-0.7	-0.5
	Time	0.0	0.0	NT	NT	-6.6	-1.6	NT	NT	0.0	0.0	NT	NT	0.0	0.0	NT	NT	6.6	1.6	NT	NT	6.4	1.6	NT	NT	0.0	0.0	NT	NT	0.0	0.0	NT	NT
	Citation rate over time	0.6	13.6	0.0	0.0	5.0	52.8	2.0	3.0	0.0	38.0	0.0	3.0	0.0	0.0	0.0	0.0	4.8	3.2	0.0	0.0	7.2	49.8	3.0	1.0	0.0	35.4	1.0	1.0	0.0	35.4	1.0	1.0
Sour	Citation	-0.3	0.7	0.0	0.1	0.6	0.8	2.2	0.4	-0.6	0.5	-0.5	0.0	2.0	2.1	2.2	1.1	-1.2	-0.3	-2.7	-0.4	1.4	1.3	0.0	0.7	2.6	1.6	2.7	1.1	2.6	1.6	2.7	1.1
	Duration	-0.7	1.2	-0.1	-0.1	0.5	0.8	1.5	0.2	-0.8	0.1	-0.2	-0.1	1.1	6.0	1.0	0.6	-1.3	-0.8	-1.7	-0.2	0.6	5.2	-0.5	0.5	1.9	6.0	1.2	0.7	1.9	6.0	1.2	0.7
	Time	0.0	0.7	NT	NT	-2.6	-1.4	NT	NT	-2.6	1.6	NT	NT	NT	NT	NT	NT	0.0	2.8	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	4.6	0.0	0.0	0.0	0.4	0.0	0.0	0.0	1.6	0.0	0.0	0.0	2.8	24.0	0.0	1.0	0.6	0.0	1.0	0.0	0.6	17.8	0.0	1.0	3.8	27.0	1.0	1.0	3.8	27.0	1.0	1.0
Spicy	Citation	0.8	0.6	3.5	1.1	-8.3	-5.7	-4.8	-6.3	-3.0	-3.6	-2.8	-3.7	1.4	0.9	2.3	0.9	5.3	2.1	2.0	2.6	9.7	6.6	7.1	7.1	4.4	4.5	5.1	4.5	4.4	4.5	5.1	4.5
	Duration	0.8	1.5	2.0	0.6	-10.9	-23.5	-1.8	-1.3	-6.3	-11.9	-1.2	-0.9	0.8	2.8	1.3	0.3	4.6	11.6	0.6	0.4	11.7	26.3	3.1	1.6	7.1	14.7	2.5	1.2	7.1	14.7	2.5	1.2
	Time	NT	NT	NT	NT	7.1	3.7	NT	NT	-4.8	-6.2	NT	NT	NT	NT	NT	NT	-11.8	-10.0	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	0.2	0.0	1.0	1.0	49.0	57.6	3.0	3.0	28.2	40.4	1.0	2.0	0.0	1.0	0.0	0.0	12.2	26.2	2.0	2.0	49.4	56.6	3.0	3.0	33.4	52.2	2.0	3.0	33.4	52.2	2.0	3.0
Sweet	Citation	-0.8	-0.8	-1.5	-1.7	-0.5	-1.6	-2.0	-1.9	-1.2	-1.3	18.0	1.9	0.0	-0.6	0.0	-2.7	-0.7	0.4	20.0	3.7	0.5	1.0	2.0	-0.8	1.2	0.7	-18.0	-4.5	1.2	0.7	-18.0	-4.5
	Duration	0.4	-1.3	-1.1	-1.2	0.5	-2.6	-1.4	-1.4	0.2	-1.6	18.3	0.7	0.3	-1.0	0.0	-1.6	-0.3	1.0	19.7	2.1	-0.2	1.6	1.4	-0.2	0.1	0.6	-18.3	-2.3	0.1	0.6	-18.3	-2.3
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	10.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Thick	Citation	0.1	-0.2	-1.3	-0.6	0.8	2.3	0.5	0.4	0.8	0.6	-2.0	-0.7	-1.8	-0.6	-1.3	-0.6	0.0	-1.7	-2.5	-1.1	-2.7	-3.0	-1.8	-1.0	-2.7	-1.3	0.6	0.1	-2.7	-1.3	0.6	0.1
	Duration	0.2	-2.2	-0.9	-0.3	-0.2	6.0	0.4	0.1	0.2	1.3	-1.2	-0.4	-1.3	-3.9	-1.2	-0.5	0.3	-4.8	-1.6	-0.5	-1.2	-9.9	-1.6	-0.6	-1.5	-5.2	0.1	-0.1	-1.5	-5.2	0.1	-0.1
	Time	3.6	4.8	NT	NT	NT	NT	NT	NT	NT	5.7	NT	NT	4.4	4.4	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	-1.4	NT	NT	NT	-1.4	NT	NT
	Citation rate over time	0.0	1.6	0.0	0.0	0.2	29.2	0.0	0.0	0.0	0.2	1.0	0.0	3.0	6.4	0.0	0.0	1.0	19.6	1.0	0.0	5.2	47.2	1.0	0.0	2.2	5.2	0.0	0.0	2.2	5.2	0.0	0.0

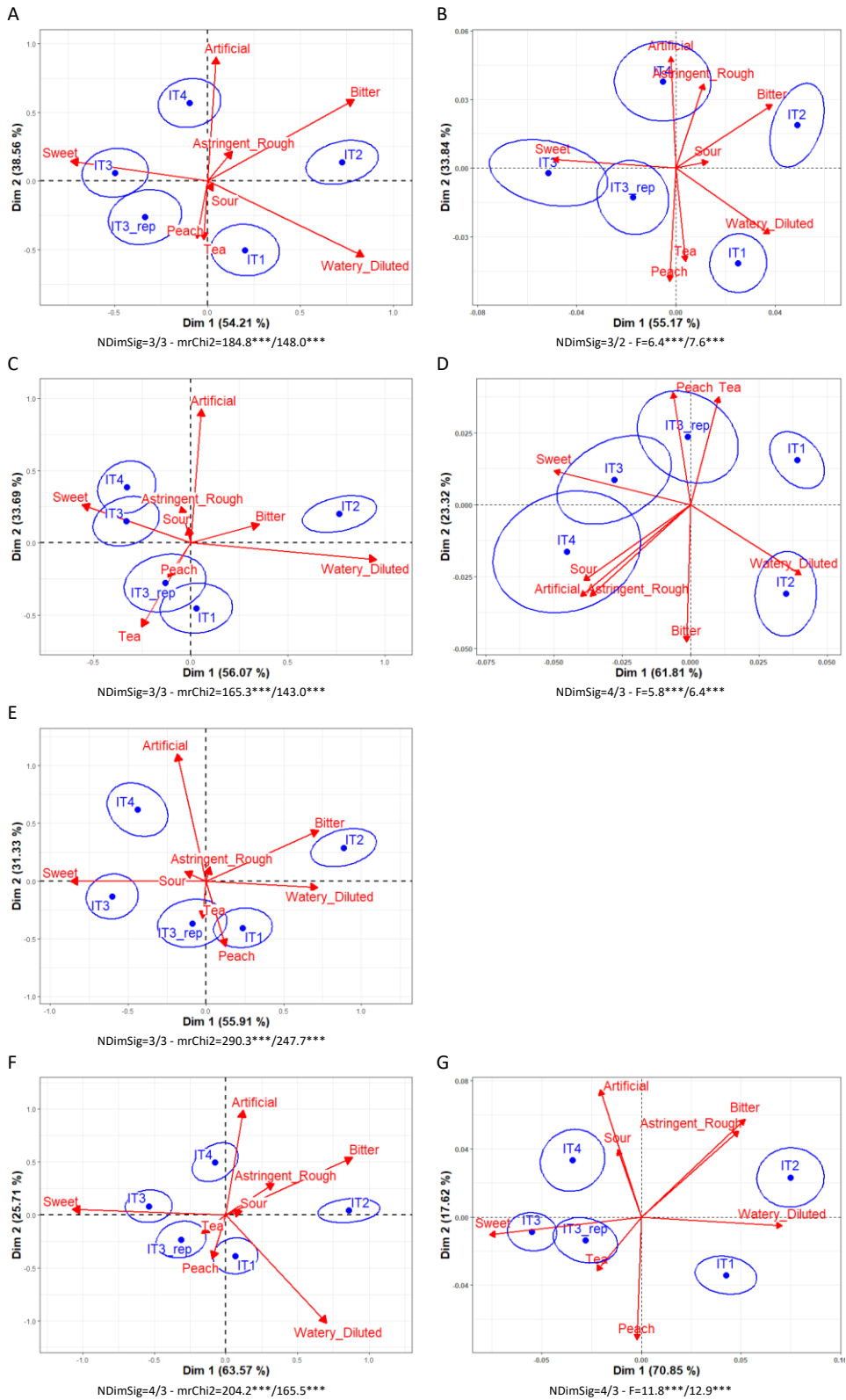


Figure 88. Ice teas - Biplots of multivariate analyses (axes 1-2), with 95% confidence ellipses. First row: TDS, second row: TCATA, third row: AEF-D, fourth row: AEF-A. First column: mrCA of citations, second column: CVA of durations. NDimsig: number of significant dimensions (all stimuli/without replicates). mrChi2: mrChi2 statistic (all stimuli/without replicates). F: MANOVA statistic (all stimuli/without replicates). ***p-value<0.001. CVAs of durations for AEF-D were not produced due to too much collinearity in durations.

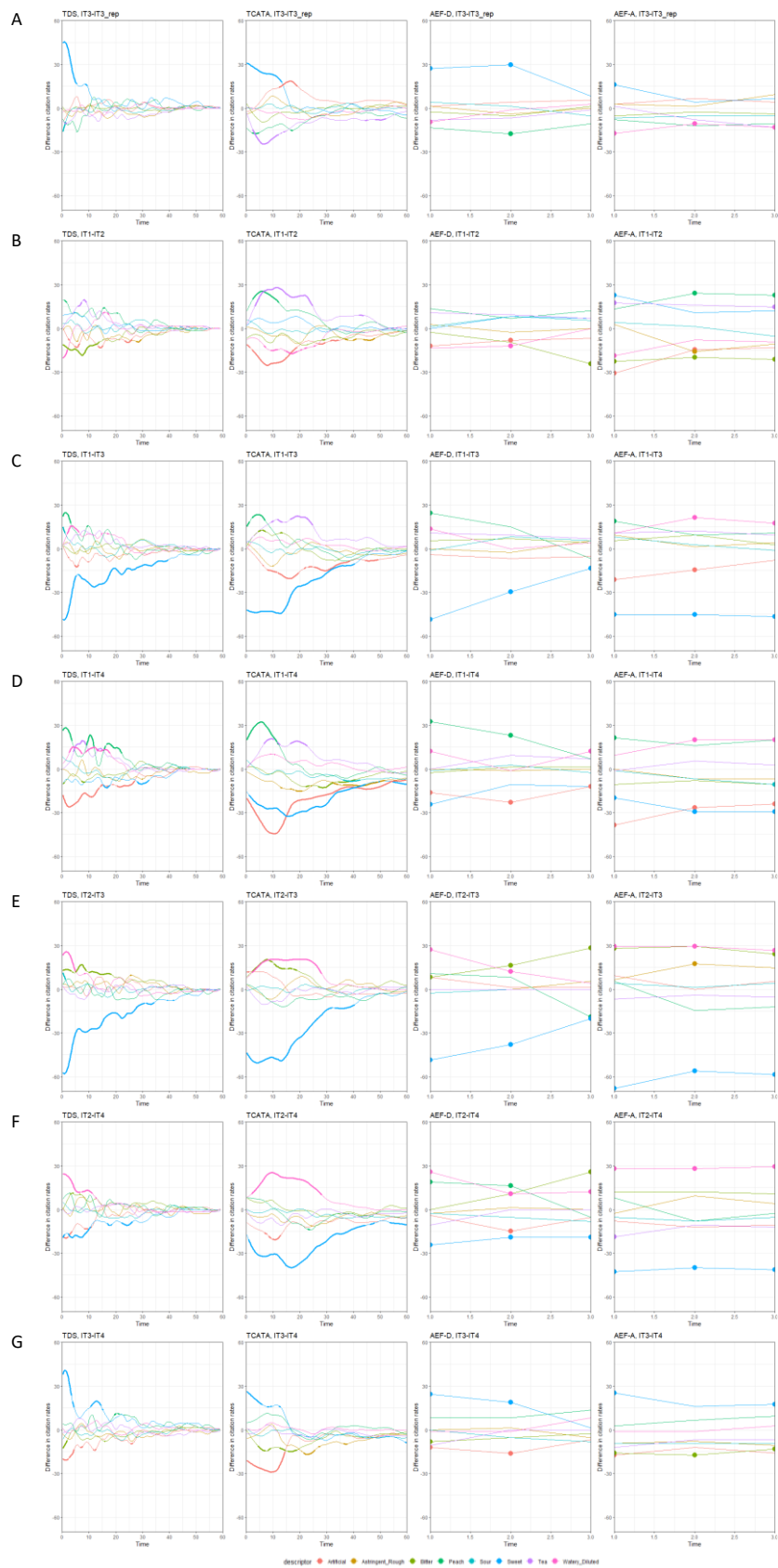


Figure 89. Ice teas - Pairwise differences in citation rates (Y-axis) plotted over time (X-axis). Rows of plots each refers to one pair comparison. Within each row, first column refers to TDS, second to TCATA, third to AEF-D, fourth to AEF-A. Bold lines (for TDS and TCATA) and points (for AEF-D and AEF-A) correspond to significant differences in citation rates ($\alpha=0.05$).

B - Ice teas		IT3-IT3_rep				IT1-IT2				IT1-IT3				IT1-IT4				IT2-IT3				IT2-IT4				IT3-IT4					
		TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA
Artificial	Citation	1.0	1.4	0.9	0.2	-1.3	-2.1	-2.5	-0.9	-1.1	-1.7	-1.8	-0.8	-3.0	-2.6	-4.0	-1.2	0.3	0.4	0.6	0.2	-1.7	-0.5	-1.5	-0.3	-1.9	-0.9	-2.2	-0.4		
	Duration	0.7	5.6	0.7	0.2	-2.7	-6.6	-1.8	-0.9	-1.9	-7.9	-1.4	-0.8	-4.9	-15.0	-2.4	-1.2	0.8	-1.3	0.4	0.2	-2.2	-8.4	-0.6	-0.3	-3.0	-7.1	-1.0	-0.4		
	Time	3.2	2.5	NT	NT	3.6	NT	NT	NT	1.4	NT	NT	NT	5.4	NT	NT	NT	-1.0	-2.0	NT	NT	1.8	1.5	NT	NT	2.8	3.4	NT	NT		
	Citation rate over time	0.0	7.2	0.0	0.0	7.8	25.0	2.0	2.0	3.2	34.4	0.0	2.0	23.6	59.6	3.0	3.0	0.4	1.6	0.0	0.0	6.2	5.4	1.0	0.0	7.4	14.4	2.0	1.0		
Astringent_Rough	Citation	-0.4	-0.7	-0.4	0.4	-0.3	-0.5	-0.2	-0.3	0.5	0.0	0.2	0.3	-0.8	-1.5	-0.4	-0.2	0.8	0.5	0.4	0.6	-0.5	-1.0	-0.2	0.1	-1.3	-1.5	-0.6	-0.5		
	Duration	-1.2	0.6	-0.1	0.4	-1.8	-2.7	0.0	-0.3	-0.4	-2.0	0.1	0.3	-1.7	-6.9	0.0	-0.2	1.4	0.7	0.1	0.6	0.1	-4.2	0.0	0.1	-1.2	-4.9	-0.2	-0.5		
	Time	1.6	-4.2	NT	NT	2.4	3.6	NT	NT	-1.2	7.8	NT	NT	0.4	3.6	NT	NT	-4.0	4.6	NT	NT	-3.2	-0.4	NT	NT	0.7	-5.0	NT	NT		
	Citation rate over time	0.0	0.0	0.0	0.0	3.0	7.2	0.0	1.0	0.0	0.4	0.0	0.0	0.6	24.4	0.0	0.0	1.0	0.0	0.0	1.0	0.0	0.6	0.0	0.0	0.2	7.6	0.0	0.0		
Bitter	Citation	-0.2	-0.7	-1.1	-0.8	-2.7	-0.9	-2.1	-1.2	0.5	1.0	1.7	1.0	-1.3	-0.2	0.1	-0.7	3.2	2.0	3.9	2.2	1.3	0.8	2.3	0.5	-1.9	-1.2	-1.6	-1.7		
	Duration	-0.1	-0.7	-0.8	-0.8	-3.0	-4.2	-1.0	-1.2	0.5	0.7	1.4	1.0	-0.9	-4.0	0.0	-0.7	3.6	5.0	2.4	2.2	2.1	0.3	1.0	0.5	-1.5	-4.7	-1.4	-1.7		
	Time	NT	NT	NT	NT	NT	-1.0	NT	NT	NT	NT	NT	NT	NT	-0.8	NT	NT	NT	NT	NT	NT	0.2	0.1	NT	NT	NT	NT	NT	NT		
	Citation rate over time	0.0	0.2	0.0	0.0	14.2	1.8	1.0	3.0	0.0	4.6	0.0	0.0	2.2	14.2	0.0	0.0	17.0	15.8	3.0	3.0	2.0	0.0	1.0	0.0	2.4	10.4	1.0	3.0		
Peach	Citation	-1.0	-0.6	-1.2	-0.2	1.3	1.4	0.9	0.4	1.0	0.7	0.8	0.2	2.1	0.9	2.1	0.3	-0.3	-0.7	-0.1	-0.1	0.7	-0.5	1.2	0.0	1.1	0.2	1.2	0.1		
	Duration	-0.1	-6.8	-0.4	-0.2	2.8	2.8	0.3	0.4	1.5	0.7	0.3	0.2	3.9	0.5	0.8	0.3	-1.2	-2.1	0.0	-0.1	1.2	-2.3	0.5	0.0	2.4	-0.2	0.5	0.1		
	Time	2.6	2.4	NT	NT	-1.9	-1.3	NT	NT	-3.8	-2.6	NT	NT	-2.6	-5.2	NT	NT	-1.7	-1.4	NT	NT	-0.8	-3.8	NT	NT	0.8	-2.3	NT	NT		
	Citation rate over time	1.2	3.8	1.0	0.0	3.2	10.2	0.0	2.0	4.2	6.2	1.0	1.0	13.6	12.0	2.0	2.0	0.0	0.0	1.0	0.0	0.0	1.0	2.0	0.0	2.6	0.0	0.0	0.0		
Sour	Citation	-0.4	-0.4	0.0	-0.7	0.8	-0.1	2.2	0.0	1.0	0.3	1.6	0.4	0.2	-0.6	0.0	-0.5	0.1	0.4	-0.6	0.4	-0.6	-0.5	-2.2	-0.5	-0.7	-0.9	-1.6	-1.0		
	Duration	-0.4	0.0	0.0	-0.7	0.6	-1.3	1.5	0.0	0.8	-1.1	1.0	0.4	-0.2	-4.0	-0.1	-0.5	0.2	0.2	-0.5	0.4	-0.8	-2.8	-1.6	-0.5	-1.0	-2.9	-1.1	-1.0		
	Time	NT	NT	NT	NT	NT	-2.6	NT	NT	NT	NT	NT	NT	-3.1	-2.4	NT	NT	NT	NT	NT	NT	NT	0.0	NT	NT	NT	NT	NT	NT		
	Citation rate over time	1.2	0.0	0.0	0.0	0.2	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	6.8	0.0	1.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		
Sweet	Citation	1.5	2.1	2.3	0.1	0.7	0.6	1.0	0.7	-4.2	-3.4	-3.7	-0.9	-1.8	-2.1	-1.9	-0.6	-4.9	-4.0	-4.7	-1.6	-2.5	-2.7	-2.8	-1.3	2.4	1.4	1.8	0.3		
	Duration	4.0	2.6	0.8	0.1	1.0	1.2	0.7	0.7	-8.1	-13.4	-1.4	-0.9	-3.2	-15.0	-1.0	-0.6	-9.1	-14.6	-2.1	-1.6	-4.2	-16.2	-1.6	-1.3	4.9	-1.5	0.5	0.3		
	Time	-2.0	-1.2	NT	NT	-0.8	0.2	NT	NT	0.0	1.1	NT	NT	0.0	0.2	NT	NT	1.2	0.0	NT	NT	0.8	-0.1	NT	NT	0.0	-1.0	NT	NT		
	Citation rate over time	6.8	13.4	2.0	1.0	0.6	0.0	0.0	1.0	39.0	40.4	3.0	3.0	7.2	49.8	1.0	3.0	35.2	39.8	3.0	3.0	15.8	52.6	3.0	3.0	8.8	9.4	2.0	2.0		
Tea	Citation	-1.3	-1.6	-0.8	-0.2	1.5	3.2	1.4	0.5	1.6	2.2	1.2	0.3	1.4	1.7	0.9	0.1	0.1	-1.0	-0.2	-0.2	-0.1	-1.5	-0.5	-0.5	-0.2	-0.5	-0.3	-0.3		
	Duration	-2.1	-6.1	-0.5	-0.2	2.2	6.8	0.7	0.5	2.2	6.2	0.7	0.3	2.1	3.6	0.4	0.1	-0.1	-0.6	0.0	-0.2	-0.2	-3.2	-0.3	-0.5	-0.1	-2.6	-0.3	-0.3		
	Time	1.8	1.2	NT	NT	-0.1	-0.8	NT	NT	-0.8	0.7	NT	NT	-2.4	-1.3	NT	NT	-0.6	1.6	NT	NT	-2.1	-0.5	NT	NT	-1.7	-2.2	NT	NT		
	Citation rate over time	1.0	16.6	0.0	0.0	3.8	23.8	0.0	2.0	0.8	12.8	0.0	0.0	7.0	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0		
Watery_Diluted	Citation	-0.7	-0.9	-0.8	-1.2	-0.3	-2.1	-1.5	-0.4	2.6	1.1	1.1	1.3	2.7	2.1	1.9	1.3	2.9	3.2	2.6	1.7	3.0	4.2	3.4	1.7	0.2	1.0	0.8	0.0		
	Duration	-0.1	-1.7	-0.5	-1.2	0.4	-3.9	-0.6	-0.4	1.9	1.8	0.8	1.3	2.4	2.0	1.3	1.3	1.5	5.6	1.4	1.7	2.0	5.9	1.9	1.7	0.5	0.3	0.5	0.0		
	Time	NT	NT	NT	NT	4.4	-2.4	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT		
	Citation rate over time	0.0	0.0	1.0	3.0	3.6	11.8	1.0	1.0	7.2	0.0	1.0	2.0	14.2	1.0	2.0	2.0	6.6	26.2	2.0	3.0	11.4	26.4	3.0	3.0	0.0	0.0	0.0	0.0		
Multivariate	mrChi2	7.4	12.9	14.9	14.3	22.9	32.4	39.7	35.9	31.2	22.0	44.5	28.8	41.5	33.4	53.7	33.0	47.2	40.2	76.0	74.5	30.2	41.4	66.3	47.7	17.9	11.5	27.9	21.5		
	F-MANOVA	1.7	2.2	NT	3.4	5.3	4.3	NT	6.3	6.3	6.5	NT	15.2	9.3	5.9	NT	13.4	9.4	8.7	NT	22.7	4.9	5.5	NT	13.8	4.3	1.4	NT	3.9		
	Discriminating ratio	1.0	1.0	1.0	1.0	1.3	1.1	1.1	2.0	1.6	1.3	1.4	1.9	1.7	1.9	1.2	1.9	1.7	1.3	1.5	2.6	1.3	1.5	1.4	2.1	1.3	1.1	1.0	1.4		

Figure 90. Ice teas - Summary of pairwise differences, by attribute (rows) and method (columns). Citation: log of the odd ratio (ANOVA citation = sample + consumer + error). Duration: mean duration (ANOVA duration = sample + consumer + error). Time: median time of citation (bootstrap test). Citation rate over time: total duration of significant differences in citation rates (difference plot). mrChi2: mrChi2 statistic restricted to the pair of samples (mrCA of citations). F-MANOVA: F statistic of the MANOVA restricted to the pair of samples (CVA of durations). Discriminating ratio: average difference between the pair of samples divided by average difference between the replicated samples (citation rates over time). Orange and red cells: significant unexpected differences between replicated samples (red cells: unexpected differences not captured by citations). Light green and dark green cells: significant differences (light green cells: differences not captured by citations). NT: discrimination not tested. Due to a technical problem, only nine attributes over twelve have been presented to the AEF-A panel on the chocolate product category.

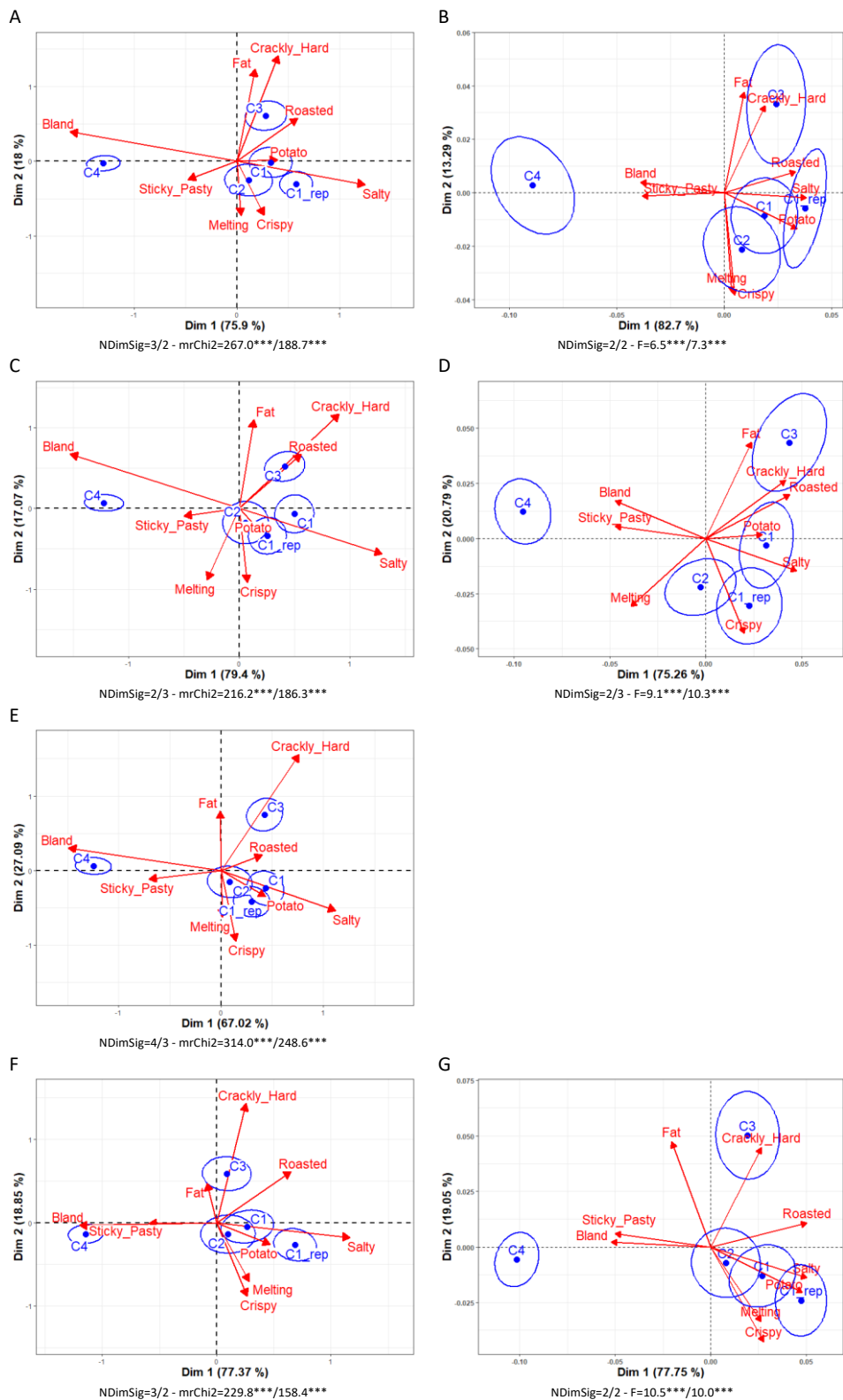


Figure 91. Crisps - Biplots of multivariate analyses (axes 1-2), with 95% confidence ellipses. First row: TDS, second row: TCATA, third row: AEF-D, fourth row: AEF-A. First column: mrCA of citations, second column: CVA of durations. NDimSig: number of significant dimensions (all stimuli/without replicates). mrChi2: mrChi2 statistic (all stimuli/without replicates). F: MANOVA statistic (all stimuli/without replicates). ***p-value<0.001. CVAs of durations for AEF-D were not produced due to too much collinearity in durations.

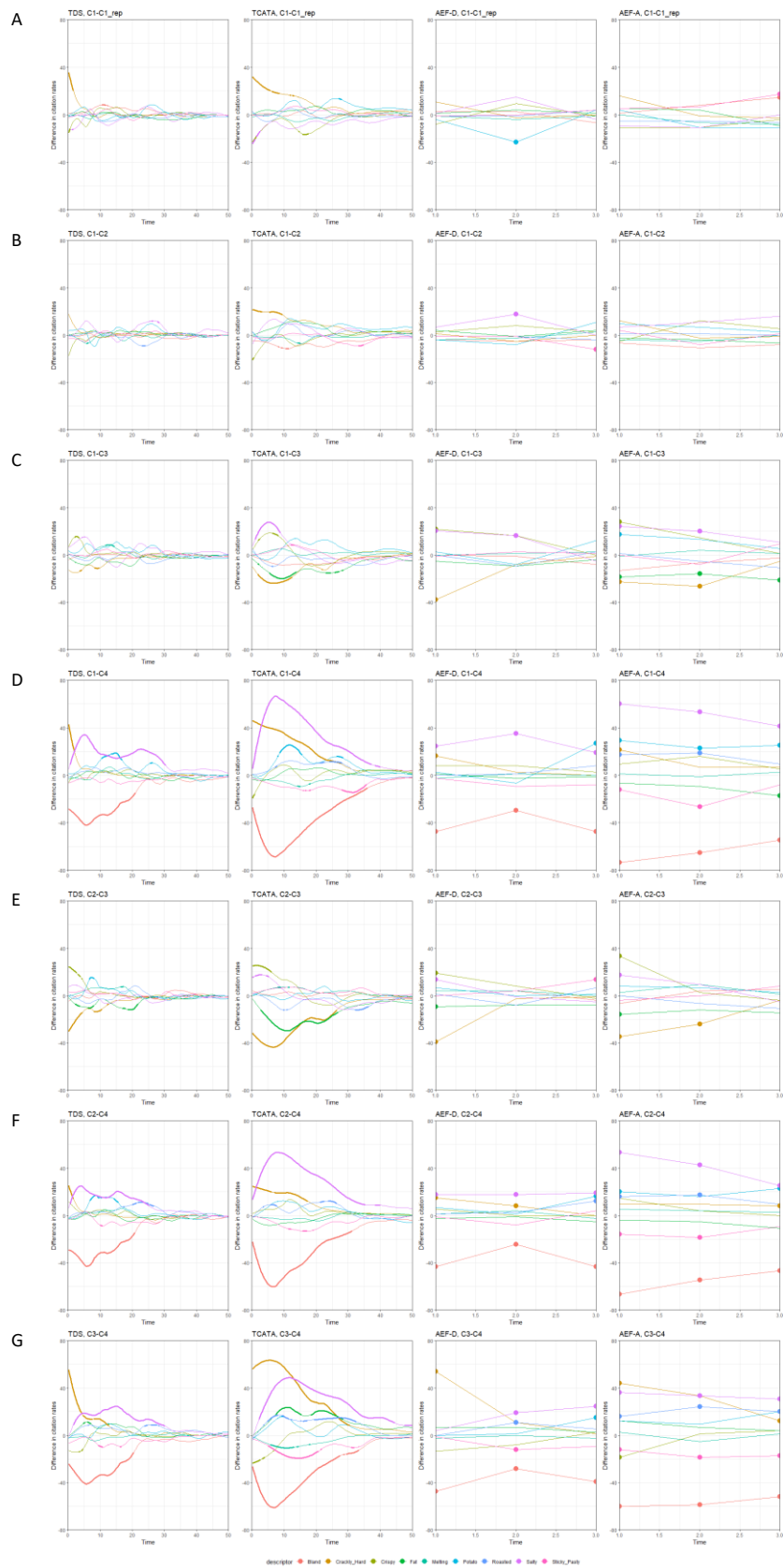


Figure 92. Crisps - Pairwise differences in citation rates (Y-axis) plotted over time (X-axis). Rows of plots each refers to one pair comparison. Within each row, first column refers to TDS, second to TCATA, third to AEF-D, fourth to AEF-A. Bold lines (for TDS and TCATA) and points (for AEF-D and AEF-A) correspond to significant differences in citation rates ($\alpha=0.05$).

C - Crisps		C1-C1_rep				C1-C2				C1-C3				C1-C4				C2-C3				C2-C4				C3-C4					
		TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA
Bland	Citation	1.4	-0.6	0.1	1.4	-0.4	-1.6	-0.5	-0.4	-0.5	-1.3	-0.3	-0.3	-5.6	-5.4	-5.8	-4.4	-0.1	0.3	0.2	0.1	-5.3	-3.8	-5.2	-4.0	-5.2	-4.1	-5.4	-4.1		
	Duration	1.4	2.0	-0.1	0.5	-0.5	-1.6	-0.5	-0.3	-0.3	-1.3	-0.4	-0.3	-7.8	-14.5	-2.0	-1.4	0.2	0.3	0.1	0.0	-7.3	-12.9	-1.5	-1.1	-7.5	-13.2	-1.6	-1.1		
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	1.2	1.0	NT	NT	6.0	6.4	NT	NT	4.8	5.6	NT	NT		
	Citation rate over time	1.8	0.0	0.0	1.0	0.0	2.2	0.0	0.0	0.0	0.0	0.0	0.0	21.0	36.0	3.0	3.0	0.0	0.0	0.0	0.0	20.8	31.2	3.0	3.0	20.4	32.2	3.0	3.0		
Crackly_Hard	Citation	2.0	1.5	0.7	0.7	0.9	1.6	-0.1	0.8	-0.8	-1.1	-2.8	-1.7	2.1	3.2	2.2	1.3	-1.7	-2.7	-2.7	-2.5	1.2	1.6	2.4	0.5	2.9	4.3	5.0	3.0		
	Duration	0.4	6.0	0.4	0.2	0.7	5.9	-0.1	0.1	-2.0	-3.3	-1.0	-0.6	1.0	10.2	1.2	0.7	-2.7	-9.2	-0.9	-0.7	0.3	4.4	1.3	0.6	3.0	13.5	2.2	1.3		
	Time	0.0	0.0	NT	NT	0.0	0.0	NT	NT	0.0	0.0	NT	NT	-0.6	NT	NT	NT	0.0	0.0	NT	NT	-0.6	NT	NT	NT	-0.6	NT	NT	NT		
	Citation rate over time	1.6	11.0	0.0	0.0	0.0	7.8	0.0	0.0	2.8	11.6	1.0	2.0	2.2	27.6	1.0	1.0	8.0	26.4	1.0	2.0	1.6	16.8	2.0	1.0	12.2	30.6	2.0	3.0		
Crispy	Citation	-0.5	-0.6	-0.3	-0.4	0.1	-0.1	0.5	0.5	1.5	1.5	1.9	2.0	1.3	0.7	1.1	1.1	1.4	1.6	1.4	1.5	1.2	0.8	0.6	0.7	-0.2	-0.8	-0.9	-0.8		
	Duration	0.0	-1.6	0.0	-0.2	-0.1	0.5	0.2	0.1	0.9	2.4	0.8	0.5	0.4	2.3	0.3	0.3	1.0	1.9	0.6	0.4	0.5	1.7	0.1	0.2	-0.6	-0.1	-0.5	-0.2		
	Time	1.2	0.8	NT	NT	1.2	0.8	NT	NT	0.2	-0.4	NT	NT	1.2	1.4	NT	NT	-1.0	-1.2	NT	NT	0.0	0.8	NT	NT	1.0	2.0	NT	NT		
	Citation rate over time	0.8	4.4	0.0	0.0	0.2	0.8	0.0	0.0	1.6	2.6	2.0	1.0	0.0	0.8	0.0	0.0	4.0	6.2	1.0	1.0	0.0	0.0	0.0	0.0	1.4	4.0	0.0	1.0		
Fat	Citation	-0.3	0.2	0.6	0.2	1.0	0.5	0.7	-0.3	-1.8	-1.3	-1.0	-0.9	0.1	0.0	-0.2	-0.3	-2.8	-1.9	-1.8	-0.6	-0.9	-0.5	-1.0	0.0	1.9	1.3	0.8	0.6		
	Duration	-0.5	1.6	0.4	0.0	0.5	2.5	0.4	-0.2	-1.8	-5.3	-0.7	-0.7	-0.2	1.3	-0.1	-0.5	-2.3	-7.8	-1.1	-0.5	-0.7	-1.1	-0.6	-0.3	1.6	6.7	0.6	0.2		
	Time	0.1	0.0	NT	NT	NT	-2.5	NT	NT	1.5	2.8	NT	NT	0.1	2.4	NT	NT	-0.2	4.9	NT	NT	NT	4.8	NT	NT	-1.6	-0.2	NT	NT		
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.6	13.6	0.0	3.0	0.0	0.0	0.0	1.0	8.4	24.6	1.0	1.0	0.0	0.2	0.0	0.0	2.8	18.6	0.0	0.0		
Melting	Citation	-0.4	-1.1	-1.1	-1.3	-0.5	-0.5	-0.9	-0.5	1.5	0.9	1.6	0.8	0.2	-0.9	0.0	0.4	2.0	1.4	2.5	1.4	0.7	-0.4	0.9	0.9	-1.3	-1.9	-1.6	-0.5		
	Duration	-0.2	-0.8	-0.8	-0.6	-0.4	-0.8	-0.6	-0.4	0.4	0.7	1.4	0.3	0.1	-1.3	-0.2	0.2	0.8	1.4	1.9	0.7	0.4	-0.5	0.3	0.6	-0.3	-1.9	-1.6	-0.1		
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT		
	Citation rate over time	0.0	0.0	0.0	0.0	0.8	1.2	0.0	0.0	3.0	0.0	0.0	0.0	0.0	1.8	0.0	0.0	1.2	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.0	0.0	0.0		
Potato	Citation	0.7	0.5	0.0	-0.1	0.9	0.4	0.4	0.8	0.8	0.9	0.6	1.2	2.3	1.2	1.5	2.2	-0.1	0.5	0.2	0.4	1.4	0.8	1.1	1.4	1.5	0.3	0.9	1.0		
	Duration	-0.1	4.1	-0.3	-0.1	0.8	4.0	0.0	0.1	1.1	3.3	0.1	0.3	2.3	5.1	0.5	0.7	0.3	-0.7	0.1	0.1	1.4	1.1	0.5	0.6	1.1	1.7	0.4	0.4		
	Time	-0.8	0.4	NT	NT	0.4	2.8	NT	NT	-1.2	0.8	NT	NT	0.1	-1.6	NT	NT	-1.5	0.4	NT	NT	-0.2	-2.0	NT	NT	1.3	-2.2	NT	NT		
	Citation rate over time	0.0	1.2	1.0	0.0	0.2	0.0	0.0	0.0	0.6	0.4	0.0	1.0	7.6	9.8	1.0	3.0	1.8	0.0	0.0	0.0	6.8	0.4	1.0	2.0	0.0	2.0	1.0	1.0		
Roasted	Citation	-1.0	0.4	0.1	-0.5	0.0	-0.2	-0.7	-0.1	-1.3	-0.9	-0.5	-0.7	1.4	1.6	1.4	2.4	-1.3	-0.7	0.1	-0.6	1.4	1.7	2.0	2.5	2.7	2.4	1.9	3.1		
	Duration	-0.8	-0.5	0.1	-0.3	-0.8	0.6	-0.4	0.0	-1.1	-2.4	-0.4	-0.2	0.9	2.2	1.0	1.7	-0.3	-3.0	0.0	-0.3	1.7	1.6	1.4	1.6	2.0	4.6	1.4	1.9		
	Time	-1.4	1.8	NT	NT	-1.6	-1.2	NT	NT	0.6	0.8	NT	NT	NT	2.2	NT	NT	2.2	2.2	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT		
	Citation rate over time	0.2	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.2	6.8	0.0	2.0	0.0	5.4	0.0	0.0	7.4	7.2	1.0	2.0	5.0	22.8	1.0	3.0		
Salty	Citation	-0.9	0.3	0.8	-0.8	0.7	1.4	1.3	0.6	0.8	1.2	1.4	1.3	6.7	41.5	37.3	6.8	0.1	-0.1	0.1	0.7	6.0	40.1	36.0	6.3	5.9	40.3	35.9	5.5		
	Duration	-0.9	-3.9	0.2	-0.1	1.9	1.1	0.4	0.2	0.9	-0.6	0.5	0.4	6.8	15.0	18.9	3.0	-1.0	-1.7	0.1	0.2	4.9	13.9	18.5	2.8	5.9	15.6	18.4	2.6		
	Time	0.9	1.4	NT	NT	0.3	0.2	NT	NT	-1.4	-2.7	NT	NT	NT	NT	NT	NT	-1.7	-3.0	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT		
	Citation rate over time	0.6	2.4	0.0	0.0	1.8	0.0	1.0	0.0	1.0	8.0	2.0	2.0	30.0	39.0	3.0	3.0	0.4	4.0	1.0	1.0	24.8	37.2	3.0	3.0	28.8	45.2	2.0	3.0		
Sticky_Pasty	Citation	0.3	0.6	0.5	2.3	0.1	0.5	-1.4	0.0	0.6	0.9	0.5	0.6	-1.2	-1.0	-2.1	-1.3	0.4	0.4	1.9	0.6	-1.4	-1.5	-0.7	-1.3	-1.8	-1.9	-2.6	-1.9		
	Duration	0.1	1.2	0.3	0.9	0.1	-1.0	-0.9	-0.1	0.2	0.5	0.3	0.0	-1.5	-4.5	-1.1	-0.7	0.2	2.3	1.3	0.1	-1.6	-2.7	-0.2	-0.6	-1.8	-5.0	-1.5	-0.7		
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT		
	Citation rate over time	0.0	0.0	0.0	1.0	0.0	1.2	1.0	0.0	0.0	1.2	0.0	0.0	0.2	10.4	0.0	2.0	0.0	0.0	1.0	0.0	1.2	3.4	0.0	2.0	2.8	18.2	1.0	3.0		
Multivariate	mrChi2	16.9	7.8	5.0	15.7	5.4	13.1	11.7	3.7	15.8	16.2	30.5	17.7	95.6	109.6	104.4	88.3	26.3	21.4	33.0	19.2	76.4	68.0	77.2	71.7	101.2	105.9	112.7	86.1		
	F-MANOVA	0.7	1.8	NT	1.6	1.1	1.5	NT	0.6	1.9	2.5	NT	3.6	11.7	20.9	NT	21.1	3.2	4.5	NT	2.7	9.2	13.5	NT	13.7	13.2	18.1	NT	21.9		
	Discriminating ratio	1.0	1.0	1.0	1.0	0.9	1.0	1.0	0.8	1.2	1.3	1.6	1.6	2.2	2.3	2.6	3.2	1.3	1.3	1.5	1.3	2.0	1.7	2.2	2.6	2.3	2.5	2.7	2.9		

Figure 93. Crisps - Summary of pairwise differences, by attribute (rows) and method (columns). Citation: log of the odd ratio (ANOVA citation = sample + consumer + error). Duration: mean duration (ANOVA duration = sample + consumer + error). Time: median time of citation (bootstrap test). Citation rate over time: total duration of significant differences in citation rates (difference plot). mrChi2: mrChi2 statistic restricted to the pair of samples (mrCA of citations). F-MANOVA: F statistic of the MANOVA restricted to the pair of samples (CVA of durations). Discriminating ratio: average difference between the pair of samples divided by average difference between the replicated samples (citation rates over time). Orange and red cells: significant unexpected differences between replicated samples (red cells: unexpected differences not captured by citations). Light green and dark green cells: significant differences (light green cells: differences not captured by citations). NT: discrimination not tested. Due to a technical problem, only nine attributes over twelve have been presented to the AEF-A panel on the chocolate product category.

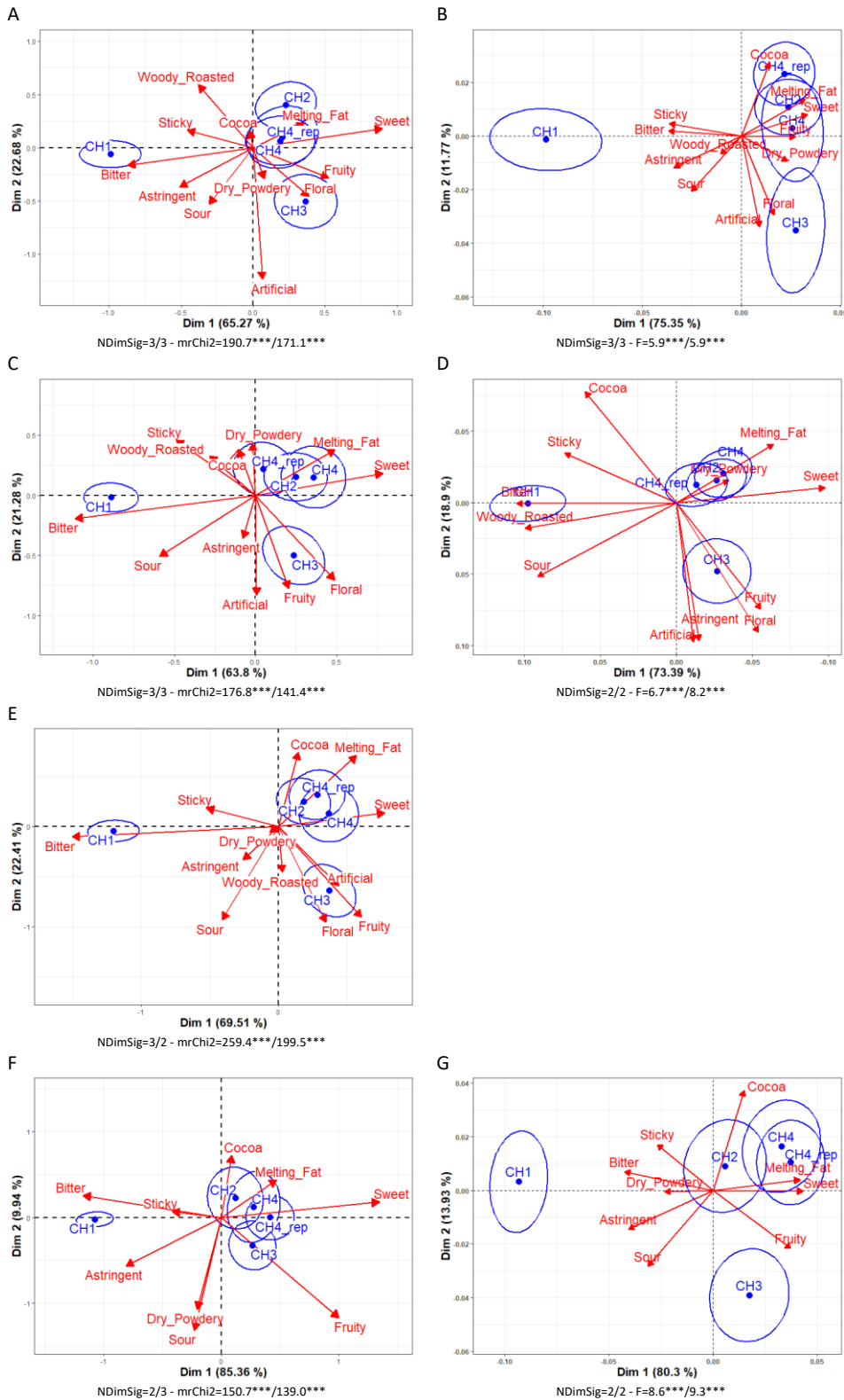


Figure 94. Chocolates - Biplots of multivariate analyses (axes 1-2), with 95% confidence ellipses. First row: TDS, second row: TCATA, third row: AEF-D, fourth row: AEF-A. First column: mrCA of citations, second column: CVA of durations. NDimSig: number of significant dimensions (all stimuli/without replicates). mrChi2: mrChi2 statistic (all stimuli/without replicates). F: MANOVA statistic (all stimuli/without replicates). ***p-value<0.001. CVAs of durations for AEF-D were not produced due to too much collinearity in durations.

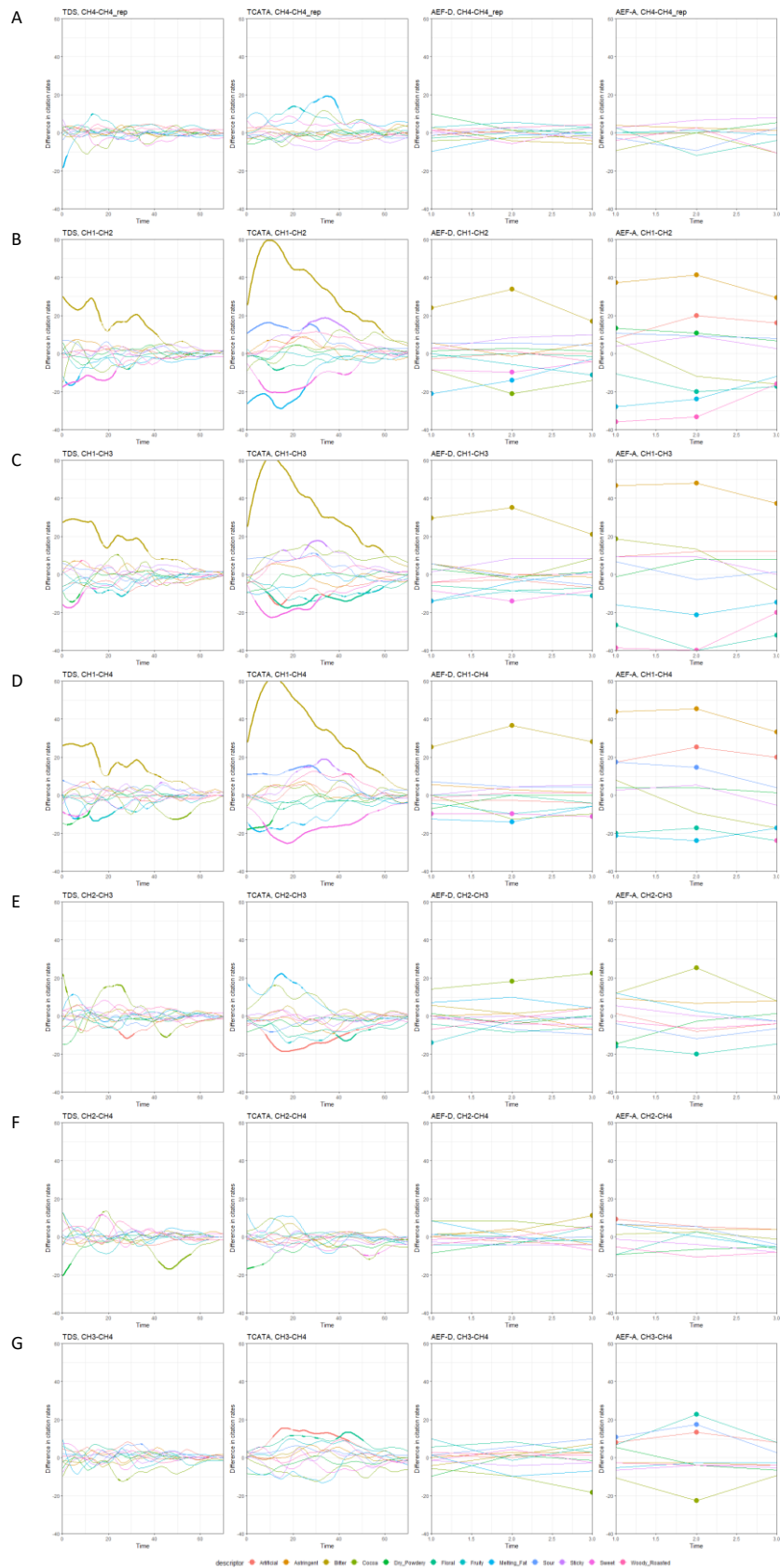


Figure 95. Chocolates - Pairwise differences in citation rates (Y-axis) plotted over time (X-axis). Rows of plots each refers to one pair comparison. Within each row, first column refers to TDS, second to TCATA, third to AEF-D, fourth to AEF-A. Bold lines (for TDS and TCATA) and points (for AEF-D and AEF-A) correspond to significant differences in citation rates ($\alpha=0.05$).

D - Chocolates		CH4-CH4_rep				CH1-CH2				CH1-CH3				CH1-CH4				CH2-CH3				CH2-CH4				CH3-CH4			
		TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A	TDS	TCATA	AEF-D	AEF-A
Artificial	Citation	0.4	0.1	0.4	NT	1.6	1.0	-19.0	NT	-1.4	-0.8	-21.5	NT	0.2	0.0	-20.5	NT	-2.9	-1.9	-2.5	NT	-1.4	-1.0	-1.5	NT	1.5	0.8	1.0	NT
	Duration	1.4	-1.4	0.6	NT	0.7	2.1	-17.8	NT	-3.0	-4.9	-19.4	NT	1.0	0.7	-19.1	NT	-3.7	-7.0	-1.6	NT	-1.7	-1.5	-1.3	NT	2.0	5.5	0.4	NT
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	0.0	0.0	0.0	NT	0.0	2.4	0.0	NT	0.4	12.6	0.0	NT	0.0	0.0	0.0	NT	8.4	38.2	0.0	NT	0.0	0.0	0.0	NT	0.0	33.0	0.0	NT
Astringent	Citation	0.4	0.1	-0.8	0.2	1.8	0.9	1.3	1.8	1.2	-0.2	0.3	1.5	1.6	0.1	1.3	2.5	-0.7	-1.2	-1.0	-0.2	-0.2	-0.8	0.0	0.7	0.5	0.4	1.0	1.0
	Duration	0.1	0.3	-0.4	0.0	1.2	0.7	0.8	0.8	1.0	-1.6	0.3	0.5	1.4	-0.1	0.8	1.5	-0.1	-2.3	-0.5	-0.3	0.3	-0.8	0.0	0.7	0.4	1.5	0.5	0.9
	Time	NT	10.1	NT	NT	NT	NT	NT	NT	NT	-8.9	NT	NT	NT	-19.2	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	-8.8	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	NT	0.8	0.0	2.0	1.2	0.0	0.0	0.0	1.0	0.0	0.0	3.0	0.0	0.6	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	2.0
Bitter	Citation	-0.4	-0.7	-0.4	0.1	4.8	4.9	3.5	3.4	4.5	4.7	4.4	4.1	4.1	5.5	5.0	3.9	-0.3	-0.2	0.9	0.8	-0.7	0.5	1.4	0.5	-0.4	0.7	0.6	-0.2
	Duration	-0.3	-0.7	-0.5	0.1	9.5	21.8	1.4	0.9	10.1	23.0	2.0	1.2	9.3	23.5	2.3	1.1	0.6	1.1	0.6	0.4	-0.2	1.7	0.9	0.2	-0.8	0.6	0.4	-0.2
	Time	-2.8	-1.9	NT	NT	-4.0	-9.2	NT	NT	-7.0	-10.4	NT	NT	-6.6	-11.0	NT	NT	-2.0	-1.4	NT	NT	-1.4	-1.2	NT	NT	0.6	-0.2	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	38.8	59.2	3.0	3.0	40.2	58.8	3.0	3.0	38.0	59.2	3.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
Cocoa	Citation	-0.6	0.0	-0.3	-0.2	-0.2	0.8	-1.5	-1.0	0.7	2.2	0.4	0.7	0.2	0.2	-0.7	-1.0	0.9	1.3	1.9	1.7	0.4	-0.6	0.8	0.0	-0.5	-1.9	-1.1	-1.7
	Duration	-1.2	2.6	-0.1	-0.1	-0.7	3.2	-0.5	-0.1	0.5	6.7	0.2	0.1	-3.0	0.7	-0.3	-0.1	1.2	3.5	0.7	0.2	-2.2	-2.5	0.2	0.0	-3.5	-6.0	-0.5	-0.2
	Time	-0.4	0.8	NT	NT	0.0	0.6	NT	NT	-2.2	0.4	NT	NT	-1.4	0.5	NT	NT	-2.4	-0.1	NT	NT	-1.6	-1.0	NT	NT	1.0	-1.0	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.8	0.6	0.0	1.0	9.8	0.0	0.0	1.0	13.8	1.0	2.0	1.0	15.4	0.8	0.0	0.0	2.6	0.0	1.0	1.0
Dry_Powdery	Citation	0.6	-0.7	0.9	0.2	0.5	0.9	0.4	1.6	-0.5	0.3	0.3	0.1	-0.8	-0.4	-0.3	0.4	-1.0	-0.5	-0.2	-1.5	-1.4	-1.3	-0.7	-1.3	-0.4	-0.8	-0.5	0.2
	Duration	0.5	-0.7	0.5	0.2	-0.1	1.2	0.3	1.2	-1.1	-1.3	0.2	0.4	-1.4	-3.3	0.3	0.2	-1.1	-2.5	-0.2	-0.8	-1.4	-4.5	-0.6	-1.0	-0.3	-2.0	-0.5	-0.2
	Time	0.0	0.0	NT	NT	NT	NT	NT	NT	0.0	NT	NT	NT	0.0	10.2	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	0.0	NT	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.4	0.0	0.0	2.0	4.0	0.0	0.0	0.0	7.8	11.0	0.0	0.0	0.8	0.0	0.0	1.0	3.8	4.2	0.0	0.0	0.0	0.0	0.0	0.0
Floral	Citation	-0.2	-0.3	-0.3	NT	-1.3	-1.6	-0.8	NT	-2.3	-2.6	-2.7	NT	-0.8	-1.4	-0.8	NT	-1.0	-1.0	-1.9	NT	0.5	0.1	0.0	NT	1.5	1.2	1.9	NT
	Duration	-0.2	-0.7	0.0	NT	-0.5	-2.2	-0.7	NT	-1.6	-8.7	-2.1	NT	-0.1	-1.7	-0.9	NT	-1.2	-6.6	-1.4	NT	0.4	0.5	-0.2	NT	1.5	7.0	1.2	NT
	Time	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	6.0	0.0	0.0	0.0	43.6	0.0	0.0	0.0	0.8	0.0	0.0	0.0	10.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	18.4	0.0	0.0
Fruity	Citation	0.3	0.8	1.2	-0.4	-1.3	-0.1	-2.5	-2.1	-2.0	-1.1	-3.9	-3.6	-1.7	-0.5	-2.8	-2.5	-0.7	-1.1	-1.4	-1.4	-0.4	-0.4	-0.3	-0.4	0.3	0.7	1.1	1.0
	Duration	2.4	3.5	0.7	-0.1	-1.7	-1.4	-1.9	-1.2	-2.1	-6.8	-2.6	-1.7	-4.2	-3.4	-2.1	-1.4	0.4	-5.0	-0.6	-0.5	-2.5	-2.0	-0.1	-0.2	-2.1	2.9	0.5	0.4
	Time	NT	NT	NT	NT	NT	2.8	NT	NT	NT	-1.2	NT	NT	NT	-1.2	NT	NT	-2.9	-3.7	NT	NT	-4.1	-1.8	NT	NT	-0.8	2.4	NT	NT
	Citation rate over time	1.4	5.0	0.0	0.0	1.6	0.0	1.0	2.0	9.6	15.8	2.0	3.0	13.0	0.0	0.0	3.0	0.0	3.2	1.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
Melting_Fat	Citation	-0.9	1.4	-0.7	-0.2	-1.8	-2.5	-2.5	-1.9	-1.2	-1.0	-1.4	-1.3	-0.6	-2.5	-2.2	-1.5	0.6	1.5	1.1	0.6	1.2	0.0	0.3	0.4	0.6	-1.5	-0.8	-0.2
	Duration	-1.6	6.7	-0.2	0.0	-3.5	-7.9	-1.1	-0.7	-2.4	-2.3	-0.7	-0.6	-2.3	-7.1	-1.0	-0.7	1.1	5.6	0.5	0.1	1.2	0.9	0.1	0.0	0.1	-4.7	-0.4	-0.1
	Time	3.2	-1.0	NT	NT	1.4	4.5	NT	NT	0.2	2.0	NT	NT	1.6	1.8	NT	NT	-1.2	-2.6	NT	NT	-0.2	-3.0	NT	NT	1.1	-0.6	NT	NT
	Citation rate over time	2.0	10.0	0.0	0.0	6.4	25.4	2.0	2.0	0.0	0.0	1.0	2.0	4.6	15.2	1.0	3.0	1.2	10.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Sour	Citation	-1.6	-1.7	0.0	-1.1	1.3	2.4	2.3	1.1	0.3	1.1	0.2	-0.1	2.6	3.3	2.0	1.6	-1.0	-1.4	-2.2	-1.2	1.3	0.8	-0.3	0.4	2.3	2.2	1.8	1.7
	Duration	-0.5	-1.5	0.0	-0.3	1.2	5.9	1.2	0.6	0.4	2.8	-0.1	0.1	2.0	5.9	1.2	0.8	-0.9	-3.2	-1.2	-0.5	0.8	0.0	0.0	0.3	1.6	3.1	1.2	0.7
	Time	NT	NT	NT	NT	NT	NT	NT	NT	-10.4	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.2	26.4	0.0	0.0	0.0	2.4	0.0	0.0	0.8	22.0	0.0	2.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.0	2.0
Sticky	Citation	0.3	-0.8	-0.1	0.7	1.6	1.4	2.5	1.2	2.4	1.8	1.9	1.1	1.2	1.2	1.1	0.6	0.8	0.4	0.5	-0.1	-0.5	-0.2	-1.4	0.7	-1.2	-0.6	-0.9	-0.6
	Duration	0.0	-2.7	-0.2	0.3	1.5	5.8	1.6	0.3	1.8	5.5	1.2	0.4	1.6	3.3	0.5	0.0	0.4	-0.3	-0.4	0.1	0.1	-2.5	-1.1	-0.3	-0.3	-2.2	-0.7	-0.3
	Time	-4.1	-2.6	NT	NT	3.7	0.2	NT	NT	NT	NT	NT	NT	10.3	4.9	NT	NT	NT	NT	NT	NT	6.1	4.5	NT	NT	NT	NT	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	20.0	0.0	0.0	0.0	12.4	0.0	0.0	1.6	14.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Sweet	Citation	-0.3	0.2	-0.1	-0.6	-4.1	-4.3	-3.4	-37.1	-4.0	-4.5	-3.8	-37.5	-3.6	-5.1	-4.0	-37.9	0.1	-0.1	-0.4	-0.3	0.4	-0.7	-0.6	-0.8	0.3	-0.6	-0.2	-0.4
	Duration	-1.7	3.2	0.0	-0.1	-4.4	-7.0	-2.2	-2.8	-3.4	-8.7	-2.5	-3.0	-2.6	-12.0	-2.5	-3.1	1.0	-1.7	-0.3	-0.1	1.7	-5.0	-0.3	-0.2	0.8	-3.3	0.0	-0.1
	Time	-5.4	-0.8	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	NT	1.2	1.2	NT	NT	-0.9	-2.4	NT	NT	-2.1	-3.6	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	23.2	31.2	1.0	3.0	11.0	29.4	1.0	3.0	8.0	47.0	3.0	3.0	0.0	0.0	0.0	0.0	3.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0
Woody_Roasted	Citation	0.1	-0.3	0.6	NT	-0.1	0.8	-0.2	NT	1.5	0.9	-0.5	NT	1.2	0.4	0.0	NT	1.5	0.1	-0.3	NT	1.3	-0.4	0.2	NT	-0.2	-0.5	0.5	NT
	Duration	1.2	-1.8	0.5	NT	-0.6	3.3	0.0	NT	0.6	2.7	-0.2	NT	0.6	4.0	0.1	NT	1.2	-0.6	-0.2	NT	1.2	0.7	0.1	NT	0.0	1.2	0.2	NT
	Time	NT	-0.8	NT	NT	-2.4	2.8	NT	NT	NT	0.2	NT	NT	NT	1.0	NT	NT	NT	-2.7	NT	NT	NT	-1.6	NT	NT	NT	1.0	NT	NT
	Citation rate over time	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	5.6	0.0	0.0	0.6	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Multivariate	mrChi2	6.1	8.6	7.3	3.7	63.8	50.8	71.8	68.4	70.7	61.7	85.7	80.3	56.4	56.4	82.7	80.5	25.9	19.3	27.9	12.9	15.7	8.9	10.8	8.8	17.4	16.2	19.9	9.5
	F-MANOVA	1.4	0.7	NT	0.7	6.4	9.0	NT	10.6	7.6	8.8	NT	15.0	6.6	9.7	NT	16.3	1.6	2.4	NT	2.7	1.9	0.8	NT	1.5	2.4	2.6	NT	2.8
	Discriminating ratio	1.0	1.0	1.0	1.0	1.8	2.0	2.7	4.3	2.0	2.3																		

III.2.5. Main findings of section III.2

In the three presented articles, TDS, TCATA and different variants of AEF were used to investigate the performances of temporal methods: validity, reliability, temporal resolution and discrimination ability.

It has been shown that with simple controlled temporal stimuli, consumer panels were able to detect and report sequences of attributes that matched the sequences of delivered sapid and aromatic compounds, confirming thus the face validity of the methods (III.2.2). Despite low individual repeatability, no significant difference was observed between the replicated stimuli regardless of the variable analysed, suggesting that the measurements were quite reliable at the panel level. However, construct and content validity were not demonstrated. The concepts of dominance and applicability were interpreted similarly by most consumers, and were essentially linked to new perceptions. Durations of dominance/applicability, citation rates and intensities have been shown to be strongly correlated with each other, but weakly with compound concentrations or durations of diffusion. Criterion validity was also unclear, probably due to a limited capacity of consumers to identify unusual sensations. These results suggest that the temporal methods can be used with consumers to measure qualitative changes in perceptions related to simple temporal stimuli, but not to quantify the magnitudes of the changes.

With complex food products, when considering confidence intervals around citation times and durations of the attributes, a great heterogeneity was observed among consumers in how they described their temporal perception of foods (III.2.3). With concurrent methods, attributes that appeared to be perceived at different times based on the visual interpretation of plots of citation rates over time (TDS or TCATA curves) were not necessarily separated by bootstrap analyses. With TDS and especially with TCATA, the most significant changes in perception were observed at the start of tasting, most of the attributes perceived afterwards cannot be separated temporally. These results demonstrated that despite a data collection in high temporal resolution, when using confirmatory data analysis, TDS and TCATA were not always effective in capturing small within product temporal differences in perception, at least with a consumer panel of usual size.

Regarding discrimination between products, it has been shown that between 70 and 85% of differences were brought by the presence/absence of sensory attributes without consideration of temporal aspects (III.2.4). No method was found to be superior to the others when considering all performances. Dominance-based methods were generally less discriminative, but more reliable than applicability-based methods with complex products.

Chapter IV. General discussion

IV.1. Limitations

Before discussing the results, it is important to acknowledge some limitations of this work.

The systematic scoping review was conducted including only research articles, so results may not be representative of industry practices. Moreover, given the time required to compile and analyse the information contained in the articles, the review has been concluded after the other articles were published.

It was chosen to focus on studies involving food products, linked to INRAE scientific themes. Temporal methods can also be used in other fields, notably in cosmetics, but it was beyond the scope of this research work to study their relevance for evaluating non-food products. It was also deliberately chosen to focus on the application of temporal methods with consumer panels. Thus, under no circumstances can the results presented be generalized to temporal methods used with other types of panels.

Not all combinations of measurements have been tested and compared. For practical reasons, it was not possible to cross all the components of temporal methods: measured concepts (dominance/applicability/intensity) x moment of evaluation (continuous time/fixed time/periods/global) x attributes (predefined list/free comment) x settings (home/laboratory/other) x categories of product x expertise of panellists (etc.). It was chosen to focus mainly on the comparison of dominance and applicability measurements (which are currently the most used with consumers) collected using the two extreme temporal resolutions (continuous time vs. periods). It cannot be excluded that the investigation of other combinations of components could have led to different conclusions.

Regarding the experiments, two datasets (II.2 and II.3) were collected opportunistically, adding a methodological question to the product-oriented research questions raised by academic or industrial partners. This explains why certain experimental conditions were not completely related to the objectives of this research work, and why certain data were not exploited. Not all the data from the dataset presented in Section II.4 were used. Due to their different nature making comparison with other methods difficult, the FC-AEF data have not yet been analysed. Although collected in most studies, hedonic data have not been analysed either. It was not beyond the scope of this work to investigate the benefit of combining hedonic and descriptive data.

In some of the featured articles, temporal data were transformed into static data (“CATA transformation”) and considered as results of static evaluations. The approach was interesting to study to what extent temporal methods discriminated between products based on temporal aspects. However, the data collection process of CATA does not involve focusing on temporal aspects, so real CATA data could have been different than data obtained after the CATA transformation. It would therefore not be relevant to conclude on the comparison between static and temporal measurements on the basis of the results presented.

In the studies presented in Section II.3, replicated samples were systematically presented after all other samples to maintain a fixed order of presentation, better suited for method comparison. It cannot be excluded that consumer repeatability was thus slightly underestimated (but in the same way for all methods) due to potential order, carryover or fatigue effects.

IV.2. Implications of key findings

IV.2.1. About the methods: what is asked vs. what is done

Dominance and applicability are most often assumed to be different concepts. TDS imposes a forced choice of a descriptor among all those perceived at the same time. This likely encourages consumers to be conservative and select the attribute they are most confident of having perceived. Conversely, TCATA invites consumers to make multiple choices supposed to be representative of all the sensations perceived. In both methods, a time constraint is imposed on the decision-making process. Although indirect, this is also a forced choice constraint that could explain why, in practice, the concepts of dominance and applicability are not so far apart, as demonstrated by the study with controlled temporal stimuli (III.2.2). It has been shown that, whatever the concept measured and the operational definition given for the task, consumers tended to report what they perceived as a function of several parameters including sensory intensity, duration of stimulation, familiarity/easiness to identify the sensation and novelty (i.e. the new sensory attribute perceived). However, due to the “forced-choice time constraint”, it appeared that the most important parameter was the novelty in TDS and TCATA (III.2.2).

The biggest difference between TDS and TCATA is in the durations of dominance/applicability of attributes. With TDS, dominance durations are computed based on the time elapsed between two consecutive citations of distinct attributes. With TCATA, durations of applicability begin when an attribute is chosen and end when consumers decide to uncheck it (or at the end of the evaluation). However, it has been shown that unchecking attributes can be difficult for panellists (Ares et al., 2016). The fading variant of TCATA (Ares et al., 2016) attempts to solve this problem by automatically unchecking attributes after a certain duration. In this way, the duration of applicability depends on a subjective choice of the experimenter. Although TCATA fading was presented as more discriminative than TCATA (Rizo et al., 2020), in view of the results presented in this work, it can be assumed that this does not necessarily mean that the method is more valid or accurate. In all cases, contrary to TDS, with TCATA (with or without fading option), the panellists must keep in memory the state (checked or unchecked) of each attribute presented in the list. Recent works demonstrated a limited capacity of gustatory working memory (Lim et al., 2022), so it is likely that tracking the state of multiple sensations (more than three) is out of reach for most consumers.

Whether with DTS or TCATA, what is being measured with the durations is unclear. Neither durations of dominance nor durations of applicability were shown to be correlated to the durations of diffusion of the compounds in the model solutions (III.2.2). At the panel level, the average durations of dominance/applicability of attributes were strongly correlated with citation rates, but weakly with the concentrations of the compounds in the controlled temporal stimuli (III.2.2). Short differences in attribute durations of stimulation (<3 s) were not consistently captured by temporal methods (III.2.2 and III.2.4). These results suggest that although the evaluation task in TDS and TCATA is concurrent with tasting, reports of perceived sensations are not simultaneous with perception, confirming that the “direct link from the tongue of the panellist to the hand moving the mouse is a fantasy” (Lawless & Heymann, 2010). Compared to a tasting where consumers simply had to signal when they no longer perceived anything (without having to describe the samples during the tasting), the total durations of the TDS and TCATA evaluations were respectively 40% and 70% longer (III.2.3). Although, in the usage questionnaires, consumers generally did not express any particular difficulty in performing the tasks (III.1.2, Ares et al., 2015), the differences observed between

the methods probably reflect a gradient of cognitive load to the disadvantage of concurrent evaluation methods and in particular TCATA.

Retrospective temporal methods do not measure durations, and are therefore not subject to the “forced-choice time constraint” in decision making. It can be assumed that this type of measurements reduces the stress that some consumers have reported feeling during concurrent measurements (III.1.2, III.1.3). On the other hand, the retrospective evaluation in AEF undoubtedly makes the task more analytical and introduces new biases linked to memory. The AEF-A task is likely very similar to a CATA task, with consumers first determining which attributes they perceived and then subsequently assigning each attribute to one or more of three periods of perception. Thus, even if the main sources of heterogeneity among panellists (citation times and durations) are removed with AEF, some variability remains in the choice of the period of dominance or applicability of the sensory attribute (III.1.3). It has been hypothesized that concurrent and retrospective measurements might align with Kahneman’s theory on the two systems of thinking (O’Brien, 2012). The first system (concurrent measurements) is fast and intuitive, while the second (retrospective measurements) is slower and more deliberative. Jaeger et al. (2017) argued that System 2 may decrease the ecological validity of the results as food choices were reported to be largely influenced by intuitive and automatic behaviours primarily determined by System 1.

The main advantage of AEF methods is their simplicity: they are easy to implement for the experimenter and easy to understand for consumers. Although the temporal resolution of AEF-A is coarse, the main within-product temporal patterns are captured and the discrimination ability of AEF-A is comparable to that of TCATA (III.2.3, III.2.4). A drawback of AEF-A is the possibility that consumers will check more attributes than they actually perceived to ensure they have not missing anything (III.1.3). This can make it difficult to distinguish between agreement and acquiescence bias also observed with CATA and CATA-like measurements (Callegaro et al., 2015; Kim et al., 2017; Mahieu et al., 2020a). FC-AEF-A is not subject to this problem, but the combined lack of agreement among consumers regarding both attributes and period choices may make the method unsuitable for measuring temporal perception of complex products (III.2.4). AEF-D and FC-AEF-D seem to suffer too much from the limitation of a single attribute choice per period, and AEF-RATA does not seem to provide additional information compared to AEF-A. Implementing these three variants of AEF was

useful for better understanding the nature of the differences between retrospective and concurrent measurements, but they appear to have little value for real use for product characterization.

IV.2.2. About the data: what is collected vs. what is analysed

TDS and TCATA ask consumers to report events related to the presence or absence of dominant/applicable attributes. This sequence of events constitutes a stochastic process, rarely analysed as such (Lecuelle et al., 2018). Indeed, events are most often transformed in discrete time series assumed as continuous if the discretization step is sufficiently small. In view of the results presented, it is suggested that the temporal resolution of the methods to capture differences (within and between products) is typically on the order of three seconds. Thus, continuous data transformation prior to statistical analysis may not be the best option, and a discretization step of one second or even more is probably sufficient in most cases, and even recommended to improve data quality (Derks et al., 2022). Larger steps would amount to dividing time into subjective and potentially meaningless periods, given that perceptual periods are unlikely to be of uniform durations (Lecuelle et al., 2018; Visalli et al., 2021) but rather linked to the stages of food oral processing and flavour release.

It has been shown that the temporal dimension adds new information, but also a new source of uncontrolled variability due to the heterogeneity among consumers with respect to temporal variables (III.1.2, III.2.2). Much of this heterogeneity is due to delays in first citation times, which are likely uninformative. Indeed, these delays are more certainly due to differences in individual behaviours or reaction times rather than to product temporal signatures. A frequent data transformation (I.3.3) is the standardization of time (Lenfant et al., 2009) which consists of aligning the start and end of perception for all panellists. However, this approach distorts individual sequences and can hide real differences in product perception (Castura, 2018), without systematic guarantee of improving reliability (III.2.3) or agreement between consumers, at least for TCATA (Dietz, Yang, et al., 2022).

The golden criterion for evaluating the relevance of results is most often limited to the level of discrimination observed between products, but the source of the discrimination is rarely examined. Reported perception times are lagged relative to perception and include random variations that are not considered with plots of citation rates over times (TDS or TCATA

curves). Although these plots rely on the unrealistic assumption that perceptions are aligned in time across consumers, these representations of TDS/TCATA data are the most frequently used (I.3.3). Due to the transformation of event data into time series, unexpected attributes may by chance achieve relatively high citation rates (III.2.2). Furthermore, although they are sometimes interpreted in this way, citation rates are not quantifications of the intensity of the perception of sensory attributes (III.2.2), but rather a reflection of the level of agreement between consumers in the actual dominance or applicability of attributes at a given time. Thus, TDS/TCATA curves should be considered exploratory data analyses, and it has been confirmed (I.3.4) that it is not appropriate to visually compare two sets of curves to conclude on differences between products, as previously suggested by Meyners (2020).

The most commonly used multivariate representations (PCA, CA or CVA) are not very informative on the nature of the differences between products. Inherently, these maps (like univariate analyses) do not capture the dynamic changes that occur during perception because they summarize temporal perception in a single variable (most often attribute duration or average citation rate). Multivariate analyses primarily capture differences in presence/absence of attributes (similar to CATA) unrelated to temporal aspects of perception (III.2.4). Alternative trajectory maps based on variables aggregated by periods can overcome this issue, but they do not resolve the problem of the subjective period choice. Furthermore, they do not allow to conclude on significant differences unless confidence ellipses around the “product x period” points are plot. However, such an approach mixes the two sources of variation (within and between product).

IV.3. Recommendations

IV.3.1. For the method choice

Assuming that the quantification of descriptors is beyond the reach of consumers, the choice of the most appropriate method necessarily involves a compromise between the complexity of the task for the consumers, the precision of the temporal descriptions, the discrimination capacity and the validity of the results (Figure 97).

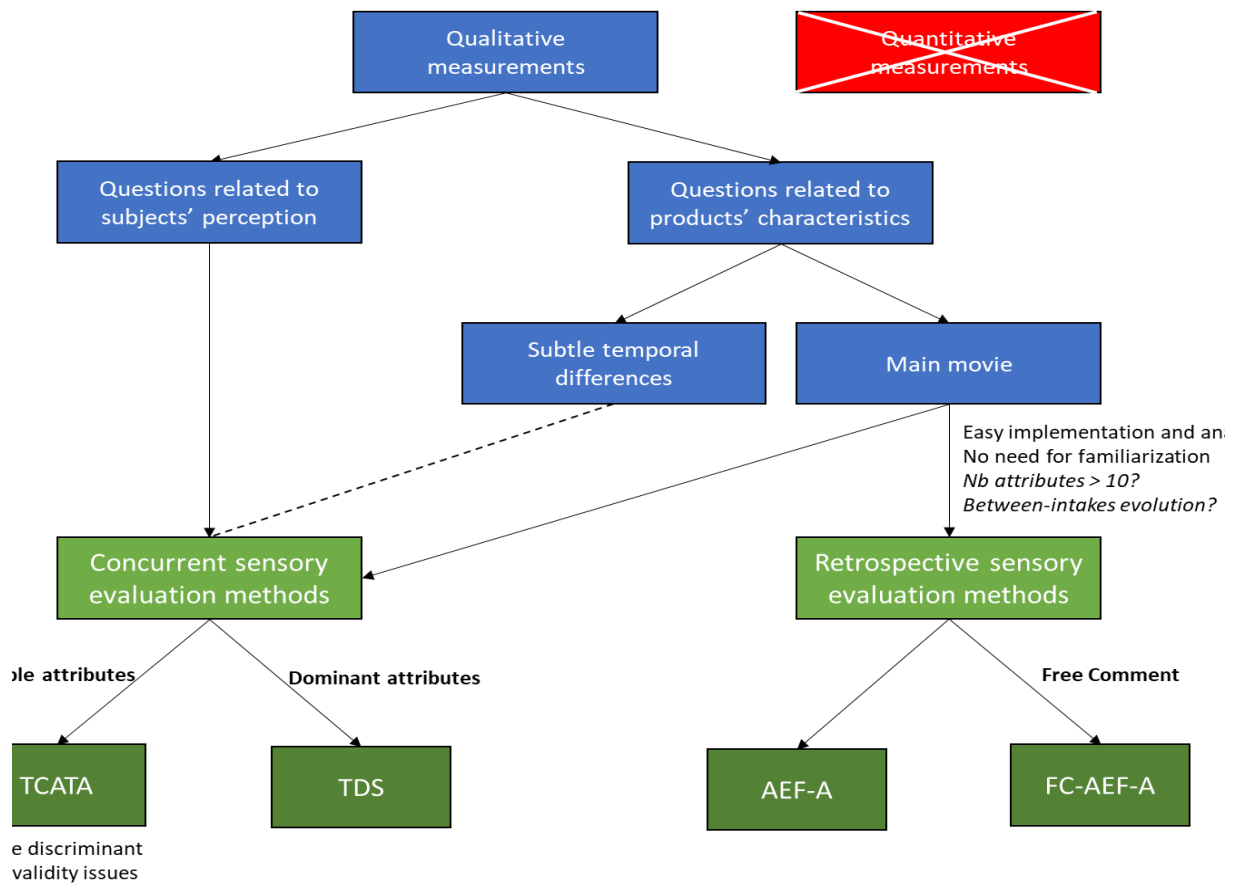


Figure 97. Flowchart for choosing the method of measuring consumer perception. Statements in italic are conjectural.

If the objective is to study the temporal perception of subjects (for example to study physiological processes), concurrent methods are required because they are more spontaneous and closer to immediate perception. If the objective is to study between-product temporal differences on samples expected with subtle differences, concurrent methods are necessary, but there is no guarantee that consumers will be able to identify these differences, especially with complex products. TCATA is overall more discriminative compared to TDS, but it has also more validity issues when used with complex food products (III.2.4).

If the objective is to study between-product differences and the temporal aspect is only secondary, or if an accurate temporal description is not required, retrospective methods such as AEF-A or FC-AEF-A can be considered. In the worst case, they will probably be equivalent to a regular CATA or free comment (with the same limitations as these methods). As they are easy to implement and do not require familiarization, they may be a reasonable choice for use in uncontrolled conditions, outside the laboratory, when no briefing is possible. Their temporal resolution could also be well suited to measurements of perception of full portions

over multiple intakes (as a qualitative variant of sequential profile). Moreover, they can be used with any survey software, or even paper and pen, and they can probably be used with a larger number of attributes than TDS and TCATA.

IV.3.2. For the implementation of the methods

Although the question was not addressed in this research work, the choice of the number of consumers is very important. The mean number of consumers included in studies involving temporal methods is about 70 (I.3.3). This number was probably chosen for practical reasons, in reference to the minimum number of tasters recommended for hedonic tests (Mammasse & Schlich, 2014) or other descriptive tests involving consumers (Ares, Tárrega, et al., 2014). However, it is very plausible that a larger number of subjects is needed to draw robust conclusions related to temporal aspects due to the additional noise brought by the temporal dimension of the data. Okamoto et al. (2022) showed that standard error around citation rates required samples of sizes larger than 100 to be correctly estimated (samples of size 50-100 being acceptable). Cardot et al. (2019) concluded that semi-Markov modelling was not very effective for small samples. Thus, it seems preferable to choose at least a number of subjects in the upper range of the recommendations of the other methods (80 to 100 consumers, or even more).

It was pointed out in the literature review that the reliability of temporal measurements was not considered in many studies. The results presented showed that individual repeatability is quite low in average, particularly for the most complex food products. Although individual repeatability cannot necessarily be expected from consumers, it is important to verify that the data contain a "strong enough" signal and are not as noisy as at the panel level, differences would appear if the test was repeated (Jaeger et al., 2013). Adding replicated samples (at least one) helps determine the level of noise in the data at the panel level (III.2.1, III.2.2, III.2.3). To limit potential adverse effects, this or these samples should be positioned among the products and presented in a random or balanced order.

The choice of attributes is also a crucial step, in particular with TDS and TCATA. The number of descriptors most frequently observed in the literature is eight, distributed into three sensory modalities (III.1.3). For TDS, Pineau et al. (2012) recommended a maximum of 10 attributes. For TCATA, Jaeger et al. (2018) showed that consumers were able to use up to 15

attributes. Among the studies presented, both methods showed the best performances when short lists of attributes were used (III.2.2, but the number of attributes was also related to the complexity of the product space). “By modality” variants of TDS and TCATA, which consist of two alternate assessments with attribute lists restricted to a single sensory modality, have so far only been implemented with trained or semi-trained panels (Dietz, Cook, et al., 2022; Lesme et al., 2020; Pittari et al., 2022; Rizo et al., 2019; Rodrigues et al., 2018). These variants seemed to give “better results” and/or present greater simplicity for panellists. Even if these conclusions remain to be confirmed, they go in the direction of limiting the size of the list of attributes to make the task easier for consumers and to avoid competition between attributes (e.g. texture vs. flavours). To do this, several approaches can be considered. First, attributes that are too “obvious” such as “coffee” in coffees or “alcohol” in distilled beverages can be omitted because they risk capturing all the citations, especially with TDS. Second, attributes with too much similarity should be avoided or grouped into categories (e.g. “fruity” or “citrus” instead of “lemon” and “orange”) or meta-attributes (e.g. “crunchy/crispy”). Identifying the sensations they perceive is not an easy task for consumers, especially when they are unfamiliar with the sensations, and a lack of capacity in attribute identification can threaten the validity criterion (III.2.2). Thus, prior to the temporal descriptive task, asking consumers to report what they perceive of a warm-up product using free comment can allow the experimenter to explore consumers' initial ability to recognize sensations prior to be influenced by the list of attributes (III.2.2). Finally, sensory attributes should be explained to consumers, even if they seem self-explanatory. Examples of day-to-day products presenting the sensory characteristics can be provided, or even onomatopes for texture attributes (Natsume et al., 2023).

Several definitions of the concept of dominance have been proposed in the literature (Varela et al., 2018). Novelty was observed to be the main trigger for dominance (III.2.2). Even if the way in which the concept is presented to consumers does not seem to have an impact on the results (Hutchings et al., 2022), a non-ambivalent instruction to give to participants could simply be to click on the attribute as soon as it is perceived. Performing the task with a video (Velázquez et al., 2020b) presenting known temporal differences can also help consumers apprehend the temporal nature of the task and panel leaders to detect consumers who have not understood it.

For AEF methods, it has been shown that the attack and finish periods correspond roughly to the first and last five seconds of the tasting (III.2.2), but the durations of these periods probably depend on the product category. To limit subjectivity in the interpretation of periods and collect more meaningful data, physiological stages (e.g. "after mouthing", "after swallowing", "aftertaste") can be proposed instead of "at the beginning", "after a few seconds" and "at the end" (III.2.3).

IV.3.3. For the analysis of temporal data

The results presented pointed out the importance of examining results collected from the temporal evaluation methods beyond face validity (I.4.1 and III.2.2). In particular, the importance of using confirmatory data analyses in addition to exploratory analyses was demonstrated. Among the available statistical tools, it is necessary to choose those which allow conclusions to be drawn which consider the dynamics of the changes that occur during perception which is the original vocation of temporal methods. Thus, a framework for confirmatory data analyses encompassing all dimensions (static and dynamic) of data is proposed below.

First, for TDS and TCATA it is recommended to discretize the data with a relatively coarse grain (one second) and to left-standardize data (i.e. consider the begin of perception as the time to first click on an attribute, not the click on START).

Second, the behaviour of panellists should be examined before data analysis. This can be done by visualizing the distributions of variables such as time to first citation (before left-standardization), total duration, and number of attributes used. These outputs provide an overall overview of the panellists' understanding of the tasks, and may possibly suggest the presence of individual outliers (e.g. individuals for whom the times before the first citation or the total duration are too long, or individuals who do not cite enough or too many attributes).

In studies involving Just-About-Right scales (Rothman & Parker, 2009), only attributes with at least 20% responses (derived from the Pareto principle) became candidate for penalty analysis (Ares, Dauber, et al., 2014; Narayanan et al., 2014; Plaehn, 2012; Xiong & Meullenet, 2006). For lack of anything better, such an arbitrary threshold can be chosen to draw more conservative conclusions with temporal data, and only attributes cited by more than 20% of consumers can be retained to limit the impact of "false positive" perceptions (III.2.2).

Third, the reliability of the measurements should be assessed. The calculation of similarity indices (i.e. distances between temporal sequences of replicated samples) can help to spot individuals which are not repeatable or not in agreement with the panel (Frascolla et al., 2023; Natsume et al., 2023). Overall reliability at the panel level can also be assessed with existing tools, for example by examining plots of significant differences in proportions of citations over time between pairs of samples. The total number of significant differences is expected close to zero for pairs of replicated products (NbSigRep) and greater than zero for pairs of different products (NbSigDiff), respectively. A large NbSigRep or a ratio (NbSigDiff/NbSigRep) lower than 1 should alert the experimenter about reliability issues (III.2.4).

Fourth, the interpretation of within product temporal evolution should be objectified. For a given product, to test whether attributes were cited at different times, it is possible to compute confidence intervals around the times to first citation of these attributes and the times of end of perception. The bootstrap approach is described in Section III.2.3, but to simplify the calculations, it is certainly possible to use parametric models by applying appropriate transformations to obtain normal distributions for the times.

Fifth, to assess the pairwise differences between products at each time, the difference curves can constitute an appropriate confirmatory analysis if certain elements relating to their interpretation are considered. Only successive significant differences accumulating over a minimum duration should be interpreted. This duration can be arbitrarily defined, for example three seconds, with regard to the assumed temporal resolution of the methods. Furthermore, it should be kept in mind that significant differences in citation rates between products are not necessarily due to differences in applicability/dominance at a given time. In fact, these differences may be due to different levels of agreement for the attribute at this time (e.g. an attribute cited by 50% of subjects in one product and 80% in the other).

Sixth, if understanding the nature of differences (presence/absence, duration or time of citation of the dominant/applicable attributes) between products is of interest, successive analyses can be carried out. Differences in perception of attributes without regard to the temporal dimension of data (as in the case of a static evaluation task) can be assessed after CATA transformation using an univariate linear model based on count data or a multivariate multiple-response Correspondence Analysis (Mahieu et al., 2020). Then, durations of dominance/applicability of attributes can be analysed in the same way by replacing mrCA by

CVA (Galmarini et al., 2017). The conclusions can then be compared to those obtained with citations to determine whether considering durations provides additional information. Finally, the time of citations of the attributes can be compared across products with bootstrap (III.2.4) to determine which attributes have offset temporal perceptions. Whatever the variable analysed, reporting the effect size of significant differences is a good practice and can help interpret results (Johnson et al., 2020). A major drawback of this approach is the accumulation of statistical tests.

IV.3.4. For reporting research findings and sharing data related to temporal measurements

The results from the systematic literature review (I.3.3) showed a great heterogeneity in the way research was reported, with many articles having forgotten to report information necessary for replicating the studies. This is why a checklist has been proposed, allowing authors and reviewers to verify that all key elements have been reported (I.3.3, Supplementary material). This work was established on the basis of the elements expected for studies involving temporal methods, but it is sufficiently generic to be used within the framework of most studies involving descriptive sensory evaluation.

Given the lack of open datasets identified in the systematic literature review (I.3.3), it appeared particularly important to make the datasets collected during this research work available to promote Open Science and FAIR principles in food science (Visalli, Schlich, Mahieu, Thomas, et al., 2023). Compiling these datasets made it possible to build a template for reporting data in a unified way, using a controlled vocabulary (I.3.3, Supplementary material). This template is also generic, it is versioned and can be downloaded on a public repository, thus it can be easily adapted and/or improved by the sensory-consumer science community. It is hoped that the FAIRification of data will generalize in such a way that a sufficient body of data may be available in the future. Only then can more substantiated recommendations be formulated for the use of sensory evaluation methods.

Chapter V. Conclusion

V.1. Temporal methods can be used with consumers, but too much should not be expected

As happened with static methods, temporal evaluation methods shifted from quantitative to qualitative measurements, resulting in rapid temporal descriptive methods which can be applied with consumers. Although this evolution is desirable to make temporal sensory measurements more representative of perception of consumers, it has probably been motivated more by practicalities than by scientific considerations as rapid methods compromise between costs, ecological validity and relevance of data collection.

Collecting data that meaningfully reflects a complex real-world experience through a simple task that consumers can perform is challenging. However, although they are increasingly used, there are still no clear recommendation on the application of the temporal sensory evaluation methods with consumers, and it is unknown to what extent the details of implementation can have an impact on the results. Indeed, it was evidenced with a thorough systematic review of the literature that despite 30 years of research involving temporal sensory measurements, knowledge on sensory evaluation methods has been built on a fragile foundation. It was therefore legitimate to ask the question "how appropriate are rapid sensory evaluation methods for measuring temporal perception of food products of consumers?".

The results presented in this research work demonstrated that it is possible to measure the temporal perception of consumers using rapid temporal sensory evaluation methods, but that one should not expect more from the methods than what they can offer given the complexity of the phenomenon studied. Indeed, the search for the most discriminative method has led to neglect other fundamental aspects of sensory measurements: reliability and validity. It has been shown that when considering these aspects, any gain in temporal resolution or in discrimination capacity is accompanied by an increased risk of overinterpretation of the results.

Although there is no serious problem when it comes to comparing food products, temporal sensory evaluation methods are also essential tools for understanding the fundamental physiological and cognitive processes underlying perception. The results presented in this research work provide new insights that may help both food manufacturers (who want to gain insight into their products) and researchers (interested in understanding perception) to

choose the most appropriate rapid method of temporal data collection depending on their question and implement it with consumers. Recommendations were also formulated to help analyse temporal data, interpret results and draw robust conclusions.

This research work has highlighted the difficulty of generalizing conclusions relating to sensory evaluation methods on the basis of a few studies. In the regrettable absence of large open databases of sensory studies that would allow meta-analyses to be carried out and perhaps help provide more definitive answers, all the datasets collected have been made public to contribute to the promotion of the FAIR principles. Although this was not the main objective of this work, a template for sharing data accompanied by guidelines for reporting research results on the topic in a more exhaustive way have been proposed.

V.2. Recommendations for future research

In this research work, it was chosen to focus on the comparison of qualitative measurements collected with consumers, in continuous time or by period. However, it would have been interesting to extend the comparisons to discrete or static measurements collected with trained panels to investigate further the resolution of temporal methods depending on their implementation conditions.

Following the scoping review, it could be possible to go further in the extraction of knowledge based on published results. A data-oriented approach based on meta-analyses could help determine (if they exist) generic food category temporal profiles or temporal signatures relating to sensory modalities (texture, basic taste, aromas, etc.). Moreover, some research gaps identified in the review have not been addressed in this manuscript. In particular, the question of the minimum number of consumers required to allow the measurement of small amplitude temporal differences is crucial. Future work should focus on this question, as well as on how to account for individual perceptual patterns rather than considering “average” perception (which may be no one's perception). This could lead to questions related to consumer segmentation based on sensory perception, in continuation of the work outlined by Cardot et al. (2019) and Peltier et al. (2023) based on Markovian modelling and categorical functional data analysis.

The framework proposed for the statistical analysis of temporal data has not been extensively tested, nor compared to the more sophisticated and integrative approaches just cited above. In this sense, it could be interesting to investigate to what extent these recently introduced statistical approaches provide usable information under the constraint of a low number of observations. It is to be noted that statistic articles are few cited compared to others (see section I.3), thus these methods are not well known by the food science community and their use remains confidential. Additional efforts should be made to make them more understandable and accessible to the food science community by sharing the codes using user-friendly interfaces.

As mentioned previously, the hedonic data collected in the studies presented were not used. Further studies could reinvestigate with a critical eye if temporal aspects really contribute to perception of product quality and preferences. The way the determinants of liking have been studied so far with temporal methods does not directly take into account the dynamics of perception. As the temporal patterns of products are often similar, it is likely that the conclusions obtained would have been the same if the data had been collected with a static method. Another way of looking at this question would consist in testing whether the same sensory attributes perceived at different times (for example the beginning and the end of tasting) by different subjects have a different impact on liking.

More generally, other temporal aspects could also be investigated. The question of multi-intake measurements has been largely studied, but as with the temporal methods presented in this research work, there remain methodological issues to address. Indeed, probably for practical reasons, multi-intake measurements have most often been carried out on a limited number of intakes, not necessarily representative of a full portion. It will therefore be necessary to find a compromise on how to evaluate the perception of a complete portion in the most natural way possible while maintaining a minimum of experimental control to obtain meaningful data.

A completely different type of measurement could be considered to explore perception on a broader temporal scale. It would consist in asking consumers what they remember about different food products (sensory attributes and liking) a few minutes/hours/days after tasting them. This “sensory image” could be more appropriate for understanding consumers' food choices which are potentially made on the basis of memories and impressions of tastings.

V.3. Use of temporal sensory evaluation methods with consumers: a position paper

This article was published in Current Opinion in Food Science:
<https://doi.org/10.1016/j.cofs.2023.101102>

M.V. Galmarini: Conceptualization, Writing - original draft; P. Schlich: Conceptualization, Writing - original draft, Supervision; M. Visalli: Conceptualization, Writing - original draft.

Temporal sensory evaluation methods aim to describe the dynamics of product perception. These methods were developed and implemented with trained panels, but they are increasingly used with consumers. However, this shift has probably been guided more by practical aspects than by scientific considerations, and the limits of application of temporal methods with consumers are not well-documented. This article discusses some of these limitations, presents recent developments looking to overcome them, and makes recommendations on the choice and implementation of methods as well as on the analysis of temporal data and their interpretation. This contribution relies on recent methodological works: a systematic scoping review on temporal methods and several articles comparing methods using controlled stimuli and commercial products with various levels of complexity.

From analytical to “rapid” temporal methods applied with consumers

Perception is a dynamic process, as a consequence numerous sensory evaluation methods have been developed attempting to capture the temporal dimension in perception. The first temporal methods were based on the same premise as the static descriptive methods: using intensity scales to quantify a perceived attribute over time (see Visalli & Galmarini, 2023 for a complete review). Then, as with static descriptive methods, temporal methods shifted towards qualitative scales, reducing or even eliminating the need for panel training. This resulted in rapid temporal descriptive methods which can be applied with consumers such as Temporal Dominance of Sensations (TDS, Pineau et al., 2009) and Temporal Check-All-That-Apply (TCATA, Castura, Antúnez, et al., 2016).

TDS and TCATA were first developed and implemented with trained panels, validated by comparison to analytical methods, and then transposed to consumer panels. However, to date, about two TDS and TCATA studies out of three (in research) are still implemented with trained or semi-trained panels.



Figure 98. Evolution of the use of multi-attribute temporal sensory evaluation methods by consumers and trained (includes semitrained as well) panels over the years.

TDS asks panellists to choose from a list of attributes which one is dominant at each moment of the tasting, while TCATA requires to check all the applicable attributes and uncheck them whenever they are no longer perceived. TDS and TCATA were also used to evaluate perception

changes over successive intakes (Cordelle et al., 2022; Machado et al., 2023; Silva et al., 2019; van Bommel, Stieger, Boelee, et al., 2019; Wakihira et al., 2020; Weerawarna et al., 2021).

Both methods present constraints which researchers tried to tackle by developing new variants. When evaluating a food product, some sensory modalities can be easier to identify than others (e.g. texture over flavor). “By modality” variants of TDS (Lesme et al., 2020; Rizo, Jimenez-Pérez, et al., 2019) and TCATA (Dietz, Cook, et al., 2022) alternate evaluations, restricting the list of attributes to a single sensory modality at each evaluation. With Dual-TDS (Pittari et al., 2022; Schlich, 2017), two lists of attributes are presented simultaneously (one by modality), allowing to track two dominant attributes in a same evaluation contrarily to “by-modality” variants that require doubling the number of intakes. Two variants of TCATA propose to address the difficulty of unselecting attributes. With TCATA-Fading (Rizo et al., 2020), applicable attributes are automatically unselected after a predefined duration. With "Discrete-time TCATA" (Visalli, Wakihira, et al., 2022), successive CATA tasks are performed at several fixed times during an intake.

Aiming at simplifying more the task of consumers, retrospective methods ask consumers to report at the end of the tasting the sensations they perceived during three periods of interest. These moments are the attack, the evolution, and the finish of the perception, which gives origin to the method's name: AEF-D (Dominance) (Visalli et al., 2020) and AEF-A (Applicability) (Visalli, Wakihira, et al., 2022). AEF-A has also been tested using RATA scales (Visalli, Béno, Nicolle, et al., 2023), and both AEF-D and AEF-A were applied using free comment (FC) (Mahieu et al., 2020b; Visalli, Dubois, Schlich, et al., 2023) instead of a predefined list of attributes. Two other methods rely on similar principles: Quessence (Jeltema et al., 2020) and Free-Temporal Order of Sensations (Free-TOS, Carrillo et al., 2021).

The need to evolve from analytical evaluations with trained panels to measurements involving consumers has been accelerated by the urge to consider the ecological validity of sensory measurements and to collect data more representative of perception of final users (Meiselman et al., 2022). Working directly with “large” panels presents several advantages. This makes possible to identify differences in perception due to physiological or individual characteristics such as expertise (Weerawarna et al., 2021), saliva (Goza et al., 2022), age, gender or ethnicity (Aguayo-Mendoza et al., 2020). It could help to understand if temporal aspects contribute to perception of product quality and preferences and explain consumers’

behaviour (Castura, 2018), by asking sensory description and preferences to the same subjects (Greis et al., 2020; Machado et al., 2023; Nguyen et al., 2023; Paglarini et al., 2020; Wakihira et al., 2023; Wang et al., 2021; Weerawarna et al., 2023) possibly out of lab (Dinnella et al., 2022).

Temporal methods can be used with consumers, but their limits of application are unclear

The transition to rapid measurements has been probably guided more by practical aspects than by scientific considerations. Regardless of their ongoing use, there are still no clear recommendation on the application of the temporal methods with consumers (Visalli & Galmarini, 2023), and it is unknown to what extent the details of implementation can impact results. Indeed, obtaining data that meaningfully reflects a complex real-world experience by means of a simple task is challenging (Castura, 2018), and rapid methods compromise between costs, ecological validity and relevance of data collection.

The continuous innovations around temporal methods suggest that there are still certain needs not met by the existing ones. However, these developments are supported by few teams (Visalli & Galmarini, 2023). In a context of academic competition (and sometimes related also to software promotion), there is a risk to push the use of temporal methods off limits that could lead to their misuse or an overinterpretation of the data. The legit quest for more ecological measurements should not make us forget that expectations on temporal measurements collected from consumers should be down to earth. Indeed, the real temporal resolution of methods (i.e. the minimum duration of statistically significant differences between times of selection/unselection of attributes that can be considered meaningful and interpretable, between two products on one attribute or between two attributes within a product) is not documented (Visalli, Mahieu, Peltier, et al., 2023). Thus, there is not always a guarantee that temporal measurements have added value compared with static ones (Velázquez et al., 2020; Vidal et al., 2019).

About the task: what is asked vs. what is done

TDS imposes a forced-choice of one descriptor over all those perceived at the same time. This could incite consumers to select the attribute they are most confident about, regardless of the definition they are given about “dominance” (Hutchings et al., 2022). Conversely to TDS,

TCATA invites consumers to make multiple choices, and selected applicable attributes are expected to be representative of all sensations perceived at any time. It is likely that, as with CATA (Jaeger et al., 2020), criteria chosen by consumers to determine what are applicable attributes are as subject-dependent as dominance with TDS. Other than the attribute choice, there is a time constraint imposed on the decision-making process of both methods which is inherent to the concurrent nature of the description. Though indirect, this is also a forced-choice constraint and it might explain why in practice dominance and applicability are not so far apart. A recent study on controlled temporal stimuli (Visalli, Béno, Nicolle, et al., 2023) confirmed that whatever the concept, consumers tend to report what they perceived as a function of several parameters including sensory intensity, duration of stimulation, familiarity/easiness to identify the sensation and novelty, this last parameter (defined as “not perceived before in the same intake”) being by far the most important.

The biggest difference between TDS and TCATA lies in the durations of dominance/applicability of attributes (Visalli, Béno, Nicolle, et al., 2023). Durations of dominance are computed based on the time elapsed between the choice of one attribute and the next one. Duration of applicability of an attribute starts when the attribute is chosen and ends when consumers uncheck it or until the end of the evaluation. As noted above, TCATA results made researchers think that it was difficult for consumers to uncheck attributes that were no longer applicable. This can be explained by recent works demonstrating a limited capacity of gustatory working memory (Lim et al., 2022), which suggest that tracking the state of multiple sensations (more than three) is probably out of reach for most consumers. TCATA-Fading tries to limit the impact of forgetting to uncheck attributes, but in this way, the durations of applicability depend on a subjective choice from the experimenter. This variant showed to be more discriminative than TCATA (Rizo et al., 2020), but as a general rule, more discrimination does not necessarily imply more validity. What duration measures is unclear. Durations do not bring information on the dynamics of the changes which occur during the perception, which is the original goal of the temporal methods. Compared to a tasting where consumers just have to report when they do no longer perceive anything, TDS and TCATA total durations are 40% and 70% longer, respectively (Visalli, Mahieu, Peltier, et al., 2023). At panel level, mean durations of dominance/applicability of attributes are highly correlated with

citation rates, but poorly with durations of stimulations and concentrations of controlled temporal stimuli (Visalli, Béno, Nicolle, et al., 2023).

Retrospective temporal methods do not record durations and are not submitted to time limit constraints. This probably reduces the stress that some consumers can feel during concurrent measurements (Visalli, Wakihira, et al., 2022), but it makes the task more analytical and introduces new biases related to memory. In this sense, AEF-A is probably very close to CATA: consumers determining what they perceived, and afterwards affect each perception to the corresponding period(s) (attack, evolution or finish). As a drawback, consumers can check more attributes than they actually perceived to make sure they do not miss anything (some kind of acquiescence bias). In return, the main sources of heterogeneity among panellist (citation times and durations) are removed and the main within-product temporal pattern (with a coarse temporal resolution) is captured, as well as differences between products (Visalli et al., 2020; Visalli et al., 2023)

Recommendations about method choice and implementation

The choice of the most appropriate method necessarily involves a compromise between temporal resolution, descriptive and discriminative capacity, validity and reliability.

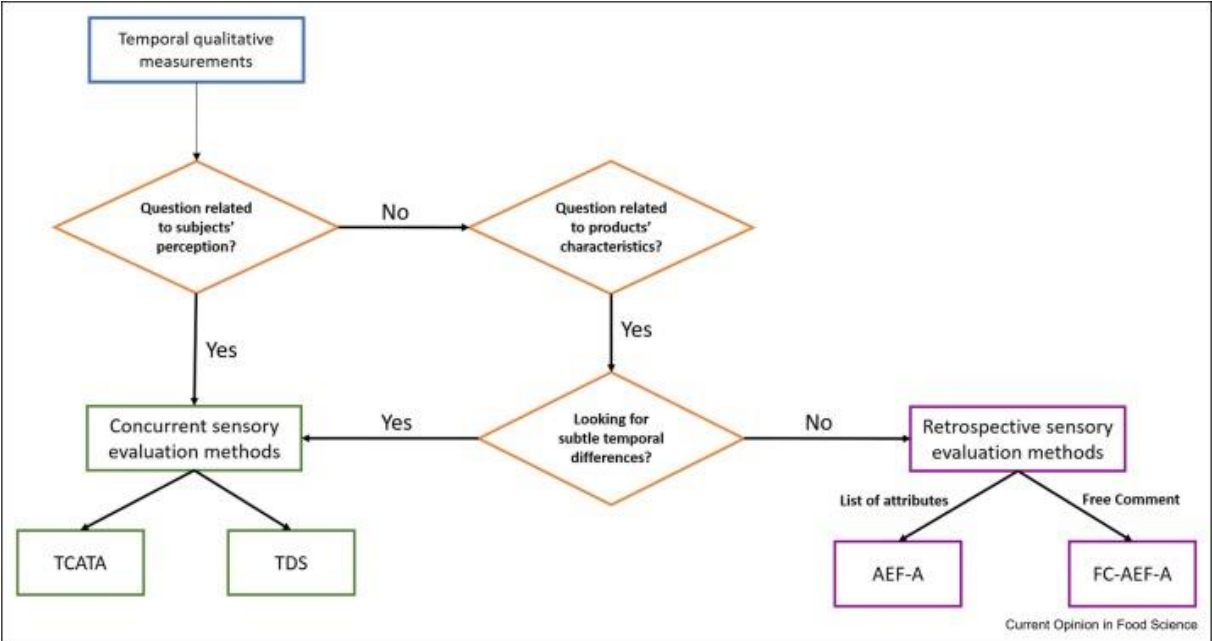


Figure 99. Flowchart for choosing the method of measuring consumer perception.

If the objective is to study subjects' temporal perception (e.g. to study physiological processes), concurrent methods should be favoured because they are closer to immediate perception and are more spontaneous.

If the objective is to study between-product temporal differences with samples expected to present small temporal differences, concurrent methods are needed but there is no guarantee that consumers will succeed in capturing differences related to temporal aspects, notably with complex products. TCATA and TDS will probably highlight the same main differences, TCATA being more discriminant but less reliable than TDS with complex products (Visalli et al., 2023).

If the objective is to study between-product differences and a fine temporal resolution is not necessary, retrospective methods can be considered as they are easy to use even out of the laboratory (no need for familiarization, briefing nor even software) (Visalli et al., 2020), they can be used with a large number of attributes or with FC and, worst case scenario, they will probably be equivalent to a regular CATA or FC. To limit the subjectivity in the interpretation, a protocol based on oral processing moments (e.g. after mouthing, after swallowing, last 10 s) can be proposed to define periods of interest.

Whatever the method, some implementation details can be considered to facilitate the task for consumers and the interpretation of the outcomes by the experimenter. The evaluation of a video (Velázquez et al., 2020) presenting known temporal differences can help consumers understand the temporal nature of the task. For TDS, to avoid dealing with subjective interpretations of dominances, a non-ambivalent instruction for participants could be simply to click on the attribute as soon as it is perceived (Hutchings et al., 2022). The sensory attributes should be explained to consumers, even if they seem simple. Examples of day-to-day products presenting the sensory characteristics can be provided, or even onomatopes for textures (Saita et al., 2021). Sensory attributes presenting too much similarity should be avoided, for example using meta-attributes (e.g. "crunchy/crispy") or high-level categories of aromas (e.g. using "citrus" instead of "lemon" and "orange").

The mean number of consumers in studies involving temporal methods is about 70 (Visalli & Galmarini, 2023). This number is probably sufficient to capture the main picture of the products, but a larger number is probably desirable to capture the movie of the perception and increase the confidence in conclusions related to temporal differences (Cardot et al., 2019; sOkamoto, 2021). Further research is needed to determine the appropriate number of

subjects regarding the complexity and the size of the differences of the evaluated products. Even if individual repeatability cannot be expected from consumers, it is recommended to replicate at least one sample to determine the level of noise in the data at panel level. If the replicated sample is served in last position to all panellists, then the data analysis can be conducted with or without the replicate.

About the data: what is collected vs. what is analysed

TDS and TCATA ask consumers to report events related to the presence or absence of dominant/applicable attributes. This sequence of events constitutes a stochastic process (Cardot et al., 2019), rarely analysed as such. Most often, these event data are represented as curves of dominance/applicability rates, as if they were continuous. However, it was observed that the temporal resolution of these methods was about four seconds, depending on the complexity of the products and the disagreement between panellists regarding temporal perception (Visalli, Béno, Nicolle, et al., 2023; Visalli, Mahieu, Peltier, et al., 2023).

A frequent data transformation is the time standardization that consists in aligning the beginning and end of the evaluation for all participants. Indeed, heterogeneity among consumers partially comes from delays in the time of the first citations, which are more likely uninformative. Indeed, these delays are more certainly due to individual differences in reaction times rather than to product signatures (Visalli, Béno, Nicolle, et al., 2023). We thus recommend to systematically remove this heterogeneity by subtracting the time of first citation to all citation times in a TDS/TCATA sequence. There are also differences among consumers in the total duration of perception. If we assume that there is a true sequence of sensations in a product, then we should get rid of this heterogeneity by doing the regular standardization in order to recover this sequence as much as possible at panel level. If we do not make that assumption, then we acknowledge the possibility of consumer segmentation in sequences of perception, as addressed by (Cardot et al., 2019; Peltier et al., 2023). With the classical TDS/TCATA curves and with a consumer panel, we recommend to keep working with the regular standardization.

The golden criterion for evaluating the relevance of results is most often restricted to the examination of the level of discrimination between products. The source of the observed differences (citations, times of citations, durations) is rarely examined, and the reliability of

temporal measurements is almost never considered (Visalli & Galmarini, 2023). Reported times of perception are delayed from actual perception, and include individual variations that are not considered with the plot of citation rates over times (TDS or TCATA curves) (Visalli, Mahieu, Peltier, et al., 2023). However, though it relies on the assumption that moments of perceptions are aligned among consumers, this representation of TDS/TCATA data is the most frequently used. As a consequence, due to the transformation of event data in time series, unexpected attributes can peak at relatively high citation rates by chance (Visalli, Béno, Nicolle, et al., 2023). Moreover, significant differences in citation rates between products observed on TDS/TCATA curves are differences in agreement at a given time, not necessarily representative of differences in temporal perception (Visalli, Mahieu, Peltier, et al., 2023). Citation rates are not at all quantifications of the intensity of the perception of sensory attributes (Visalli, Béno, Nicolle, et al., 2023), but rather reflections of the level of confidence in the dominance or applicability of attributes at a given time. For these reasons, TDS/TCATA curves should be considered as exploratory data analyses, and visually comparing two sets of curves is inappropriate to conclude on differences between products (Meyners, 2020b).

Multivariate representations (PCA, CA or CVA) are little informative about the source of temporal differences. Indeed they summarize the temporal perception in a single variable (most often duration or average citation rate) and they mainly capture differences in citations of attributes (similar to CATA) not related to temporal aspects of perception (Visalli et al., 2023). Trajectory maps with variables aggregated by periods are more representative of the dynamics of perception, but they do not consider the different nature of within and between product sources of variation. To model the evolution of citation rates over time, some authors used generalized linear models for binomial data followed by analysis of deviance (Weerawarna et al., 2021; Weerawarna et al., 2023). This approach avoids the issue related to the lack of independence among observations, but as trajectory maps, it requires data to be split into periods of subjective sizes.

Recommendations about data analysis and interpretation

It is important to examine results beyond face validity (Meyners, 2020b; Visalli, Béno, Nicolle, et al., 2023) to conclude on temporal differences. Prior to data analysis, it is recommended to examine panellist behaviour, by visualizing the distributions of variables such as delay to first citation, total duration and number of used attributes (Hutchings et al., 2022). Computing individual similarity indexes (i.e. distances between temporal sequences, expected null between replicated samples) can then help to spot panellists who are not in agreement with the panel (Frascolla et al., 2023) or who substantially bias the average responses (Natsume et al., 2023).

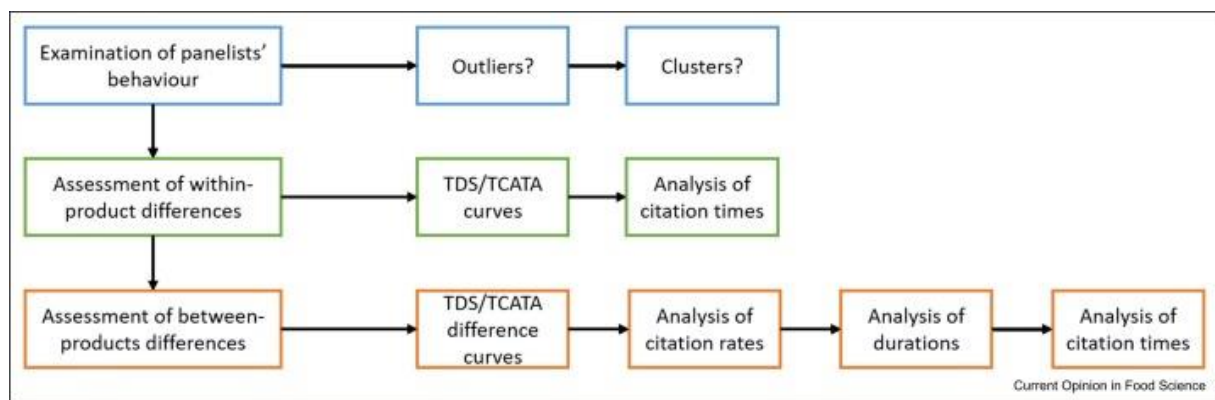


Figure 100. Flowchart for data analyses.

To objectify the interpretation of within-product temporal evolution and test if attributes are perceived at same or different times, it is possible to complete TDS/TCATA curves by confirmatory data analyses considering confidence intervals around times of citations (Visalli, Mahieu, Peltier, et al., 2023). To draw more robust conclusions, the Pareto principle can be applied, and attributes cited by less than 20% of consumers can be ignored (Visalli et al., 2023).

To compare pairs of products in a holistic way, difference curves are appropriate confirmatory analyses. However, it should be ensured that the attribute considered was perceived (with a sufficiently high citation rate) in at least one of the two products. For TDS this means that the dominance rate should be significant in at least one of the two products, unfortunately that cannot be done with applicability curves for which no significance level can be defined (except if applying the 20% threshold suggested above). Moreover, significance of a dominance rate or significance of a difference between two dominance/applicability rates should not be

considered when it lasts for less than the expected resolution of the method; in general, we suggest three seconds.

If understanding the source of differences between products is of interest, a multi-step approach can be performed. First, differences in perception of attributes (without regard to the temporal dimension of data) can be assessed using an univariate linear model based on count data (Visalli et al., 2023) or a multivariate multiple-response Correspondence Analysis (mrCA, Mahieu et al., 2020b). Then, durations of perception of attributes can be analysed in the same way replacing mrCA by CVA, and conclusions compared to those obtained with citations to determine if durations bring additional information to citations. Finally, the moments when the attributes are perceived can be compared between products to determine if differences are due to offsets in perception of specific attributes (Visalli et al., 2023). Whatever the variable analysed, reporting the effect size of significant differences can help interpreting results (Johnson et al., 2020).

A major drawback of this multi-step approach is the accumulation of statistical tests. New types of statistical analyses based on semi-Markov models (Frascolla et al., 2022) allow to properly test equality of two models drawn from two sets of TDS sequences. Recently, Categorical Functional Data Analysis opened a new integrative way of analysing TDS data, but does not apply to TCATA, nor offer direct statistical inference yet (Peltier et al., 2023). However, these frameworks of data analysis need to be applied to numerous datasets in order to demonstrate that their outcomes are meaningful in terms of sensory interpretation.

Conclusion

It is possible to measure the temporal perception of consumers with qualitative sensory analysis methods, but one should not expect more from the methods than they can offer given the complexity of temporal data and perception process itself. It seems that the minimum these methods give are the sequences according to which attributes are perceived. There is no ultimate method, and every gain in temporal resolution or discrimination is probably at the cost of a loss of validity and/or reliability. Rather than developing new methods, we should now make sure the existing ones are used correctly and encourage the published data to be FAIR (Visalli et al., 2023) in such a way that meta-analysis of large amounts of temporal datasets could bring new knowledge about the power and the limits of these methods.

References

- Abraham, I., & MacDonald, K. (2011). Quasi-experimental research. In *Encyclopedia Of Nursing Research*. Springer Publishing Company.
- Adams, J., Williams, A., Lancaster, N., & Foley, M. (2007). Advantages and uses of check-all-that-apply response compared to traditional scaling of attributes for salty snacks. 7th Pangborn Sensory Science Symposium. Minneapolis, USA, 12–16 August, 2007.
- Agresti, A. (2013). *Categorical data analysis*. Wiley-Interscience.
- Aguayo-Mendoza, M. G., Chatonidi, G., Piqueras-Fizman, B., & Stieger, M. (2021). Linking oral processing behavior to bolus properties and dynamic sensory perception of processed cheeses with bell pepper pieces. *Food Quality and Preference*, 88. <https://doi.org/10.1016/J.FOODQUAL.2020.104084>
- Aguayo-Mendoza, M. G., Martinez-Almaguer, E. F., Piqueras-Fizman, B., & Stieger, M. (2020). Differences in oral processing behavior of consumers varying in age, gender and ethnicity lead to changes in bolus properties but only to small differences in dynamic texture perception of sausages. *Food and Function*, 11(11), 10022–10032. <https://doi.org/10.1039/D0FO01835J>
- Agudelo, A., Varela, P., & Fizman, S. (2015). Methods for a deeper understanding of the sensory perception of fruit fillings. *Food Hydrocolloids*, 46, 160–171. <https://doi.org/10.1016/J.FOODHYD.2014.12.024>
- Aguinis, H., Vassar, M., & Wayant, C. (2021). On reporting and interpreting statistical significance and p values in medical research. *BMJ Evidence-Based Medicine*, 26(2), 39. <https://doi.org/10.1136/BMJEBM-2019-111264>
- Akdag, H. (2019). Identifying quality criteria of a scientific research adopted by academic community: a case study. *International Journal Of Eurasia Social Sciences*, 10(36), 516–527. <https://doi.org/10.35826/IJOESS.2479>
- Albert, A., Salvador, A., Schlich, P., & Fizman, 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. <https://doi.org/10.1016/J.FOODQUAL.2011.10.003>
- Alcaire, F., Antúnez, L., Vidal, L., Zorn, S., Giménez, A., Castura, J. C., & Ares, G. (2017). Comparison of static and dynamic sensory product characterizations based on check-all-that-apply questions with consumers. *Food Research International*, 97, 215–222. <https://doi.org/10.1016/J.FOODRES.2017.04.012>
- Aldrovandi, S., Poirier, M., Kusev, P., & Ayton, P. (2015). Retrospective evaluations of sequences: Testing the predictions of a memory-based analysis. *Experimental Psychology*, 62(5), 320–334. <https://doi.org/10.1027/1618-3169/A000301>
- Allès, B., Péneau, S., Kesse-Guyot, E., Baudry, J., Hercberg, S., & Méjean, C. (2017). Food choice motives including sustainability during purchasing are associated with a healthy dietary

pattern in French adults. *Nutrition Journal* 2017 16:1, 16(1), 1–12. <https://doi.org/10.1186/S12937-017-0279-9>

Amerine, M. A. Edward B. (1976). *Wines, their sensory evaluation*. 230.

Anandharamakrishnan, C., Sethupathy, P., Sivakamasundari, S. K., & Moses, J. A. (2021). Effect of varietal differences on the oral processing behavior and bolus properties of cooked rice. *International Journal of Food Engineering*, 17(3), 177–188. <https://doi.org/10.1515/IJFE-2020-0097>

Andersen, C. A., Alfine, L., Ohla, K., & Höchenberger, R. (2019). A new gustometer: Template for the construction of a portable and modular stimulator for taste and lingual touch. *Behavior Research Methods*, 51(6), 2733–2747. <https://doi.org/10.3758/S13428-018-1145-1>

Andrade, A. C., Martins, M. B., Rodrigues, J. F., Coelho, S. B., Pinheiro, A. C. M., & Bastos, S. C. (2019). Effect of different quantities of miracle fruit on sour and bitter beverages. *LWT*, 99, 89–97. <https://doi.org/10.1016/J.LWT.2018.09.054>

Antúnez, L., Giménez, A., Vidal, L., & Ares, G. (2018). Partial replacement of NaCl with KCl in bread: Effect on sensory characteristics and consumer perception. *Journal of Sensory Studies*, 33(5). <https://doi.org/10.1111/JOSS.12441>

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 3. <https://doi.org/10.1037/AMP0000191>

Appelqvist, I. A. M., Poelman, A. A. M., Cochet-Broch, M., & Delahunty, C. M. (2016). Impact of model fat emulsions on sensory perception using repeated spoon to spoon ingestion. *Physiology and Behavior*, 160, 80–86. <https://doi.org/10.1016/J.PHYSBEH.2016.03.035>

Ares, G., & Jaeger, S. R. (2015). Check-all-that-apply (CATA) questions with consumers in practice: experimental considerations and impact on outcome. *Rapid sensory profiling techniques and related methods: applications in new product development and consumer research*, 227–245. <https://doi.org/10.1533/9781782422587.2.227>

Ares, G. (2015). Methodological challenges in sensory characterization. *Current Opinion in Food Science*, 3, 1–5. <https://doi.org/10.1016/J.COFS.2014.09.001>

Ares, G. (2018). Methodological issues in cross-cultural sensory and consumer research. *Food Quality and Preference*, 64, 253–263. <https://doi.org/10.1016/J.FOODQUAL.2016.10.007>

Ares, G., Alcaire, F., Antúnez, L., Vidal, L., Giménez, A., & Castura, J. C. (2017). Identification of drivers of (dis)liking based on dynamic sensory profiles: Comparison of Temporal Dominance of Sensations and Temporal Check-all-that-apply. *Food Research International*, 92, 79–87. <https://doi.org/10.1016/J.FOODRES.2016.12.016>

Ares, G., Antúnez, L., Giménez, A., Roigard, C. M., Pineau, B., Hunter, D. C., & Jaeger, S. R. (2014). Further investigations into the reproducibility of check-all-that-apply (CATA) questions for sensory product characterization elicited by consumers. *Food Quality and Preference*, 36, 111–121. <https://doi.org/10.1016/J.FOODQUAL.2014.03.010>

- Ares, G., Bruzzone, F., Vidal, L., Cadena, R. S., Giménez, A., Pineau, B., Hunter, D. C., Paisley, A. G., & Jaeger, S. R. (2014). Evaluation of a rating-based variant of check-all-that-apply questions: Rate-all-that-apply (RATA). *Food Quality and Preference*, 36, 87–95. <https://doi.org/10.1016/J.FOODQUAL.2014.03.006>
- 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., Dauber, C., Fernández, E., Giménez, A., & Varela, P. (2014). Penalty analysis based on CATA questions to identify drivers of liking and directions for product reformulation. *Food Quality and Preference*, 32, 65–76. <https://doi.org/10.1016/J.FOODQUAL.2013.05.014>
- 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.
- Ares, G., Tárrega, A., Izquierdo, L., & Jaeger, S. R. (2014). Investigation of the number of consumers necessary to obtain stable sample and descriptor configurations from check-all-that-apply (CATA) questions. *Food Quality and Preference*, 31(1), 135–141. <https://doi.org/10.1016/J.FOODQUAL.2013.08.012>
- Ares, G., & Varela, P. (2017). Trained vs. consumer panels for analytical testing: Fueling a long lasting debate in the field. *Food Quality and Preference*, 61, 79–86. <https://doi.org/10.1016/J.FOODQUAL.2016.10.006>
- Arksey, H., & O'Malley, L. (2007). Scoping studies: towards a methodological framework, 8(1), 19–32. <https://doi.org/10.1080/1364557032000119616>
- Aschemann-Witzel, J., Ares, G., Thøgersen, J., & Monteleone, E. (2019). A sense of sustainability? – How sensory consumer science can contribute to sustainable development of the food sector. *Trends in Food Science and Technology*, 90, 180–186. <https://doi.org/10.1016/J.TIFS.2019.02.021>
- Asioli, D., Mignani, A., & Alfnes, F. (2021). Quick and easy? Respondent evaluations of the Becker–DeGroot–Marschak and multiple price list valuation mechanisms. *Agribusiness*, 37(2), 215–234. <https://doi.org/10.1002/AGR.21668>
- Auriema, B. E., Correa, F. J., Silva, R., Soares, P. T. S., Lima, A. L., Vidal, V. A. S., Raices, R. S. L., Pollonio, M. A. R., Luchese, R. H., Esmerino, E. A., & Mathias, S. P. (2022). Fat replacement by green banana biomass: Impact on the technological, nutritional and dynamic sensory profiling of chicken mortadella. *Food Research International*, 152. <https://doi.org/10.1016/J.FOODRES.2021.110890>
- Baker, A. K., Castura, J. C., & Ross, C. F. (2016). Temporal Check-All-That-Apply characterization of Syrah wine. *Journal of Food Science*, 81(6), S1521–S1529. <https://doi.org/10.1111/1750-3841.13328>
- Barbosa, M. de S. G., Francisco, J. S., dos Santos Scholz, M. B., Kitzberger, C. S. G., & Benassi, M. de T. (2019). Dynamics of sensory perceptions in arabica coffee brews with different

roasting degrees. *Journal of Culinary Science and Technology*, 17(5), 453–464. <https://doi.org/10.1080/15428052.2018.1489321>

Barker, S., & McSweeney, M. B. (2022). Sensory characterization of yellow pea and ground chicken hybrid meat burgers using static and dynamic methodologies. *Journal of Food Science*, 87(12), 5390–5401. <https://doi.org/10.1111/1750-3841.16380>

Barton, A., Hayward, L., Richardson, C. D., & McSweeney, M. B. (2020). Use of different panellists (experienced, trained, consumers and experts) and the projective mapping task to evaluate white wine. *Food Quality and Preference*, 83, 103900. <https://doi.org/10.1016/J.FOODQUAL.2020.103900>

Batista, N. N., Ramos, C. L., Dias, D. R., Pinheiro, A. C. M., & Schwan, R. F. (2016). The impact of yeast starter cultures on the microbial communities and volatile compounds in cocoa fermentation and the resulting sensory attributes of chocolate. *Journal of Food Science and Technology*, 53(2), 1101–1110. <https://doi.org/10.1007/S13197-015-2132-5>

Beaton, D., & Meyners, M. (2020). Powerful visualization of product-attribute associations for temporal data. *Food Quality and Preference*, 79, 103572. <https://doi.org/10.1016/J.FOODQUAL.2018.09.002>

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science*, 9(3), 226–232. <https://doi.org/10.1002/BS.3830090304>

Bécue-Bertaut, M., Álvarez-Esteban, R., & Pagès, J. (2008). Rating of products through scores and free-text assertions: Comparing and combining both. *Food Quality and Preference*, 19(1), 122–134. <https://doi.org/10.1016/J.FOODQUAL.2007.07.006>

Bell, A. (2013). Randomized or fixed order for studies of behavioral syndromes? *Behavioral Ecology*, 24(1), 16. <https://doi.org/10.1093/BEHECO/ARS148>

Bemfeito, R. M., Rodrigues, J. F., Silva, J. G. E., & Abreu, L. R. (2016). Temporal dominance of sensations sensory profile and drivers of liking of artisanal Minas cheese produced in the region of Serra da Canastra, Brazil. *Journal of Dairy Science*, 99(10), 7886–7897. <https://doi.org/10.3168/JDS.2016-11056>

Béno, N., Nicolle, L., & Visalli, M. (2023). A dataset of consumer perceptions of gustometer-controlled stimuli measured with three temporal sensory evaluation methods. *Data in Brief*, 109271. <https://doi.org/10.1016/J.DIB.2023.109271>

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. <https://doi.org/10.1016/j.foodqual.2019.103790>

Bhandari, P. (2022a). Construct Validity. Scribbr. <https://www.scribbr.com/methodology/construct-validity/>

Bhandari, P. (2022b). What Is Face Validity?. Scribbr. <https://www.scribbr.com/methodology/face-validity/>

- Bianchi, I., Branchini, E., Torquati, S., Fermani, A., Capitani, E., Barnaba, V., Savardi, U., & Burro, R. (2021). Non experts' understanding of terms frequently used by experts to describe the sensory properties of wine: An investigation based on opposites. *Food Quality and Preference*, 92, 104215. <https://doi.org/10.1016/J.FOODQUAL.2021.104215>
- Bikos, D., Samaras, G., Cann, P., Masen, M., Hardalupas, Y., Charalambides, M. N., Hartmann, C., German, J., & Vieira, J. (2022). Effect of structure on the mechanical and physical properties of chocolate considering time scale phenomena occurring during oral processing. *Food Structure*, 31. <https://doi.org/10.1016/J.FOOSTR.2021.100244>
- Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta Psychologica*, 134(3), 330–343. <https://doi.org/10.1016/J.ACTPSY.2010.03.006>
- Boinbaser, L., Parente, M. E., Castura, J. C., & Ares, G. (2015). Dynamic sensory characterization of cosmetic creams during application using Temporal Check-All-That-Apply (TCATA) questions. *Food Quality and Preference*, 45, 33–40. <https://doi.org/10.1016/J.FOODQUAL.2015.05.003>
- Bommel, R. van, Stieger, M. P. dr., Graaf, K. de P. dr., & Jager, G. D. C. (2020). Time will tell : dynamic sensory characteristics, hedonic perceptions and food-evoked emotions from first to last bite. Wageningen University.
- Bord, C., Guerinon, D., & Lebecque, A. (2019). Application of two sensory methods to investigate the impact of heating on the flavor perception of a French blue cheese. *Journal of Sensory Studies*, 34(5). <https://doi.org/10.1111/JOSS.12509>
- Borgogno, M., Dinnella, C., Iaconisi, V., Fusi, R., Scarpaleggia, C., Schiavone, A., Monteleone, E., Gasco, L., & Parisi, G. (2017). Inclusion of *Hermetia illucens* larvae meal on rainbow trout (*Oncorhynchus mykiss*) feed: effect on sensory profile according to static and dynamic evaluations. *Journal of the Science of Food and Agriculture*, 97(10), 3402–3411. <https://doi.org/10.1002/JSFA.8191>
- Botha, J. J., Cannon, P., & Hort, J. (2021). Comparing a new rapid combined method (RapCoTT) with traditional approaches for phenotyping thermal taste. *Physiology & Behavior*, 238, 113482. <https://doi.org/10.1016/J.PHYSBEH.2021.113482>
- Bouteille, R., Cordelle, S., Laval, C., Tournier, C., Lecanu, B., This, H., & Schlich, P. (2013). Sensory exploration of the freshness sensation in plain yoghurts and yoghurt-like products. *Food Quality and Preference*, 30(2), 282–292. <https://doi.org/10.1016/J.FOODQUAL.2013.06.012>
- Bradbury, J. W., & Vehrencamp, S. L. (1998). Principles of animal communication.
- Braghieri, A., Piazzolla, N., Galgano, F., Condelli, N., De Rosa, G., & Napolitano, F. (2016). Effect of preservative addition on sensory and dynamic profile of Lucanian dry-sausages as assessed by quantitative descriptive analysis and temporal dominance of sensations. *Meat Science*, 122, 68–75. <https://doi.org/10.1016/J.MEATSCI.2016.07.020>

- Brand, A., Allen, L., Altman, M., Hlava, M., & Scott, J. (2015). Beyond authorship: attribution, contribution, collaboration, and credit. *Learned Publishing*, 28(2), 151–155. <https://doi.org/10.1087/20150211>
- Breslin, P. A. S. (2013). An evolutionary perspective on food and human taste. *Current Biology*, 23(9), R409–R418. <https://doi.org/10.1016/J.CUB.2013.04.010>
- Brock, J. (2019). “A love letter to your future self”: What scientists need to know about FAIR data. *Nature Index*. <https://www.nature.com/nature-index/news/what-scientists-need-to-know-about-fair-data>
- Bruwer, J., & Buller, C. (2013). Product involvement, brand loyalty, and country-of-origin brand preferences of Japanese wine consumers, 24(1), 38–58. <https://doi.org/10.1080/09571264.2012.717221>
- Bruwer, J., Burrows, N., Chaumont, S., Li, E., & Saliba, A. (2014). Consumer involvement and associated behaviour in the UK high-end retail off-trade wine market, 24(2), 145–165. <https://doi.org/10.1080/09593969.2013.839464>
- Bruzzone, F., Ares, G., & Giménez, A. (2012). Consumers’ texture perception of milk desserts. II – Comparison with trained assessor’s data. *Journal of Texture Studies*, 43(3), 214–226. <https://doi.org/10.1111/J.1745-4603.2011.00332.X>
- Bruzzone, F., Ares, G., & Giménez, A. (2013). Temporal aspects of yoghurt texture perception. *International Dairy Journal*, 29(2), 124–134. <https://doi.org/10.1016/J.IDAIRYJ.2012.10.012>
- Bull, S. P., Hong, Y., Khutoryanskiy, V. V., Parker, J. K., Faka, M., & Methven, L. (2017). Whey protein mouth drying influenced by thermal denaturation. *Food Quality and Preference*, 56, 233–240. <https://doi.org/10.1016/J.FOODQUAL.2016.03.008>
- Burse, K. M. M., Camacho, S., Knoop, J., & Bult, J. H. F. (2010). Sweet taste intensity is enhanced by temporal fluctuation of aroma and taste, and depends on phase shift. *Physiology and Behavior*, 101(5), 726–730. <https://doi.org/10.1016/J.PHYSBEH.2010.08.014>
- Busch, J. L. H. C., Tournier, C., Knoop, J. E., Kooyman, G., & Smit, G. (2009). Temporal contrast of salt delivery in mouth increases salt perception. *Chemical Senses*, 34(4), 341–348. <https://doi.org/10.1093/CHEMSE/BJP007>
- Cadoret, M., & Husson, F. (2013). Construction and evaluation of confidence ellipses applied at sensory data. *Food Quality and Preference*, 28(1), 106–115. <https://doi.org/10.1016/J.FOODQUAL.2012.09.005>
- Cairncross, S. E., & Sjostrom, L. B. (1950). Flavor profiles—A new approach to flavor problems. *Food Technology*, 4, 308–311.
- Callegaro, M., Murakami, M. H., Tepman, Z., & Henderson, V. (2015). Yes–no answers versus check-all in self-administered modes: a systematic review and analyses., 57(2), 203–224. <https://doi.org/10.2501/IJMR-2015-014A>
- Caltabiano, M. L., & Shellshear, J. (1998). Palatability versus healthiness as determinants of food preferences in young adults: a comparison of nomothetic and idiographic analytic

- approaches. *Australian and New Zealand Journal of Public Health*, 22(5), 547–551. <https://doi.org/10.1111/J.1467-842X.1998.TB01436.X>
- Câmara, A. K. F. I., Vidal, V. A. S., Santos, M., Bernardinelli, O. D., Sabadini, E., & Pollonio, M. A. R. (2020). Reducing phosphate in emulsified meat products by adding chia (*Salvia hispanica* L.) mucilage in powder or gel format: A clean label technological strategy. *Meat Science*, 163. <https://doi.org/10.1016/J.MEATSCI.2020.108085>
- Candau, J., & Wathelet, O. (2011). The categories of odours are they real categories? *Langages*, 181(1). <https://doi.org/10.3917/LANG.181.0037>
- 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.
- Carrillo, E., Laguna, L., Arancibia, C., & Tárrega, A. (2021). Rescuing flavor identity and dynamic perception in puréed dishes; a restructuring solution for the purée diet. *Foods* 2021, Vol. 10, Page 905, 10(4), 905. <https://doi.org/10.3390/FOODS10040905>
- Carroll, C., & Booth, A. (2015). Quality assessment of qualitative evidence for systematic review and synthesis: Is it meaningful, and if so, how should it be performed? *Research Synthesis Methods*, 6(2), 149–154. <https://doi.org/10.1002/JRSM.1128>
- Castura, J. C. (2020). Investigating temporal sensory data via a graph theoretic approach. *Food Quality and Preference*, 79. <https://doi.org/10.1016/J.FOODQUAL.2019.103787>
- Castura, J. C., Rutledge, D. N., Ross, C. F., & Næs, T. (2022). Discriminability and uncertainty in principal component analysis (PCA) of temporal check-all-that-apply (TCATA) data. *Food Quality and Preference*, 96. <https://doi.org/10.1016/J.FOODQUAL.2021.104370>
- Castura, J. C. (2018). Dynamics of consumer perception. *Methods in consumer research*, Volume 1: new approaches to classic methods, 211–240. <https://doi.org/10.1016/B978-0-08-102089-0.00009-1>
- Castura, J. C., Baker, A. K., & Ross, C. F. (2016). Using contrails and animated sequences to visualize uncertainty in dynamic sensory profiles obtained from temporal check-all-that-apply (TCATA) data. *Food Quality and Preference*, 54, 90–100. <https://doi.org/10.1016/J.FOODQUAL.2016.06.011>
- Castura, J. C., & Li, M. (2016). Using TDS dyads and other dominance sequences to characterize products and investigate liking changes. *Food Quality and Preference*, 47, 109–121. <https://doi.org/10.1016/J.FOODQUAL.2015.06.019>
- 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.
- Cerri, J., Thøgersen, J., & Testa, F. (2019). Social desirability and sustainable food research: A systematic literature review. *Food Quality and Preference*, 71, 136–140.

- Chadha, D., Hamid, N., Kantono, K., & Marsan, M. (2022). Changes in temporal sensory profile, liking, satiety, and postconsumption attributes of yogurt with natural sweeteners. *Journal of Food Science*, 87(7), 3190–3206. <https://doi.org/10.1111/1750-3841.16224>
- Charles, M., Endrizzi, I., Aprea, E., Zambanini, J., Betta, E., & Gasperi, F. (2017). Dynamic and static sensory methods to study the role of aroma on taste and texture: A multisensory approach to apple perception. *Food Quality and Preference*, 62, 17–30. <https://doi.org/10.1016/J.FOODQUAL.2017.06.014>
- Charles, M., Romano, A., Yener, S., Barnabà, M., Navarini, L., Märk, T. D., Biasoli, F., & Gasperi, F. (2015). Understanding flavour perception of espresso coffee by the combination of a dynamic sensory method and in-vivo nosespace analysis. *Food Research International*, 69, 9–20. <https://doi.org/10.1016/J.FOODRES.2014.11.036>
- Chen, J. (2020). It is important to differentiate sensory property from the material property. *Trends in Food Science & Technology*, 96, 268–270. <https://doi.org/10.1016/J.TIFS.2019.12.014>
- Cheong, J. N., Foster, K. D., Morgenstern, M. P., Grigor, J. M. V., Bronlund, J. E., Hutchings, S. C., & Hedderley, D. I. (2014). The application of temporal dominance of sensations (TDS) for oral processing studies: An initial investigation. *Journal of Texture Studies*, 45(6), 409–419. <https://doi.org/10.1111/JTXS.12091>
- Church, B. A., Jackson, B. N., Beran, M. J., & Smith, J. D. (2019). Simultaneous versus prospective/retrospective uncertainty monitoring: The effect of response competition across cognitive levels. *Journal of Experimental Psychology: Animal Learning and Cognition*, 45(3), 311–321. <https://doi.org/10.1037/XAN0000207>
- Clark, C. C., & Lawless, H. T. (1994). Limiting response alternatives in time-intensity scaling: an examination of the halo-dumping effect. *Chemical Senses*, 19(6), 583–594. <https://doi.org/10.1093/CHEMSE/19.6.583>
- Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368), 829–836. <https://doi.org/10.1080/01621459.1979.10481038>
- Cliceri, D., Aprea, E., Menghi, L., Endrizzi, I., & Gasperi, F. (2021). Variability in the temporal perception of polyphenol-related sensations in extra virgin olive oil and impact on flavor perception. *Food Quality and Preference*, 93. <https://doi.org/10.1016/J.FOODQUAL.2021.104249>
- Cliff, M., & Heymann, H. (1993). Development and use of time-intensity methodology for sensory evaluation: A review. *Food Research International*, 26(5), 375–385. [https://doi.org/10.1016/0963-9969\(93\)90081-5](https://doi.org/10.1016/0963-9969(93)90081-5)
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. *Statistical Power Analysis for the Behavioral Sciences*. <https://doi.org/10.4324/9780203771587>
- Cordelle, S., Redl, A., & Schlich, P. (2022). Sensory acceptability of new plant protein meat substitutes. *Food Quality and Preference*, 98. <https://doi.org/10.1016/J.FOODQUAL.2021.104508>

- Cornelis, M. C., Tordoff, M. G., El-Sohehy, A., & van Dam, R. M. (2017). Recalled taste intensity, liking and habitual intake of commonly consumed foods. *Appetite*, 109, 182–189. <https://doi.org/10.1016/J.APPET.2016.11.036>
- Corrêa Simioni, S. C., Ribeiro, M. N., de Souza, V. R., Nunes, C. A., & Pinheiro, A. C. M. (2018). Multiple-sip temporal dominance of sensations associated with acceptance test: a study on special beers. *Journal of Food Science and Technology*, 55(3), 1164–1174. <https://doi.org/10.1007/S13197-018-3032-2>
- Correia, E., Amorim, E., & Vilela, A. (2022). Structural Equation Modeling (SEM) and Temporal Dominance of Sensations (TDS) in the evaluation of DOC Douro red wine's sensory profile. *Foods*, 11(8). <https://doi.org/10.3390/FOODS11081168>
- Cosson, A., Souchon, I., Richard, J., Descamps, N., & Anne, S. E. (2020). Using multiple sensory profiling methods to gain insight into temporal perceptions of pea protein-based formulated foods. *Foods*, 9(8). <https://doi.org/10.3390/FOODS9080969>
- Coulon-Leroy, C., Symoneaux, R., Lawrence, G., Mehinagic, E., & Maitre, I. (2017). Mixed Profiling: A new tool of sensory analysis in a professional context. Application to wines. *Food Quality and Preference*, 57, 8–16. <https://doi.org/10.1016/J.FOODQUAL.2016.11.005>
- Crepalde, L. T., Vidigal, M. C. T. R., Carneiro, J. de D. S., & Minim, V. P. R. (2021). Integration of the optimized descriptive profile and temporal dominance of sensations methodologies. *Journal of Sensory Studies*, 36(3). <https://doi.org/10.1111/JOSS.12651>
- Criado, C., Chaya, C., Fernández-Ruiz, V., Álvarez, M. D., Herranz, B., & Pozo-Bayón, M. Á. (2019). Effect of saliva composition and flow on inter-individual differences in the temporal perception of retronasal aroma during wine tasting. *Food Research International*, 126. <https://doi.org/10.1016/J.FOODRES.2019.108677>
- Croijmans, I., Hendrickx, I., Lefever, E., Majid, A., & van Den Bosch, A. (2019). Uncovering the language of wine experts. *Natural Language Engineering*, 26(5), 511–530. <https://doi.org/10.1017/S1351324919000500>
- Croijmans, I., & Majid, A. (2016). Not all flavor expertise is equal: the language of wine and coffee experts. *PLOS ONE*, 11(6), e0155845. <https://doi.org/10.1371/JOURNAL.PONE.0155845>
- Da Silva, T. L. T., De Souza, V. R., Pinheiro, A. C. M., Nunes, C. A., & Freire, T. V. M. (2014). Equivalence salting and temporal dominance of sensations analysis for different sodium chloride substitutes in cream cheese. *International Journal of Dairy Technology*, 67(1), 31–38. <https://doi.org/10.1111/1471-0307.12100>
- Daniel, T. A., & Katz, J. S. (2018). Primacy and recency effects for taste. *Journal of Experimental Psychology: Learning Memory and Cognition*, 44(3), 399–405. <https://doi.org/10.1037/XLM0000437>
- Danner, L., Johnson, T. E., Ristic, R., Meiselman, H. L., & Bastian, S. E. P. (2020). Consumption context effects on fine wine consumer segments' liking and emotions. *Foods*, 9(12). <https://doi.org/10.3390/FOODS9121798>

- Daudt, H. M. L., Van Mossel, C., & Scott, S. J. (2013). Enhancing the scoping study methodology: A large, inter-professional team's experience with Arksey and O'Malley's framework. *BMC Medical Research Methodology*, 13(1), 1–9. <https://doi.org/10.1186/1471-2288-13-48/PEER-REVIEW>
- de Lavergne, M. D., Tournier, C., Bertrand, D., Salles, C., Van de Velde, F., & Stieger, M. (2016). Dynamic texture perception, oral processing behaviour and bolus properties of emulsion-filled gels with and without contrasting mechanical properties. *Food Hydrocolloids*, 52, 648–660. <https://doi.org/10.1016/J.FOODHYD.2015.07.022>
- De Loubens, C., Panouillé, M., Saint-Eve, A., Déléris, I., Tréléa, I. C., & Souchon, I. (2011). Mechanistic model of in vitro salt release from model dairy gels based on standardized breakdown test simulating mastication. *Journal of Food Engineering*, 105(1), 161–168. <https://doi.org/10.1016/J.JFOODENG.2011.02.020>
- de Medeiros, A. C., Tavares Filho, E. R., & Bolini, H. M. A. (2021). Temporal profile of low calorie lactose-free ice cream chocolate flavor: temporal dominance sensation and multiple time–intensity analysis. *Journal of Food Science and Technology*, 58(8), 3164–3173. <https://doi.org/10.1007/S13197-020-04819-2>
- de Oliveira Paula, M. M., Massingue, A. A., de Moura, A. P. R., de Deus Souza Carneiro, J., de Lemos Souza Ramos, A., & Ramos, E. M. (2021). Temporal dominance of sensations and check-all-that-apply analysis of restructured cooked hams elaborated with different salt content and pork quality meats. *Food Science and Technology International*, 27(1), 73–83. <https://doi.org/10.1177/1082013220932355>
- De Paula, D. G. C., Vasconcelos, C. M., Pereira, A. F., Da Cunha Quintão, A. L., Chaves, J. O., Parreiras, P. M., & Menezes, C. C. (2022). Maceration time over antioxidant activity, volatile compounds and temporal dominant sensation of fermented jabuticaba. *Food Science and Technology (Brazil)*, 42. <https://doi.org/10.1590/FST.43221>
- De Souza, V. R., Freire, T. V. M., Saraiva, C. G., De Deus Souza Carneiro, J., Pinheiro, A. C. M., & Nunes, C. A. (2013). Salt equivalence and temporal dominance of sensations of different sodium chloride substitutes in butter. *Journal of Dairy Research*, 80(3), 319–325. <https://doi.org/10.1017/S0022029913000204>
- Deegan, K. C., Heikintalo, N., Ritvanen, T., Putkonen, T., Rekonen, J., McSweeney, P. L. H., Alatossava, T., & Tuorila, H. (2013). Effects of low-pressure homogenisation on the sensory and chemical properties of Emmental cheese. *Innovative Food Science and Emerging Technologies*, 19, 104–114. <https://doi.org/10.1016/J.IFSET.2013.04.008>
- del Barrio-Galán, R., Medel-Marabolí, M., & Peña-Neira, Á. (2016). Effect of different ageing techniques on the polysaccharide and phenolic composition and sensorial characteristics of Chardonnay white wines fermented with different selected *Saccharomyces Cerevisiae* yeast strains. *European Food Research and Technology*, 242(7), 1069–1085. <https://doi.org/10.1007/S00217-015-2612-X>
- Delarue, J., Lawlor, J. Ben, & Rogeaux, M. (2015). Rapid sensory profiling techniques and related methods: applications in new product development and consumer research. *Rapid*

Sensory Profiling Techniques and Related Methods: Applications in New Product Development and Consumer Research, 1–555. <https://doi.org/10.1016/C2013-0-16502-6>

Déléris, I., Saint-Eve, A., Guo, Y., Lieben, P., Cypriani, M. L., Jacquet, N., Brunerie, P., & Souchon, I. (2011). Impact of swallowing on the dynamics of aroma release and perception during the consumption of alcoholic beverages. *Chemical Senses*, 36(8), 701–713. <https://doi.org/10.1093/CHEMSE/BJR038>

Delompré, T., Lenoir, L., Martin, C., Briand, L., & Salles, C. (2020). Characterizing the dynamic taste and retro-nasal aroma properties of oral nutritional supplements using temporal dominance of sensation and temporal check-all-that-apply methods. *Foods*, 9(10). <https://doi.org/10.3390/FOODS9101456>

Demeter, M., Jele, A., & Major, Z. B. (2021). The international development of open access publishing: a comparative empirical analysis over seven world regions and nine academic disciplines. *Publishing Research Quarterly*, 37(3), 364–383. <https://doi.org/10.1007/S12109-021-09814-9/FIGURES/4>

Demi, L. M., Taylor, B. W., Reading, B. J., Tordoff, M. G., & Dunn, R. R. (2021). Understanding the evolution of nutritive taste in animals: Insights from biological stoichiometry and nutritional geometry. *Ecology and Evolution*, 11(13), 8441–8455. <https://doi.org/10.1002/ECE3.7745>

den Boer, A., Boesveldt, S., & Lawlor, J. Ben. (2019). How sweetness intensity and thickness of an oral nutritional supplement affects intake and satiety. *Food Quality and Preference*, 71, 406–414. <https://doi.org/10.1016/J.FOODQUAL.2018.08.009>

Derks, E., Ramnarain, S., Zhang, T., van Doorn, R., Nijmeijer, M., & van den Berg, M. (2022). Visualizing dynamic (after)taste effects by means of time-discrete TCATA data analysis. *Journal of Sensory Studies*, 37(3). <https://doi.org/10.1111/JOSS.12737>

Devezeaux de Lavergne, M., van Delft, M., van de Velde, F., van Boekel, M. A. J. S., & Stieger, M. (2015). Dynamic texture perception and oral processing of semi-solid food gels: Part 1: Comparison between QDA, progressive profiling and TDS. *Food Hydrocolloids*, 43, 207–217. <https://doi.org/10.1016/J.FOODHYD.2014.05.020>

Devezeaux de Lavergne, M., Derks, J. A. M., Ketel, E. C., de Wijk, R. A., & Stieger, M. (2015). Eating behaviour explains differences between individuals in dynamic texture perception of sausages. *Food Quality and Preference*, 41, 189–200. <https://doi.org/10.1016/J.FOODQUAL.2014.12.006>

Devezeaux de Lavergne, M., Van de Velde, F., & Stieger, M. (2017). Bolus matters: the influence of food oral breakdown on dynamic texture perception. *Food & Function*, 8(2), 464–480. <https://doi.org/10.1039/C6FO01005A>

Di Monaco, R., Miele, N. A., Volpe, S., Picone, D., & Cavella, S. (2014). Temporal sweetness profile of MNEI and comparison with commercial sweeteners. *Journal of Sensory Studies*, 29(6), 385–394. <https://doi.org/10.1111/JOSS.12119>

- Di Monaco, R., Galiñanes Plaza, A., Miele, N. A., Picone, D., & Cavella, S. (2016). Temporal sweetness profile of MNEI protein in gelled model systems. *Journal of Sensory Studies*, 31(5), 382–392. <https://doi.org/10.1111/JOSS.12222>
- Di Monaco, R., Miele, N. A., Volpe, S., Masi, P., & Cavella, S. (2016). Temporal dominance of sensations and dynamic liking evaluation of polenta sticks. *British Food Journal*, 118(3), 749–760. <https://doi.org/10.1108/BFJ-07-2015-0236>
- 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.
- Dietz, C., Cook, D., Yang, Q., Wilson, C., & Ford, R. (2022). A TCATA by modality approach to study the multisensory temporal profile of hop bitter and flavour products applied in lager. *Food Quality and Preference*, 97, 104470. <https://doi.org/10.1016/J.FOODQUAL.2021.104470>
- Dietz, C., Yang, Q., & Ford, R. (2022). The impact of time standardising TCATA by modality data on the multisensory profile of beer. *Food Quality and Preference*, 98. <https://doi.org/10.1016/J.FOODQUAL.2021.104506>
- Dijksterhuis, G. B., & Piggott, J. R. (2000). Dynamic methods of sensory analysis. *Trends in Food Science & Technology*, 11(8), 284–290. [https://doi.org/10.1016/S0924-2244\(01\)00020-6](https://doi.org/10.1016/S0924-2244(01)00020-6)
- Dijksterhuis, G. B., & Eilers, P. (1997). Modelling time-intensity curves using prototype curves. *Food Quality and Preference*, 8(2), 131–140. [https://doi.org/10.1016/S0950-3293\(96\)00039-0](https://doi.org/10.1016/S0950-3293(96)00039-0)
- 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. <https://doi.org/10.1016/J.FOODQUAL.2013.04.006>
- 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. <https://doi.org/10.1016/J.FOODQUAL.2012.04.013>
- Dinnella, C., Pierguidi, L., Spinelli, S., Borgogno, M., Gallina Toschi, T., Predieri, S., Lavezzi, G., Trapani, F., Tura, M., Magli, M., Bendini, A., & Monteleone, E. (2022). Remote testing: Sensory test during Covid-19 pandemic and beyond. *Food Quality and Preference*, 96. <https://doi.org/10.1016/J.FOODQUAL.2021.104437>
- Dixon-Woods, M., Shaw, R. L., Agarwal, S., & Smith, J. A. (2004). The problem of appraising qualitative research. *BMJ Quality & Safety*, 13(3), 223–225. <https://doi.org/10.1136/QSHC.2003.008714>
- Djekic, I., Ilić, J., Chen, J., Djekic, R., Sołowiej, B. G., Vujadinović, D., & Tomasevic, I. (2021). Analysis of pungency sensation effects from an oral processing, sensorial and emotions detection perspective—case study with grilled pork meat. *Applied Sciences (Switzerland)*, 11(21). <https://doi.org/10.3390/APP112110459>

- Djekic, I., Ilic, J., Lorenzo, J. M., & Tomasevic, I. (2021). How do culinary methods affect quality and oral processing characteristics of pork ham? *Journal of Texture Studies*, 52(1), 36–44. <https://doi.org/10.1111/JTXS.12557>
- Djekic, I. V., Ilic, J. G., Sołowiej, B. G., Djekić, R. I., & Tomasevic, I. B. (2022). Application of food mechanics and oral processing in modelling first bite of grilled meat. *Journal of Food Quality*, 2022. <https://doi.org/10.1155/2022/9176628>
- Dong, M., Zhang, Y. Y., Huang, X. H., Xin, R., Dong, X. P., Konno, K., Zhu, B. W., Fisk, I., & Qin, L. (2022). Dynamic sensations of fresh and roasted salmon (*Salmo salar*) during chewing. *Food Chemistry*, 368. <https://doi.org/10.1016/J.FOODCHEM.2021.130844>
- Doyennette, M., Aguayo-Mendoza, M. G., Williamson, A. M., Martins, S. I. F. S., & Stieger, M. (2019). Capturing the impact of oral processing behaviour on consumption time and dynamic sensory perception of ice creams differing in hardness. *Food Quality and Preference*, 78, 103721. <https://doi.org/10.1016/J.FOODQUAL.2019.103721>
- Du, P., Yuan, H., Chen, Y., Zhou, H., Zhang, Y., Huang, M., Jiangfang, Y., Su, R., Chen, Q., Lai, J., Guan, L., Ding, Y., Hu, H., & Luo, J. (2023). Identification of key aromatic compounds in basil (*Ocimum L.*) using sensory evaluation, metabolomics and volatilomics analysis. *Metabolites*, 13(1), 85. <https://doi.org/10.3390/METABO13010085/S1>
- Duizer, L. M., Bloom, K., & Findlay, C. J. (1996). Dual-attribute time-intensity measurement of sweetness and peppermint perception of chewing gum. *Journal of Food Science*, 61(3), 636–638. <https://doi.org/10.1111/J.1365-2621.1996.TB13175.X>
- Duizer, L. M., Bloom, K., & Findlay, C. J. (1997). Dual-attribute time-intensity sensory evaluation: A new method for temporal measurement of sensory perceptions. *Food Quality and Preference*, 8(4), 261–269. [https://doi.org/10.1016/S0950-3293\(96\)00052-3](https://doi.org/10.1016/S0950-3293(96)00052-3)
- Dunn, P. K., & Smyth, G. K. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, 5(3), 236. <https://doi.org/10.2307/1390802>
- Dupas de Matos, A., Marangon, M., Magli, M., Cianciabella, M., Predieri, S., Curioni, A., & Vincenzi, S. (2019). Sensory characterization of cucumbers pickled with verjuice as novel acidifying agent. *Food Chemistry*, 286, 78–86. <https://doi.org/10.1016/J.FOODCHEM.2019.01.216>
- Emorine, M., Septier, C., Martin, C., Cordelle, S., Sémon, E., Thomas-Danguin, T., & Salles, C. (2021). Salt and aroma compound distributions influence flavour release and temporal perception while eating hot-served flans. *Molecules*, 26(5). <https://doi.org/10.3390/MOLECULES26051300>
- Engelen, L., & Van Der Bilt, A. (2008). Oral physiology and texture perception of semisolids. *Journal of Texture Studies*, 39(1), 83–113. <https://doi.org/10.1111/J.1745-4603.2007.00132.X>
- Erwig, M., & Gopinath, R. (2012). Explanations for regular expressions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 7212 LNCS, 394–408. https://doi.org/10.1007/978-3-642-28872-2_27/COVER/

- 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. <https://doi.org/10.1016/J.FOODRES.2017.09.012>
- Etaio, I., Meillon, S., Pérez-Elortondo, F. J., & Schlich, P. (2016). Dynamic sensory description of Rioja Alavesa red wines made by different winemaking practices by using Temporal Dominance of Sensations. *Journal of the Science of Food and Agriculture*, 96(10), 3492–3499. <https://doi.org/10.1002/JSFA.7533>
- Evangelista, S. R., da Cruz Pedrozo Miguel, M. G., de Souza Cordeiro, C., Silva, C. F., Marques Pinheiro, A. C., & Schwan, R. F. (2014). Inoculation of starter cultures in a semi-dry coffee (*Coffea arabica*) fermentation process. *Food Microbiology*, 44, 87–95. <https://doi.org/10.1016/J.FM.2014.05.013>
- Evangelista, S. R., Miguel, M. G. da C. P., Silva, C. F., Pinheiro, A. C. M., & Schwan, R. F. (2015). Microbiological diversity associated with the spontaneous wet method of coffee fermentation. *International Journal of Food Microbiology*, 210, 102–112. <https://doi.org/10.1016/J.IJFOODMICRO.2015.06.008>
- Evangelista, S. R., Silva, C. F., Miguel, M. G. P. da C., Cordeiro, C. de S., Pinheiro, A. C. M., Duarte, W. F., & Schwan, R. F. (2014). Improvement of coffee beverage quality by using selected yeasts strains during the fermentation in dry process. *Food Research International*, 61, 183–195. <https://doi.org/10.1016/J.FOODRES.2013.11.033>
- Feltrin, A. C., De Souza, V. R., Saraiva, C. G., Nunes, C. A., & Pinheiro, A. C. M. (2015). Sensory study of different sodium chloride substitutes in aqueous solution. *International Journal of Food Science and Technology*, 50(3), 730–735. <https://doi.org/10.1111/IJFS.12670>
- Fiches, G., Saint Eve, A., Jourden, S., Déléris, I., Brunerie, P., & Souchon, I. (2016). Temporality of perception during the consumption of French grape brandies with different aging times in relation with aroma compound release. *Flavour and Fragrance Journal*, 31(1), 31–40. <https://doi.org/10.1002/FFJ.3265>
- Fife, D. A., & Rodgers, J. L. (2021). Understanding the exploratory/confirmatory data analysis continuum: moving beyond the “replication crisis.” *American Psychologist*, 77(3), 453–466. <https://doi.org/10.1037/AMP0000886>
- Fischer, U., Boulton, R. B., & Noble, A. C. (1994). Physiological factors contributing to the variability of sensory assessments: Relationship between salivary flow rate and temporal perception of gustatory stimuli. *Food Quality and Preference*, 5(1–2), 55–64. [https://doi.org/10.1016/0950-3293\(94\)90008-6](https://doi.org/10.1016/0950-3293(94)90008-6)
- Fizman, S., & Tarrega, A. (2018). The dynamics of texture perception of hard solid food: A review of the contribution of the temporal dominance of sensations technique. *Journal of Texture Studies*, 49(2), 202–212. <https://doi.org/10.1111/JTXS.12273>

- Forde, C. G. (2016). Sensory science: From measuring perception to understanding ingestive behavior. Reference Module in Food Science. <https://doi.org/10.1016/B978-0-08-100596-5.03396-5>
- Foster, K. D., Grigor, J. M. V., Cheong, J. N., Yoo, M. J. Y., Bronlund, J. E., & Morgenstern, M. P. (2011). The role of oral processing in dynamic sensory perception. *Journal of Food Science*, 76(2), R49–R61. <https://doi.org/10.1111/J.1750-3841.2010.02029.X>
- Frank, D., Eyres, G. T., Piyasiri, U., Cochet-Broch, M., Delahunty, C. M., Lundin, L., & Appelqvist, I. M. (2015). Effects of agar gel strength and fat on oral breakdown, volatile release, and sensory perception using in vivo and in vitro systems. *Journal of Agricultural and Food Chemistry*, 63(41), 9093–9102. <https://doi.org/10.1021/ACS.JAFC.5B03441>
- Frascolla, C., Lecuelle, G., Schlich, P., & Cardot, H. (2022). Two sample tests for Semi-Markov processes with parametric sojourn time distributions: an application in sensory analysis. *Computational Statistics*, 37(5), 2553–2580. <https://doi.org/10.1007/S00180-022-01210-X>
- Frascolla, C., Visalli, M., Cardot, H., & Schlich, P. (2023). Indexes of individual repeatability and product discrimination in TDS and TCATA and their statistical inference. *Food Quality and Preference*, 110, 104929. <https://doi.org/10.1016/J.FOODQUAL.2023.104929>
- 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.
- Frost, S. C., Harbertson, J. F., & Heymann, H. (2017). A full factorial study on the effect of tannins, acidity, and ethanol on the temporal perception of taste and mouthfeel in red wine. *Food Quality and Preference*, 62, 1–7. <https://doi.org/10.1016/J.FOODQUAL.2017.05.010>
- Fryer, J. A., Collins, T. S., & Tomasino, E. (2021). Evaluation of different interstimulus rinse protocols on smoke attribute perception in wildfire-affected wines. *Molecules*, 26(18). <https://doi.org/10.3390/MOLECULES26185444>
- Furey, A. E., Hoeche, U., McLaughlin, C., & Noci, F. (2022). Incorporation of roe, milt and liver from plaice (*Pleuronectes platessa*), herring (*Clupea harengus*) and cod (*Gadus morhua*) in newly developed seafood Pâtés: Sensory evaluation by teenage consumers in Ireland and their attitudes to seafood. *International Journal of Gastronomy and Food Science*, 28. <https://doi.org/10.1016/J.IJGFS.2022.100524>
- Galiñanes Plaza, A., Delarue, J., & Saulais, L. (2019). The pursuit of ecological validity through contextual methodologies. *Food Quality and Preference*, 73, 226–247. <https://doi.org/10.1016/J.FOODQUAL.2018.11.004>
- Galmarini, M. V., Silva Paz, R. J., Enciso Choquehuanca, D., Zamora, M. C., & Mesz, B. (2021). Impact of music on the dynamic perception of coffee and evoked emotions evaluated by temporal dominance of sensations (TDS) and emotions (TDE). *Food Research International*, 150. <https://doi.org/10.1016/J.FOODRES.2021.110795>
- Galmarini, M. V., Dufau, L., Loiseau, A. L., Visalli, M., & Schlich, P. (2018). Wine and cheese: Two products or one association? a new method for assessing wine-cheese pairing. *Beverages*, 4(1). <https://doi.org/10.3390/BEVERAGES4010013>

- Galmarini, M. V., Loiseau, A. L., Debreyer, D., Visalli, M., & Schlich, P. (2017). Use of multi-intake Temporal Dominance of Sensations (TDS) to evaluate the influence of wine on cheese perception. *Journal of Food Science*, 82(11), 2669–2678. <https://doi.org/10.1111/1750-3841.13932>
- Galmarini, M. V., Loiseau, A. L., Visalli, M., & Schlich, P. (2016). Use of multi-intake Temporal Dominance of Sensations (TDS) to evaluate the influence of cheese on wine perception. *Journal of Food Science*, 81(10), S2566–S2577. <https://doi.org/10.1111/1750-3841.13500>
- Galmarini, M. V., Symoneaux, R., Visalli, M., Zamora, M. C., & Schlich, P. (2016). Could Time-Intensity by a trained panel be replaced with a progressive profile done by consumers? A case on chewing-gum. *Food Quality and Preference*, 48.
- 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. <https://doi.org/10.1016/j.foodqual.2016.01.011>
- Gao, J., Ong, J. J. X., Henry, J., & Zhou, W. (2017). Physical breakdown of bread and its impact on texture perception: A dynamic perspective. *Food Quality and Preference*, 60, 96–104. <https://doi.org/10.1016/J.FOODQUAL.2017.03.014>
- Gao, J., Tay, S. L., Koh, A. H. S., & Zhou, W. (2018). Dough and bread making from high- and low-protein flours by vacuum mixing: Part 3. Oral processing of bread. *Journal of Cereal Science*, 79, 408–417. <https://doi.org/10.1016/J.JCS.2017.12.002>
- Ginty, A. T. (2013). Construct validity. *Encyclopedia of Behavioral Medicine*, 487–487. https://doi.org/10.1007/978-1-4419-1005-9_861
- Glanz, K., Basil, M., Maibach, E., Goldberg, J., & Snyder, D. (1998). Why Americans eat what they do: taste, nutrition, cost, convenience, and weight control concerns as influences on food consumption. *Journal of the American Dietetic Association*, 98(10), 1118–1126. [https://doi.org/10.1016/S0002-8223\(98\)00260-0](https://doi.org/10.1016/S0002-8223(98)00260-0)
- Gonçalves, C., Rodrigues, J., Júnior, H., Carneiro, J., Freire, T., & Freire, L. (2017). Sodium reduction in margarine using NaCl substitutes. *Anais Da Academia Brasileira de Ciencias*, 89(3), 2505–2513. <https://doi.org/10.1590/0001-3765201720150618>
- Gonçalves, G. A. S., Resende, N. S., Gonçalves, C. S., de Carvalho, E. E. N., de Resende, J. V., & Boas, E. V. de B. V. (2020). How freezing methods affect the sensory profile of pasteurized and unpasteurized mangaba pulp. *Ciência e Agrotecnologia*, 44(003020), e003020. <https://doi.org/10.1590/1413-7054202044003020>
- Gonzalez-Estanol, K., Clicerí, D., Biasioli, F., & Stieger, M. (2022). Differences in dynamic sensory perception between reformulated hazelnut chocolate spreads decrease when spreads are consumed with breads and wafers. *Food Quality and Preference*, 98. <https://doi.org/10.1016/J.FOODQUAL.2022.104532>
- González-Mohino, A., Ventanas, S., Estévez, M., & Olegario, L. S. (2021). Sensory characterization of iberian dry-cured loins by using check-all-that-apply (CATA) analysis and multiple-intake temporal dominance of sensations (TDS). *Foods*, 10(9). <https://doi.org/10.3390/FOODS10091983>

- Goodman, L. A. (1965). On simultaneous confidence intervals for multinomial proportions. *Technometrics*, 7(2), 247–254. <https://doi.org/10.1080/00401706.1965.10490252>
- Gordin, H. H. (1987). Intensity variation descriptive methodology: development and application of a new sensory evaluation technique. *Journal of Sensory Studies*, 2(3), 187–198. <https://doi.org/10.1111/J.1745-459X.1987.TB00416.X>
- Gotow N., Kobayakawa T. (2021). Context Effect on Temporal Resolution of Olfactory–Gustatory, Visual–Gustatory, and Olfactory–Visual Synchrony Perception. *Chemosensory Perception*, 14 (1), 27 - 40, <https://doi.org/10.1007/s12078-020-09282-z>
- Gower, J. C., & Legendre, P. (1986). Metric and Euclidean properties of dissimilarity coefficients. *Journal of Classification*, 3(1), 5–48. <https://doi.org/10.1007/BF01896809/METRICS>
- Goza, J. L., Ziegler, G. R., Wee, J., Hayes, J. E., & Hopfer, H. (2022). Salivary α -amylase activity and flow rate explain differences in temporal flavor perception in a chewing gum matrix comprising starch-limonene inclusion complexes. *Food Research International*, 158. <https://doi.org/10.1016/J.FOODRES.2022.111573>
- Grainger, K. (2009). *Wine quality tasting and selection*. Wiley-Blackwell.
- Greenacre, M. J. (1988). Clustering the rows and columns of a contingency table. *Journal of Classification*, 5(1), 39–51. <https://doi.org/10.1007/BF01901670>
- Greis, M., Kukkonen, R., Lampi, A. M., Seppä, L., Partanen, R., & Sandell, M. (2022). The impact of vanilla and lemon aromas on sensory perception in plant-based yogurts measured with static and dynamic methods. *Foods*, 11(14). <https://doi.org/10.3390/FOODS11142030>
- Greis, M., Sainio, T., Katina, K., Kinchla, A. J., Nolden, A., Partanen, R., & Seppä, L. (2020). Dynamic texture perception in plant-based yogurt alternatives: Identifying temporal drivers of liking by TDS. *Food Quality and Preference*, 86. <https://doi.org/10.1016/J.FOODQUAL.2020.104019>
- Harrington, R. J. (2008). *Food and wine pairing : a sensory experience*. John Wiley.
- Harwood, W. S., & Drake, M. A. (2021). Application of temporal penalty analysis for the optimization of sugar reduction in protein beverages. *Journal of Sensory Studies*, 36(3). <https://doi.org/10.1111/JOSS.12644>
- Harwood, W. S., Parker, M. N., & Drake, M. A. (2020). Influence of ethanol concentration on sensory perception of rums using temporal check-all-that-apply. *Journal of Sensory Studies*, 35(1). <https://doi.org/10.1111/JOSS.12546>
- Hawthornthwaite, D., Ramjan, Y., & Rosenthal, A. (2015). Oral processing of low water content foods - a development to hutchings and lillford's breakdown path. *Journal of Texture Studies*, 46(3), 212–218. <https://doi.org/10.1111/JTXS.12126>
- Hays, J. C. (2010). Eight Recommendations for Writing Titles of Scientific Manuscripts. *Public Health Nursing*, 27(2), 101–103. <https://doi.org/10.1111/J.1525-1446.2010.00832.X>

- Hayward, L., Jantzi, H., Smith, A., & McSweeney, M. B. (2020). How do consumers describe cool climate wines using projective mapping and ultra-flash profile? *Food Quality and Preference*, 86, 104026. <https://doi.org/10.1016/J.FOODQUAL.2020.104026>
- He, Y., Chen, S., Tang, K., Qian, M., Yu, X., & Xu, Y. (2021). Sensory characterization of Baijiu pungency by combined time-intensity (TI) and temporal dominance of sensations (TDS). *Food Research International*, 147. <https://doi.org/10.1016/J.FOODRES.2021.110493>
- Heale, R., & Twycross, A. (2015). Validity and reliability in quantitative studies. *Evidence-Based Nursing*, 18(3), 66–67. <https://doi.org/10.1136/EB-2015-102129>
- Helm, E., & Troille, B. (1946). Selection of a taste panel. *Wallerstein Laboratory Communications*, 9(181–194).
- Higgins, M. J., Gipple, J. T., & Hayes, J. E. (2021). Common bitter stimuli show differences in their temporal profiles before and after swallowing. *Food Quality and Preference*, 87, 104041. <https://doi.org/10.1016/J.FOODQUAL.2020.104041>
- Higgs, S. (2016). Cognitive processing of food rewards. *Appetite*, 104, 10–17. <https://doi.org/10.1016/J.APPET.2015.10.003>
- Hoek, A. C., Pearson, D., James, S. W., Lawrence, M. A., & Friel, S. (2017). Healthy and environmentally sustainable food choices: Consumer responses to point-of-purchase actions. *Food Quality and Preference*, 58, 94–106. <https://doi.org/10.1016/J.FOODQUAL.2016.12.008>
- Honkanen, P., & Frewer, L. (2009). Russian consumers' motives for food choice. *Appetite*, 52(2), 363–371. <https://doi.org/10.1016/J.APPET.2008.11.009>
- Hough, G., Wakeling, I., Mucci, A., Chambers IV, E., Gallardo, I. M., & Alves, L. R. (2006). Number of consumers necessary for sensory acceptability tests. *Food Quality and Preference*, 17(6), 522–526. <https://doi.org/10.1016/J.FOODQUAL.2005.07.002>
- Hughson, A. L., & Boakes, R. A. (2002). The knowing nose: The role of knowledge in wine expertise. *Food Quality and Preference*, 13(7–8), 463–472. [https://doi.org/10.1016/S0950-3293\(02\)00051-4](https://doi.org/10.1016/S0950-3293(02)00051-4)
- Humes, L. E., Busey, T. A., Craig, J., & Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Attention, Perception, and Psychophysics*, 75(3), 508–524. <https://doi.org/10.3758/S13414-012-0406-9/TABLES/7>
- Hutchings, S. C., Cha, W., Dunshea, F. R., Sharma, C., & Torrico, D. D. (2022). Understanding dominance: The effect of changing the definition of dominance when using TDS with consumers. *Journal of Sensory Studies*, 37(4), e12750. <https://doi.org/10.1111/JOSS.12750>
- Hutchings, S. C., de Casanove, A., Schlich, P., & O'Riordan, D. (2017). The effect of training on the temporal dominance of sensations method: A study with milk protein hydrolysates. *Journal of Sensory Studies*, 32(6), e12303. <https://doi.org/10.1111/JOSS.12303>
- 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.

- Hutchings, S. C., Foster, K. D., Hedderley, D. I., & Morgenstern, M. P. (2014). Differences between age groups in the use of the temporal dominance of sensations technique across a range of food textures. *Journal of Texture Studies*, 45(3), 206–219. <https://doi.org/10.1111/JTXS.12066>
- Hutchings, S. C., Horner, K. M., Dible, V. A., Grigor, J. M. V., & O’Riordan, D. (2017). Modification of aftertaste with a menthol mouthwash reduces food wanting, liking, and ad libitum intake of potato crisps. *Appetite*, 108, 57–67. <https://doi.org/10.1016/J.APPET.2016.09.022>
- Ilic, J., Tomasevic, I., & Djekic, I. (2021a). Influence of boiling, steaming, and sous-vide on oral processing parameters of celeriac (*Apium graveolens* var. *rapaceum*). *International Journal of Gastronomy and Food Science*, 23. <https://doi.org/10.1016/J.IJGFS.2021.100308>
- Ilic, J., Tomasevic, I., & Djekic, I. (2021b). Influence of water-based and contact heating preparation methods on potato mechanical properties, mastication, and sensory perception. *International Journal of Gastronomy and Food Science*, 25. <https://doi.org/10.1016/J.IJGFS.2021.100401>
- Ilic, J., Tomasevic, I., & Djekic, I. (2022a). Purple eggplant and zucchini color, mechanical properties, mastication, and sensory perception influenced by boiling and grilling. *Journal of Texture Studies*, 53(2), 174–184. <https://doi.org/10.1111/JTXS.12651>
- Ilic, J., Tomasevic, I., & Djekic, I. (2022b). Influence of boiling, grilling, and sous-vide on mastication, bolus formation, and dynamic sensory perception of wild boar ham. *Meat Science*, 188. <https://doi.org/10.1016/J.MEATSCI.2022.108805>
- Ilic, J., Tomasevic, I., & Djekic, I. (2022c). Purple eggplant and zucchini color, mechanical properties, mastication, and sensory perception influenced by steaming and Sous-vide. *International Journal of Gastronomy and Food Science*, 28, 100549. <https://doi.org/10.1016/J.IJGFS.2022.100549>
- INRAE. (2022). TimeSens© software 2.0. <https://www.chemosenstools.com/timesens/>
- Ishida, Y., Shimizu, T., & Yoshikawa, M. (2020). An analysis and comparison of keyword recommendation methods for scientific data. *International Journal on Digital Libraries*, 21(3), 307–327. <https://doi.org/10.1007/S00799-020-00279-3/FIGURES/25>
- ISO. (2016). ISO 13299:2016 - Sensory analysis — Methodology — General guidance for establishing a sensory profile. <https://www.iso.org/standard/58042.html>
- ISO 5492:2008. (2008). ISO 5492:2008 (E/F/R) Sensory analysis – vocabulary. Sensory analysis – vocabulary.
- Jaccard, P. (1901). Distribution de la flore alpine dans le Bassin des Dranses et dans quelques regions voisines. *Bull Soc Vaudoise Sci Nat*, 37, 241–272. <https://cir.nii.ac.jp/crid/1570009751307275008>
- 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. <https://doi.org/10.1016/J.FOODQUAL.2016.03.006>
- Jaeger, S. 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. <https://doi.org/10.1016/J.FOODQUAL.2017.09.013>
- Jaeger, S. 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. <https://doi.org/10.1016/J.FOODQUAL.2016.09.001>
- Jaeger, S. R., Beresford, M. K., Lo, K. R., Hunter, D. C., Chheang, S. L., & Ares, G. (2020). What does it mean to check-all-that-apply? Four case studies with beverages. *Food Quality and Preference*, 80, 103794. <https://doi.org/10.1016/J.FOODQUAL.2019.103794>
- Jaeger, S. R., Chheang, S. L., Yin, J., Bava, C. M., Gimenez, A., Vidal, L., & Ares, G. (2013). Check-all-that-apply (CATA) responses elicited by consumers: Within-assessor reproducibility and stability of sensory product characterizations. *Food Quality and Preference*, 30(1), 56–67. <https://doi.org/10.1016/J.FOODQUAL.2013.04.009>
- Jaeger, S. 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, S. 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.
- Jager, G., Schlich, P., Tijssen, I., Yao, J., Visalli, M., de Graaf, C., & Stieger, M. (2014). Temporal dominance of emotions: Measuring dynamics of food-related emotions during consumption. *Food Quality and Preference*, 37.
- Januszewska, R., Giret, E., Clement, F., Van Leuven, I., Goncalves, C., Vladislavleva, E., Pradal, P., Nåbo, R., Landuyt, A., D'Heer, G., Frommenwiler, S., & Haefliger, H. (2020). Impact of vanilla origins on sensory characteristics of chocolate. *Food Research International*, 137. <https://doi.org/10.1016/J.FOODRES.2020.109313>
- Januszewska, R., Pieniak, Z., & Verbeke, W. (2011). Food choice questionnaire revisited in four countries. Does it still measure the same? *Appetite*, 57(1), 94–98. <https://doi.org/10.1016/J.APPET.2011.03.014>
- Jeltema, M., Beckley, J., Vahalik, J., & Garza, J. (2020). Consumer textural food perception over time based on Mouth Behavior. *Journal of Texture Studies*, 51(1), 185–194. <https://doi.org/10.1111/JTXS.12479>
- Johnson, F., & Wardle, J. (2014). Variety, palatability, and obesity. *Advances in Nutrition*, 5(6), 851. <https://doi.org/10.3945/AN.114.007120>

- Johnson, S. L., Stone, W. J., Bunn, J. A., Lyons, T. S., & Navalta, J. W. (2020). New author guidelines in statistical reporting: embracing an era beyond $p < .05$. *International Journal of Exercise Science*, 13(1), 1. /pmc/articles/PMC7523905/
- Josse, J., Pagès, J., & Husson, F. (2008). Testing the significance of the RV coefficient. *Computational Statistics & Data Analysis*, 53(1), 82–91. <https://doi.org/10.1016/J.CSDA.2008.06.012>
- Jourdren, S., Saint-Eve, A., Panouillé, M., Lejeune, P., Déléris, I., & Souchon, I. (2016). Respective impact of bread structure and oral processing on dynamic texture perceptions through statistical multiblock analysis. *Food Research International*, 87, 142–151. <https://doi.org/10.1016/J.FOODRES.2016.06.021>
- Jourdren, S., Saint-Eve, A., Pollet, B., Panouillé, M., Lejeune, P., Guichard, E., Déléris, I., & Souchon, I. (2017). Gaining deeper insight into aroma perception: An integrative study of the oral processing of breads with different structures. *Food Research International*, 92, 119–127. <https://doi.org/10.1016/J.FOODRES.2017.01.001>
- Jourdren, Solenne, Masson, M., Saint-Eve, A., Panouillé, M., Blumenthal, D., Lejeune, P., Déléris, I., & Souchon, I. (2017). Effect of bread crumb and crust structure on the in vivo release of volatiles and the dynamics of aroma perception. *Journal of Agricultural and Food Chemistry*, 65(16), 3330–3340. <https://doi.org/10.1021/ACS.JAFC.7B00287>
- Junge, J. Y., Mielby, L. A., Zeng, Y., Sun, Y. X., Byrne, D. V., Castura, J. C., & Kidmose, U. (2022). Investigating the temporality of binary taste interactions in blends of sweeteners and citric acid in solution. *Journal of Sensory Studies*, 37(6). <https://doi.org/10.1111/JOSS.12785>
- Kahneman, Daniel, Fredrickson, B. L., Schreiber, C. A., & Redelmeier, D. A. (1993). When more pain is preferred to less: adding a better end. *Psychological Science*, 4(6), 401–405. <https://doi.org/10.1111/J.1467-9280.1993.TB00589.X>
- Kang, W., Bindon, K. A., Wang, X., Muhlack, R. A., Smith, P. A., Niimi, J., & Bastian, S. E. P. (2020). Chemical and sensory impacts of Accentuated Cut Edges (ACE) grape must polyphenol extraction technique on shiraz wines. *Foods*, 9(8). <https://doi.org/10.3390/FOODS9081027>
- Kang, W., Niimi, J., Muhlack, R. A., Smith, P. A., & Bastian, S. E. P. (2019). Dynamic characterization of wine astringency profiles using modified progressive profiling. *Food Research International*, 120, 244–254. <https://doi.org/10.1016/J.FOODRES.2019.02.041>
- Kantono, K., Hamid, N., Chadha, D., Ma, Q., Oey, I., & Farouk, M. M. (2021). Pulsed electric field (Pef) processing of chilled and frozen-thawed lamb meat cuts: Relationships between sensory characteristics and chemical composition of meat. *Foods*, 10(5), 1148. <https://doi.org/10.3390/FOODS10051148/S1>
- Kantono, K., Hamid, N., Oey, I., Wang, S., Xu, Y., Ma, Q., Faridnia, F., & Farouk, M. (2019). Physicochemical and sensory properties of beef muscles after Pulsed Electric Field processing. *Food Research International*, 121, 1–11. <https://doi.org/10.1016/J.FOODRES.2019.03.020>
- Kantono, K., Hamid, N., Shepherd, D., Lin, Y. H. T., Brard, C., Grazioli, G., & Thomas Carr, B. (2018). The effect of music on gelato perception in different eating contexts. *Food Research International*, 113, 43–56. <https://doi.org/10.1016/J.FOODRES.2018.06.030>

- Kantono, K., Hamid, N., Shepherd, D., Lin, Y. H. T., Skiredj, S., & Carr, B. T. (2019). Emotional and electrophysiological measures correlate to flavour perception in the presence of music. *Physiology and Behavior*, 199, 154–164. <https://doi.org/10.1016/J.PHYSBEH.2018.11.012>
- Kantono, K., Hamid, N., Shepherd, D., Yoo, M. J. Y., Grazioli, G., & Carr, B. T. (2016). Listening to music can influence hedonic and sensory perceptions of gelati. *Appetite*, 100, 244–255. <https://doi.org/10.1016/J.APPET.2016.02.143>
- Kapferer, J. -N, & Laurent, G. (1993). Further evidence on the consumer involvement profile: Five antecedents of involvement. *Psychology & Marketing*, 10(4), 347–355. <https://doi.org/10.1002/MAR.4220100408>
- Kawasaki, H., Sekizaki, Y., Hirota, M., Sekine-Hayakawa, Y., & Nonaka, M. (2016). Analysis of binary taste-taste interactions of MSG, lactic acid, and NaCl by temporal dominance of sensations. *Food Quality and Preference*, 52, 1–10. <https://doi.org/10.1016/J.FOODQUAL.2016.03.010>
- 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. <https://doi.org/10.1016/J.FOODQUAL.2019.05.009>
- Keefer, H. M., Harwood, W. S., Castura, J. C., & Drake, M. A. (2022). Temporal ranking for characterization and improved discrimination of protein beverages. *Journal of Sensory Studies*, 37(4). <https://doi.org/10.1111/JOSS.12751>
- Keefer, H. R.M., Nishku, S., Gerard, P. D., & Drake, M. A. (2020). Role of sweeteners on temporality and bar hardening of protein bars. *Journal of Dairy Science*, 103(7), 6032–6053. <https://doi.org/10.3168/JDS.2019-17995>
- Keefer, Heather R.M., Rovai, D., & Drake, M. A. (2023). A timely application—Temporal methods, past, present, and future. *Journal of Food Science*, 88(S1), A21–A52. <https://doi.org/10.1111/1750-3841.16491>
- Kelling, S. T., & Halpern, B. P. (1988). Taste judgements and gustatory stimulus duration: Taste quality, taste intensity, and reaction time. *Chemical Senses*, 13(4), 559–586. <https://doi.org/10.1093/CHEMSE/13.4.559>
- Kemp, B., Trussler, S., Willwerth, J., & Inglis, D. (2019). Applying temporal check-all-that-apply (TCATA) to mouthfeel and texture properties of red wines. *Journal of Sensory Studies*, 34(4). <https://doi.org/10.1111/JOSS.12503>
- Kemp, S. E., Hort, J., & Hollowood, T. (2017). Time-dependent measures of perception in sensory evaluation. John Wiley & Sons Ltd.
- Khemacheevakul, K., Wolodko, J., Nguyen, H., & Wismer, W. (2021). Temporal sensory perceptions of sugar-reduced 3D printed chocolates. *Foods*, 10(9). <https://doi.org/10.3390/FOODS10092082>

- Kim, I. A., Hopkinson, A., van Hout, D., & Lee, H. S. (2017). A novel two-step rating-based 'double-faced applicability' test. Part 1: Its performance in sample discrimination in comparison to simple one-step applicability rating. *Food Quality and Preference*, 56, 189–200. <https://doi.org/10.1016/J.FOODQUAL.2016.10.010>
- Kiumarsi, M., Majchrzak, D., Jäger, H., Song, J., Lieleg, O., & Shahbazi, M. (2021). Comparative study of instrumental properties and sensory profiling of low-calorie chocolate containing hydrophobically modified inulin. Part II: Proton mobility, topological, tribological and dynamic sensory properties. *Food Hydrocolloids*, 110. <https://doi.org/10.1016/J.FOODHYD.2020.106144>
- Knaapila, A. (2022). Sensory and consumer research has a role in supporting sustainability of the food system. *Foods*, 11(13), 11. <https://doi.org/10.3390/FOODS11131958>
- Kobue-Lekalake, R. I., Taylor, J. R. N., & de Kock, H. L. (2012). Application of the dual attribute time-intensity (DATI) sensory method to the temporal measurement of bitterness and astringency in sorghum. *International Journal of Food Science and Technology*, 47(3), 459–466. <https://doi.org/10.1111/J.1365-2621.2011.02862.X>
- Köster, E. P. (2003). The psychology of food choice: some often encountered fallacies. *Food Quality and Preference*, 14(5–6), 359–373. [https://doi.org/10.1016/S0950-3293\(03\)00017-X](https://doi.org/10.1016/S0950-3293(03)00017-X)
- Kuesten, C., & Bi, J. (2018). Temporal drivers of liking based on functional data analysis and non-additive models for multi-attribute time-intensity data of fruit chews. *Foods*, 7(6). <https://doi.org/10.3390/FOODS7060084>
- 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. <https://doi.org/10.1016/J.FOODQUAL.2013.06.007>
- Kurata, S., Kuroda, R., & Komaki, F. (2022). Statistical modeling for temporal dominance of sensations data incorporating individual characteristics of panelists: an application to data of milk chocolate. *Journal of Food Science and Technology*, 59(6), 2420–2428. <https://doi.org/10.1007/S13197-021-05260-9>
- Kuusela, H., & Paul, P. (2000). A comparison of concurrent and retrospective verbal protocol analysis. *American Journal of Psychology*, 113(3), 387–404. <https://doi.org/10.2307/1423365>
- Labbe, D., Schlich, P., Pineau, N., Gilbert, F., & Martin, N. (2009). Temporal dominance of sensations and sensory profiling: A comparative study. *Food Quality and Preference*, 20(3), 216–221. <https://doi.org/10.1016/J.FOODQUAL.2008.10.001>
- Laguna, L., Varela, P., Salvador, A., & Fiszman, S. (2013). A new sensory tool to analyse the oral trajectory of biscuits with different fat and fibre contents. *Food Research International*, 51(2), 544–553. <https://doi.org/10.1016/J.FOODRES.2013.01.003>
- Lahne, J. (2016). Sensory science, the food industry, and the objectification of taste. 10. <https://doi.org/10.4000/AOF.7956>

- Lahne, J. (2018). Standard sensations: the production of objective experience from industrial technique. *The Senses and Society*, 13(1), 6–18. <https://doi.org/10.1080/17458927.2017.1420842>
- Laleg, K., Barron, C., Cordelle, S., Schlich, P., Walrand, S., & Micard, V. (2017). How the structure, nutritional and sensory attributes of pasta made from legume flour is affected by the proportion of legume protein. *LWT*, 79, 471–478. <https://doi.org/10.1016/J.LWT.2017.01.069>
- Lawless, H. T. (1984). Flavor Description of White Wine by “Expert” and Nonexpert Wine Consumers. *Journal of Food Science*, 49(1), 120–123. <https://doi.org/10.1111/J.1365-2621.1984.TB13686.X>
- Lawless, H. T., & Heymann, H. (2010). *Sensory Evaluation of Food* (2nd ed.). Springer New York. <https://doi.org/10.1007/978-1-4419-6488-5>
- Lawrence, G., Symoneaux, R., Maitre, I., Brossaud, F., Maestrojuan, M., & Mehinagic, E. (2013). Using the free comments method for sensory characterisation of Cabernet Franc wines: Comparison with classical profiling in a professional context. *Food Quality and Preference*, 30(2), 145–155. <https://doi.org/10.1016/J.FOODQUAL.2013.04.005>
- Le Calvé, B., Saint-Léger, C., Gaudreau, N., & Cayeux, I. (2019). Capturing key sensory moments during biscuit consumption: Using TDS to evaluate several concurrent sensory modalities. *Journal of Sensory Studies*, 34(6). <https://doi.org/10.1111/JOSS.12529>
- Le Révérend, F. M., Hidrio, C., Fernandes, A., & Aubry, V. (2008). Comparison between temporal dominance of sensations and time intensity results. *Food Quality and Preference*, 19(2), 174–178. <https://doi.org/10.1016/J.FOODQUAL.2007.06.012>
- Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software*, 25(1), 1–18. <https://doi.org/10.18637/JSS.V025.I01>
- 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*, 40, 78–82.
- Lehrer, A. (1975). Talking about Wine. *Language*, 51(4), 901. <https://doi.org/10.2307/412700>
- Lenfant, F., Loret, C., Pineau, N., Hartmann, C., & Martin, N. (2009). Perception of oral food breakdown. The concept of sensory trajectory. *Appetite*, 52(3), 659–667.
- Lepage, M., Neville, T., Rytz, A., Schlich, P., Martin, N., & Pineau, N. (2014). Panel performance for Temporal Dominance of Sensations. *Food Quality and Preference*, 38, 24–29.
- Lesme, H., Courcoux, P., Alleaume, C., Famelart, M. H., Bouhallab, S., Prost, C., & Rannou, C. (2020). Contribution of temporal dominance of sensations performed by modality (M-TDS) to the sensory perception of texture and flavor in semi-solid products: A case study on fat-free strawberry yogurts. *Food Quality and Preference*, 80, 103789. <https://doi.org/10.1016/J.FOODQUAL.2019.103789>

- Lesschaeve, I., & Noble, A. C. (2022). Sensory analysis of wine. *Managing Wine Quality*, 243–277. <https://doi.org/10.1016/B978-0-08-102067-8.00010-5>
- Lester, S., Hurst, K., Cornacchia, L., Kleijn, M., Ayed, C., Dinu, V., Taylor, M. A., & Fisk, I. (2021). The relation between stimulated salivary flow and the temporal consumption experience of a liquid oral nutritional supplement. *Appetite*, 166. <https://doi.org/10.1016/J.APPET.2021.105325>
- Liang, L., Duan, W., Zhang, J., Huang, Y., Zhang, Y., & Sun, B. (2022). Characterization and molecular docking study of taste peptides from chicken soup by sensory analysis combined with nano-LC-Q-TOF-MS/MS. *Food Chemistry*, 383. <https://doi.org/10.1016/J.FOODCHEM.2022.132455>
- Liang, P., Jiang, J., Ding, Q., Tang, X., & Roy, S. (2018). Memory load influences taste sensitivities. *Frontiers in Psychology*, 0(DEC), 2533. <https://doi.org/10.3389/FPSYG.2018.02533>
- Likert, R. A. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 40, 5–53.
- Lim, S. X. L., Höchenberger, R., Ruda, I., Fink, G. R., Viswanathan, S., & Ohla, K. (2022). The capacity and organization of gustatory working memory. *Scientific Reports* 2022 12:1, 12(1), 1–14. <https://doi.org/10.1038/s41598-022-12005-x>
- Lin, Y. H. T., Hamid, N., Shepherd, D., Kantono, K., & Spence, C. (2019). Environmental sounds influence the multisensory perception of chocolate gelati. *Foods*, 8(4). <https://doi.org/10.3390/FOODS8040124>
- Lin, Y. H. T., Hamid, N., Shepherd, D., Kantono, K., & Spence, C. (2022). Musical and non-musical sounds influence the flavour perception of chocolate ice cream and emotional responses. *Foods*, 11(12). <https://doi.org/10.3390/FOODS11121784>
- Liu, S. L., Jaw, Y. M., Wang, L. F., Chuang, G. C. C., Zhuang, Z. Y., Chen, Y. S., & Liou, B. K. (2021). Evaluation of sensory quality for taiwanese specialty teas with cold infusion using cata and temporal cata by taiwanese consumers. *Foods*, 10(10). <https://doi.org/10.3390/FOODS10102344>
- Lockshin, L., & Corsi, A. M. (2019). Consumer Research For Wine. *Handbook of Eating and Drinking*, 1–18. https://doi.org/10.1007/978-3-319-75388-1_162-1
- Lohéac, Y., & Issanchou, S. (2007). Using auctions to estimate prices and value of food products. *Consumer-Led Food Product Development*, 471–491. <https://doi.org/10.1533/9781845693381.3.471>
- Long, H. A., French, D. P., & Brooks, J. M. (2020). Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis: 1(1), 31–42. <https://doi.org/10.1177/2632084320947559>
- Lorido, L., Estévez, M., & Ventanas, S. (2018). Fast and dynamic descriptive techniques (Flash Profile, Time-intensity and Temporal Dominance of Sensations) for sensory characterization

of dry-cured loins. *Meat Science*, 145, 154–162. <https://doi.org/10.1016/J.MEATSCI.2018.06.028>

Lorido, L., Hort, J., Estévez, M., & Ventanas, S. (2016). Reporting the sensory properties of dry-cured ham using a new language: Time intensity (TI) and temporal dominance of sensations (TDS). *Meat Science*, 121, 166–174. <https://doi.org/10.1016/J.MEATSCI.2016.06.009>

Lorido, L., Pizarro, E., Estévez, M., & Ventanas, S. (2019). Emotional responses to the consumption of dry-cured hams by Spanish consumers: A temporal approach. *Meat Science*, 149, 126–133. <https://doi.org/10.1016/J.MEATSCI.2018.11.015>

Lunardo, R., Saintives, C., & Chaney, D. (2021). Food packaging and the color red: How negative cognitive associations influence feelings of guilt. *Journal of Business Research*, 134, 589–600. <https://doi.org/10.1016/J.JBUSRES.2021.05.052>

Lytra, G., Tempere, S., Marchand, S., De Revel, G., & Barbe, J. C. (2016). How do esters and dimethyl sulphide concentrations affect fruity aroma perception of red wine? Demonstration by dynamic sensory profile evaluation. *Food Chemistry*, 194, 196–200. <https://doi.org/10.1016/J.FOODCHEM.2015.07.143>

Lyu, J., Fu, J., Chen, S., Xu, Y., Nie, Y., & Tang, K. (2022). Impact of tannins on intraoral aroma release and retronasal perception, including detection thresholds and temporal perception by taste, in model wines. *Food Chemistry*, 375. <https://doi.org/10.1016/J.FOODCHEM.2021.131890>

Ma, Q., Hamid, N., Oey, I., Kantono., K., Faridnia, F., Yoo, M., & Farouk, M. (2016). Effect of chilled and freezing pre-treatments prior to pulsed electric field processing on volatile profile and sensory attributes of cooked lamb meats. *Innovative Food Science and Emerging Technologies*, 37, 359–374. <https://doi.org/10.1016/J.IFSET.2016.04.009>

Machado, J. C., Martins, Z. E., & Ferreira, I. M. P. L. V. O. (2023). Dynamic sensory analysis by Temporal Dominance of Sensations paired with dynamic liking and wanting methodologies to understand the consumers' preference between two beer styles enriched with elderberries. *LWT*, 173. <https://doi.org/10.1016/J.LWT.2022.114266>

Magalhães da Veiga Moreira, I., de Figueiredo Vilela, L., da Cruz Pedroso Miguel, M. G., Santos, C., Lima, N., & Freitas Schwan, R. (2017). Impact of a microbial cocktail used as a starter culture on cocoa fermentation and chocolate flavor. *Molecules (Basel, Switzerland)*, 22(5). <https://doi.org/10.3390/MOLECULES22050766>

Mahieu, B. (2021). MultiResponseR. <https://github.com/MahieuB/MultiResponseR>

Mahieu, B., Schlich, P., Visalli, M., & Cardot, H. (2021). A multiple-response chi-square framework for the analysis of Free-Comment and Check-All-That-Apply data. *Food Quality and Preference*, 93, 104256. <https://doi.org/10.1016/J.FOODQUAL.2021.104256>

Mahieu, B., Visalli, M., Thomas, A., & Schlich, P. (2020a). Free-comment outperformed check-all-that-apply in the sensory characterisation of wines with consumers at home. *Food Quality and Preference*, 84, 103937. <https://doi.org/10.1016/J.FOODQUAL.2020.103937>

- Mahieu, B., Visalli, M., Thomas, A., & Schlich, P. (2020b). Using Free-Comment with consumers to obtain temporal sensory descriptions of products. *Food Quality and Preference*, 86, 104008. <https://doi.org/10.1016/J.FOODQUAL.2020.104008>
- Mahieu, B., Visalli, M., Thomas, A., & Schlich, P. (2021). An investigation of the stability of Free-Comment and Check-All-That-Apply in two consumer studies on red wines and milk chocolates. *Food Quality and Preference*, 90, 104159. <https://doi.org/10.1016/J.FOODQUAL.2020.104159>
- Makame, J., Cronje, T., Emmambux, N. M., & De Kock, H. (2019). Dynamic oral texture properties of selected indigenous complementary porridges used in African communities. *Foods*, 8(6). <https://doi.org/10.3390/FOODS8060221>
- Mammasse, N., & Schlich, P. (2014). Adequate number of consumers in a liking test. Insights from resampling in seven studies. *Food Quality and Preference*, 31(1), 124–128. <https://doi.org/10.1016/J.FOODQUAL.2012.01.009>
- Marcano, J., Varela, P., Cunha, L. M., & Fiszman, S. (2015). Relating dynamic perception of reformulated cheese pies to consumers' expectations of satiating ability. *Food Research International*, 78, 369–377. <https://doi.org/10.1016/J.FOODRES.2015.09.005>
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160–1177. <https://doi.org/10.1016/J.JOI.2018.09.002>
- Martínez-García, N., Ramírez-Rivera, E. de J., Ramón-Canul, L. G., Servín-Juárez, R., López-Espíndola, M., & Herrera-Corredor, J. A. (2020). Sweetener solutions as binding agents for amaranth bars: evolution of temporal dominance of sensory attributes. *International Journal of Food Science and Technology*, 55(6), 2570–2579. <https://doi.org/10.1111/IJFS.14510>
- Martinez, S. J., Batista, N. N., Ramos, C. L., Dias, D. R., & Schwan, R. F. (2021). Brazilian cocoa hybrid-mix fermentation: Impact of microbial dominance as well as chemical and sensorial properties. *Journal of Food Science*, 86(6), 2604–2614. <https://doi.org/10.1111/1750-3841.15758>
- Martins, P. M. M., Batista, N. N., Miguel, M. G. da C. P., Simão, J. B. P., Soares, J. R., & Schwan, R. F. (2020). Coffee growing altitude influences the microbiota, chemical compounds and the quality of fermented coffees. *Food Research International*, 129. <https://doi.org/10.1016/J.FOODRES.2019.108872>
- Mayhew, E. J., Schmidt, S. J., Schlich, P., & Lee, S. Y. (2017). Temporal texture profile and identification of glass transition temperature as an instrumental predictor of stickiness in a caramel system. *Journal of Food Science*, 82(9), 2167–2176. <https://doi.org/10.1111/1750-3841.13822>
- Mayhew, E. J., Schmidt, S. J., Schlich, P., & Lee, S. Y. (2018). Correlation of consumer perception of stickiness and contributing texture attributes to trained panelist temporal evaluations in a caramel system. *Food Quality and Preference*, 65, 72–80. <https://doi.org/10.1016/J.FOODQUAL.2017.11.009>

- McCain-Keefer, H. R., Meals, S., & Drake, M. A. (2020). The sensory properties and consumer acceptance of cold brew coffee. *Journal of Sensory Studies*, 35(6). <https://doi.org/10.1111/JOSS.12604>
- Mccrickerd, K., & Forde, C. G. (2016). Sensory influences on food intake control: moving beyond palatability. *Obesity Reviews*, 17(1), 18–29. <https://doi.org/10.1111/OBR.12340>
- McGowan, J., Sampson, M., Salzwedel, D. M., Cogo, E., Foerster, V., & Lefebvre, C. (2016). PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement. *Journal of Clinical Epidemiology*, 75, 40–46. <https://doi.org/10.1016/J.JCLINEPI.2016.01.021>
- McMahon, K. M., Culver, C., Castura, J. C., & Ross, C. F. (2017). Perception of carbonation in sparkling wines using descriptive analysis (DA) and temporal check-all-that-apply (TCATA). *Food Quality and Preference*, 59, 14–26. <https://doi.org/10.1016/J.FOODQUAL.2017.01.017>
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12, 153-157.
- Medel-Marabolí, M., Romero, J. L., Obreque-Slier, E., Contreras, A., & Peña-Neira, A. (2017). Effect of a commercial tannin on the sensorial temporality of astringency. *Food Research International*, 102, 341–347. <https://doi.org/10.1016/J.FOODRES.2017.09.099>
- Medel-Marabolí, Marcela, López-Solís, R., Valenzuela-Prieto, D., Vargas-Silva, S., & Obreque-Slier, E. (2021). Limited relationship between temporality of sensory perception and phenolic composition of red wines. *LWT*, 142. <https://doi.org/10.1016/J.LWT.2021.111028>
- Medoro, C., Cianciabella, M., Camilli, F., Magli, M., Gatti, E., Predieri, S., Medoro, C., Cianciabella, M., Camilli, F., Magli, M., Gatti, E., & Predieri, S. (2016). Sensory profile of Italian craft beers, beer taster expert versus sensory methods: a comparative study. *Food and Nutrition Sciences*, 7(6), 454–465. <https://doi.org/10.4236/FNS.2016.76047>
- Meillon, S., Urbano, C., & Schlich, P. (2009). Contribution of the Temporal Dominance of Sensations (TDS) method to the sensory description of subtle differences in partially dealcoholized red wines. *Food Quality and Preference*, 20(7), 490–499. <https://doi.org/10.1016/J.FOODQUAL.2009.04.006>
- Meillon, S., Viala, D., Medel, M., Urbano, C., Guillot, G., & Schlich, P. (2010). Impact of partial alcohol reduction in Syrah wine on perceived complexity and temporality of sensations and link with preference. *Food Quality and Preference*, 21(7), 732–740. <https://doi.org/10.1016/J.FOODQUAL.2010.06.005>
- Meiselman, H.L., Jaeger, S. R., Carr, B. T., & Churchill, A. (2022). Approaching 100 years of sensory and consumer science: Developments and ongoing issues. *Food Quality and Preference*, 100, 104614. <https://doi.org/10.1016/J.FOODQUAL.2022.104614>
- Meiselman, Herbert L., & Schutz, H. G. (2003). History of food acceptance research in the US Army. *Appetite*, 40(3), 199–216. [https://doi.org/10.1016/S0195-6663\(03\)00007-2](https://doi.org/10.1016/S0195-6663(03)00007-2)
- Menegali, B. S., Selani, M. M., Saldaña, E., Patinho, I., Diniz, J. P., Melo, P. S., Pimentel Filho, N. de J., & Contreras-Castillo, C. J. (2020). Pink pepper extract as a natural antioxidant in

chicken burger: Effects on oxidative stability and dynamic sensory profile using Temporal Dominance of Sensations. *LWT*, 121. <https://doi.org/10.1016/J.LWT.2019.108986>

Menezes, A. G. T., Batista, N. N., Ramos, C. L., de Andrade e Silva, A. R., Efraim, P., Pinheiro, A. C. M., & Schwan, R. F. (2016). Investigation of chocolate produced from four different Brazilian varieties of cocoa (*Theobroma cacao* L.) inoculated with *Saccharomyces cerevisiae*. *Food Research International*, 81, 83–90. <https://doi.org/10.1016/J.FOODRES.2015.12.036>

Menis-Henrique, M. E. C., Janzantti, N. S., Andriot, I., Sémon, E., Berdeaux, O., Schlich, P., & Conti-Silva, A. C. (2019). Cheese-flavored expanded snacks with low lipid content: Oil effects on the in vitro release of butyric acid and on the duration of the dominant sensations of the products. *LWT*, 105, 30–36. <https://doi.org/10.1016/J.LWT.2019.01.052>

Mesurole, J., Saint-Eve, A., Déléris, I., & Souchon, I. (2013). Impact of fruit piece structure in yogurts on the dynamics of aroma release and sensory perception. *Molecules*, 18(5), 6035–6056. <https://doi.org/10.3390/molecules18056035>

Methven, L., Rahelu, K., Economou, N., Kinneavy, L., Ladbroke-Davis, L., Kennedy, O. B., Mottram, D. S., & Gosney, M. A. (2010). The effect of consumption volume on profile and liking of oral nutritional supplements of varied sweetness: Sequential profiling and boredom tests. *Food Quality and Preference*, 21(8), 948–955. <https://doi.org/10.1016/J.FOODQUAL.2010.04.009>

Meyners, M., & Castura, J. (2018). The analysis of temporal check-all-that-apply (TCATA) data. *Food Quality and Preference*, 67, 67–76. <https://doi.org/10.1016/J.FOODQUAL.2017.02.003>

Meyners, M. (2011). Panel and panelist agreement for product comparisons in studies of Temporal Dominance of Sensations. *Food Quality and Preference*, 22(4), 365–370. <https://doi.org/10.1016/J.FOODQUAL.2011.01.006>

Meyners, M. (2016). Temporal liking and CATA analysis of TDS data on flavored fresh cheese. *Food Quality and Preference*, 47, 101–108. <https://doi.org/10.1016/J.FOODQUAL.2015.02.005>

Meyners, M. (2020a). Controversy regarding relevance and rigor of Sensometrics for industrial applications. *Sensometrics 2020 – E3S*. <http://www.e3sensory.eu/sensometrics-2020/>

Meyners, M. (2020b). Temporal methods: Are we comparing apples and oranges? *Food Quality and Preference*, 79. <https://doi.org/10.1016/j.foodqual.2018.11.022>

Meyners, M., & Castura, J. C. (2018). The analysis of temporal check-all-that-apply (TCATA) data. *Food Quality and Preference*, 67, 67–76. <https://doi.org/10.1016/J.FOODQUAL.2017.02.003>

Meyners, M., & Castura, J. C. (2019). Did assessors select attributes by chance alone in your TDS study, and how relevant is it to know? *Food Research International*, 119, 571–583. <https://doi.org/10.1016/j.foodres.2018.10.035>

Meyners, M., & Pineau, N. (2010). Statistical inference for temporal dominance of sensations data using randomization tests. *Food Quality and Preference*, 21(7), 805–814. <https://doi.org/10.1016/J.FOODQUAL.2010.04.004>

- Middleton, F. (2022). The 4 Types of Validity in Research. Scribbr. <https://www.scribbr.com/methodology/types-of-validity/>
- Middleton, F. (2023). The 4 Types of Reliability in Research. Scribbr. <https://www.scribbr.com/methodology/types-of-reliability/>
- Miguel, M. G. da C. P., Reis, L. V. de C., Efraim, P., Santos, C., Lima, N., & Schwan, R. F. (2017). Cocoa fermentation: Microbial identification by MALDI-TOF MS, and sensory evaluation of produced chocolate. *LWT*, 77, 362–369. <https://doi.org/10.1016/J.LWT.2016.11.076>
- Milošević, J., Žeželj, I., Gorton, M., & Barjolle, D. (2012). Understanding the motives for food choice in Western Balkan Countries. *Appetite*, 58(1), 205–214. <https://doi.org/10.1016/J.APPET.2011.09.012>
- Missbach, B., Majchrzak, D., Sulzner, R., Wansink, B., Reichel, M., & Koenig, J. (2017). Exploring the flavor life cycle of beers with varying alcohol content. *Food Science and Nutrition*, 5(4), 889–895. <https://doi.org/10.1002/FSN3.472>
- Mitchell, J., Castura, J. C., Thibodeau, M., & Pickering, G. (2019). Application of TCATA to examine variation in beer perception due to thermal taste status. *Food Quality and Preference*, 73, 135–142. <https://doi.org/10.1016/J.FOODQUAL.2018.11.016>
- Morais, E. C., Pinheiro, A. C. M., Nunes, C. A., & Bolini, H. M. A. (2014). Multiple Time-Intensity analysis and Temporal Dominance of Sensations of chocolate dairy dessert using prebiotic and different high-intensity sweeteners. *Journal of Sensory Studies*, 29(5), 339–350. <https://doi.org/10.1111/JOSS.12115>
- Morata, A., Escott, C., Bañuelos, M. A., Loira, I., Del Fresno, J. M., González, C., & Suárez-lepe, J. A. (2020). Contribution of non-saccharomyces yeasts to wine freshness. A Review. *Biomolecules*, 10(1). <https://doi.org/10.3390/BIOM10010034>
- Morell, P., Ramírez-López, C., Vélez-Ruiz, J. F., & Fiszman, S. (2015). Relating HPMC concentration to elicited expected satiation in milk-based desserts. *Food Hydrocolloids*, 45, 158–167. <https://doi.org/10.1016/J.FOODHYD.2014.11.011>
- Moskowitz, H. R. (2003). The intertwining of psychophysics and sensory analysis: historical perspectives and future opportunities—a personal view. *Food Quality and Preference*, 14(2), 87–98. [https://doi.org/10.1016/S0950-3293\(02\)00072-1](https://doi.org/10.1016/S0950-3293(02)00072-1)
- Moskowitz, H. R. (2008). Validity and reliability in sensory science. Viewpoints and controversies in sensory science and consumer product testing, 97–101. <https://doi.org/10.1002/9780470385128.CH5>
- Moss, R., Barker, S., & McSweeney, M. B. (2021). The use of temporal check-all-that-apply and category scaling by experienced panellists to evaluate sweet and dry ciders. *Beverages*, 7(2). <https://doi.org/10.3390/BEVERAGES7020024>
- Mueller Loose, S., Francis, I. L. F., & Lockshin, L. (2008). The relationship between wine liking, subjective and objective wine knowledge: Does it matter who is in your “consumer” sample? 4th International Conference of the Academy of Wine Business Research, Siena 17-19th July 2008.

- Nair, A. (2019). Publication bias - Importance of studies with negative results! *Indian Journal of Anaesthesia*, 63(6), 505. https://doi.org/10.4103/IJA.IJA_142_19
- Narayanan, P., Chinnasamy, B., Jin, L., & Clark, S. (2014). Use of just-about-right scales and penalty analysis to determine appropriate concentrations of stevia sweeteners for vanilla yogurt. *Journal of Dairy Science*, 97(6), 3262–3272. <https://doi.org/10.3168/JDS.2013-7365>
- Nascimento, L. Á. do, Avila, B. P., Colussi, R., & Elias, M. C. (2020). Effect of abiotic stress on bioactive compound production in germinated brown rice. *Cereal Chemistry*, 97(4), 868–876. <https://doi.org/10.1002/CCHE.10310>
- Natsume, H., Okamoto, S., & Nagano, H. (2023). TDS similarity: outlier analysis using a similarity index to compare time-series responses of Temporal Dominance of Sensations tasks. *Foods* 2023, Vol. 12, Page 2025, 12(10), 2025. <https://doi.org/10.3390/FOODS12102025>
- Ng, M., Lawlor, J. B., Chandra, S., Chaya, C., Hewson, L., & Hort, J. (2012). Using quantitative descriptive analysis and temporal dominance of sensations analysis as complementary methods for profiling commercial blackcurrant squashes. *Food Quality and Preference*, 25(2), 121–134. <https://doi.org/10.1016/J.FOODQUAL.2012.02.004>
- Nguyen, H., & Wismer, W. V. (2022). Temporal sensory profiles of regular and sodium-reduced foods elicited by Temporal Dominance of Sensations (TDS) and Temporal Check-All-That-Apply (TCATA). *Foods*, 11(3). <https://doi.org/10.3390/FOODS11030457>
- Nguyen, Q. C., Castura, J. C., Nguyen, D. D. Le, & Varela, P. (2023). Identifying temporal sensory drivers of liking of biscuit supplemented with brewer's spent grain for young consumers. *Food Research International*, 170, 113049. <https://doi.org/10.1016/J.FOODRES.2023.113049>
- Nguyen, Q. C., Næs, T., & Varela, P. (2018). When the choice of the temporal method does make a difference: TCATA, TDS and TDS by modality for characterizing semi-solid foods. *Food Quality and Preference*, 66, 95–106.
- Nguyen, Q. C., Wahlgren, M. B., Almlí, V. L., & Varela, P. (2017). Understanding the role of dynamic texture perception in consumers' expectations of satiety and satiation. A case study on barley bread. *Food Quality and Preference*, 62, 218–226. <https://doi.org/10.1016/J.FOODQUAL.2017.06.006>
- Niimi, J., Eddy, A. I., Overington, A. R., Silcock, P., Bremer, P. J., & Delahunty, C. M. (2014). Cross-modal interaction between cheese taste and aroma. *International Dairy Journal*, 39(2), 222–228. <https://doi.org/10.1016/J.IDAIRYJ.2014.07.002>
- Niimi, J., Overington, A. R., Silcock, P., Bremer, P. J., & Delahunty, C. M. (2015). Influence of cross-modal sensory interactions on cheese flavour intensity and character. *ACS Symposium Series*, 1191, 15–25. <https://doi.org/10.1021/BK-2015-1191.CH002>
- Ningtyas, D. W., Bhandari, B., Bansal, N., & Prakash, S. (2019). Sequential aspects of cream cheese texture perception using temporal dominance of sensations (TDS) tool and its relation with flow and lubrication behaviour. *Food Research International*, 120, 586–594. <https://doi.org/10.1016/J.FOODRES.2018.11.009>

- Noble, A. C., Arnold, R. A., Buechsenstein, J., Leach, E. J., Schmidt, J. O., & Stern, P. M. (1987). Modification of a standardized system of wine aroma Terminology. *American Journal of Enology and Viticulture*, 38(2).
- Nogueira Soares Souza, F., Rocha Vieira, S., Leopoldina Lamounier Campidelli, M., Abadia Reis Rocha, R., Milani Avelar Rodrigues, L., Henrique Santos, P., de Deus Souza Carneiro, J., Maria de Carvalho Tavares, I., & Patrícia de Oliveira, C. (2022). Impact of using cocoa bean shell powder as a substitute for wheat flour on some of chocolate cake properties. *Food Chemistry*, 381. <https://doi.org/10.1016/J.FOODCHEM.2022.132215>
- Norris, J. M., Plonsky, L., Ross, S. J., & Schoonen, R. (2015). Guidelines for reporting quantitative methods and results in primary research. *Language Learning*, 65(2), 470–476. <https://doi.org/10.1111/LANG.12104>
- Nygren, T., Nilsen, A. N., & Öström, Å. (2017). Dynamic changes of taste experiences in wine and cheese combinations. *Journal of Wine Research*, 28(2), 105–122. <https://doi.org/10.1080/09571264.2017.1284655>
- O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, 89(9), 1245–1251. <https://doi.org/10.1097/ACM.0000000000000388>
- O'Brien, D. T. (2012). Thinking, fast and slow by Daniel Kahneman. *Journal of Social, Evolutionary, and Cultural Psychology*, 6(2), 253. <https://doi.org/10.1037/H0099210>
- O'Mahony, M. (1979). Psychophysical aspects of sensory analysis of dairy product: a critique. *Journal of Dairy Science*, 62(12), 1954–1962. [https://doi.org/10.3168/JDS.S0022-0302\(79\)83528-6](https://doi.org/10.3168/JDS.S0022-0302(79)83528-6)
- Oberrauter, L. M., Januszewska, R., Schlich, P., & Majchrzak, D. (2018). Sensory evaluation of dark origin and non-origin chocolates applying Temporal Dominance of Sensations (TDS). *Food Research International*, 111, 39–49. <https://doi.org/10.1016/J.FOODRES.2018.05.007>
- Okada, T., Okamoto, S., & Yamada, Y. (2022). Affective dynamics: causality modeling of temporally evolving perceptual and affective responses. *IEEE Transactions on Affective Computing*, 13(2), 628–639. <https://doi.org/10.1109/TAFFC.2019.2942931>
- Okamoto, S. (2021). Bootstrap resampling of Temporal Dominance of Sensations curves to compute uncertainties. *Foods* 2021, Vol. 10, Page 2472, 10(10), 2472. <https://doi.org/10.3390/FOODS10102472>
- Okamoto, S., Ehara, Y., Okada, T., & Yamada, Y. (2022). Affective dynamics: Principal Motion Analysis of Temporal Dominance of Sensations and Emotions data. *IEEE Transactions on Affective Computing*, 13(2), 871–880. <https://doi.org/10.1109/TAFFC.2020.2971700>
- Olatujoye, J. B., Methven, L., & Jauregi, P. (2020). Effect of β -lactoglobulin on perception of astringency in red wine as measured by sequential profiling. *LWT*, 130. <https://doi.org/10.1016/J.LWT.2020.109611>
- Olegario, L. S., González-Mohino, A., Estévez, M., Madruga, M. S., & Ventanas, S. (2020). Impact of 'free-from' and 'healthy choice' labeled versions of chocolate and coffee on

temporal profile (multiple-intake TDS) and liking. *Food Research International*, 137. <https://doi.org/10.1016/J.FOODRES.2020.109342>

Oliveira, A. A. A., Andrade, A. C., Bastos, S. C., Condino, J. P. F., Curzi Júnior, A., & Pinheiro, A. C. M. (2021). Use of strawberry and vanilla natural flavors for sugar reduction: A dynamic sensory study with yogurt. *Food Research International*, 139. <https://doi.org/10.1016/J.FOODRES.2020.109972>

Oliveira, D., Antúnez, L., Giménez, A., Castura, J. C., Deliza, R., & Ares, G. (2015). Sugar reduction in probiotic chocolate-flavored milk: Impact on dynamic sensory profile and liking. *Food Research International*, 75, 148–156. <https://doi.org/10.1016/J.FOODRES.2015.05.050>

Oliveira Mendes, M., Ferreira de Moraes, M., & Ferreira Rodrigues, J. (2019). A2A2 milk: Brazilian consumers' opinions and effect on sensory characteristics of Petit Suisse and Minas cheeses. *LWT*, 108, 207–213. <https://doi.org/10.1016/J.LWT.2019.03.064>

Oliver, P., Cicerale, S., Pang, E., & Keast, R. (2018). A Comparison of Temporal Dominance of Sensation (TDS) and Quantitative Descriptive Analysis (QDATM) to Identify Flavors in Strawberries. *Journal of Food Science*, 83(4), 1094–1102. <https://doi.org/10.1111/1750-3841.14096>

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

Ovejero-López, I., Bro, R., & Bredie, W. L. P. (2005). Univariate and multivariate modelling of flavour release in chewing gum using time-intensity: a comparison of data analytical methods. *Food Quality and Preference*, 16(4), 327–343. <https://doi.org/10.1016/J.FOODQUAL.2004.05.014>

Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review* 1994 6:4, 6(4), 351–371. <https://doi.org/10.1007/BF02213420>

Pacheco, M. H. S., Esmerino, E. A., Capobianco, C. S. C., Cruz, A. G., Leddomado, L. S., Pimentel, T. C., Benevides Filho, I. M., & de Freitas, M. Q. (2018). Bottled mineral water: classic and temporal descriptive sensory analysis associated with liking. *British Food Journal*, 120(7), 1547–1560. <https://doi.org/10.1108/BFJ-11-2017-0655>

Pae, C. U. (2015). Why Systematic Review rather than Narrative Review? *Psychiatry Investigation*, 12(3), 417. <https://doi.org/10.4306/PI.2015.12.3.417>

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *PLOS Medicine*, 18(3), e1003583. <https://doi.org/10.1371/JOURNAL.PMED.1003583>

Paglarini, C. de S., Vidal, V. A. S., dos Santos, M., Coimbra, L. O., Esmerino, E. A., Cruz, A. G., & Pollonio, M. A. R. (2020). Using dynamic sensory techniques to determine drivers of liking in

- sodium and fat-reduced Bologna sausage containing functional emulsion gels. *Food Research International*, 132. <https://doi.org/10.1016/J.FOODRES.2020.109066>
- Palczak, J., Blumenthal, D., & Delarue, J. (2019). From consumption behaviour to sensory measurement: Sensory characterization of the perceived flavour complexity of a chocolate dessert experience. *Food Quality and Preference*, 78, 103734. <https://doi.org/10.1016/J.FOODQUAL.2019.103734>
- Palczak, J., Blumenthal, D., & Delarue, J. (2020). Influence of sensory complexity on preferences for novel gourmet dairy desserts. Does Berlyne's theory apply to desserts? *Food Quality and Preference*, 84. <https://doi.org/10.1016/J.FOODQUAL.2020.103957>
- Panouillé, M., Saint-Eve, A., Déléris, I., Le Bleis, F., & Souchon, I. (2014). Oral processing and bolus properties drive the dynamics of salty and texture perceptions of bread. *Food Research International*, 62, 238–246. <https://doi.org/10.1016/J.FOODRES.2014.02.031>
- Panzarini, A. C., Menis-Henrique, M. E. C., & Conti-Silva, A. C. (2020). Effect of salt and monosodium glutamate on the sensory characteristics of low-sodium cheese-flavored corn grits expanded snacks. *Journal of Food Processing and Preservation*, 44(12). <https://doi.org/10.1111/JFPP.14936>
- Parker, M. N., Lopetcharat, K., & Drake, M. A. (2018). Consumer acceptance of natural sweeteners in protein beverages. *Journal of Dairy Science*, 101(10), 8875–8889. <https://doi.org/10.3168/JDS.2018-14707>
- Patel, M., Lee, R., Merchant, E. V., Juliani, H. R., Simon, J. E., & Tepper, B. J. (2021). Descriptive aroma profiles of fresh sweet basil cultivars (*Ocimum* spp.): Relationship to volatile chemical composition. *Journal of Food Science*, 86(7), 3228–3239. <https://doi.org/10.1111/1750-3841.15797>
- Patterson, N. M., Zhong, Y., James, B. J., Gant, N., & Hautus, M. J. (2021). Effect of basic structural variation, aimed at increasing perceivable textures in model foods, on the perception of textural complexity. *Food Quality and Preference*, 91. <https://doi.org/10.1016/J.FOODQUAL.2021.104196>
- Paulsen, M. T., Næs, T., Ueland, Ø., Rukke, E. O., & Hersleth, M. (2013). Preference mapping of salmon-sauce combinations: The influence of temporal properties. *Food Quality and Preference*, 27(2), 120–127. <https://doi.org/10.1016/J.FOODQUAL.2012.09.010>
- Paulsen, M. T., Nys, A., Kvarberg, R., & Hersleth, M. (2014). Effects of NaCl substitution on the sensory properties of sausages: Temporal aspects. *Meat Science*, 98(2), 164–170. <https://doi.org/10.1016/J.MEATSCI.2014.05.020>
- Paz, L. I., Januszewska, R., Schouteten, J. J., & Van Impe, J. (2021). Challenges of pairing chocolates and nuts: Perceptions, interactions and dynamics of contrasting chocolates with nuts. *Food Research International*, 148. <https://doi.org/10.1016/J.FOODRES.2021.110620>
- Pecore, S., Rathjen-Nowak, C., & Tamminen, T. (2011). Temporal Order of Sensations. 9th Pangborn Symposium.

- Pedrotti, M., Spaccasassi, A., Biasioli, F., & Fogliano, V. (2019). Ethnicity, gender and physiological parameters: Their effect on in vivo flavour release and perception during chewing gum consumption. *Food Research International*, 116, 57–70. <https://doi.org/10.1016/J.FOODRES.2018.12.019>
- Peltier, C., Visalli, M., & Thomas, A. (2019). Using temporal dominance of emotions at home. Impact of coffee advertisements on consumers' behavior and methodological perspectives. *Food Quality and Preference*, 71, 311–319.
- Peltier, C., Visalli, M., Labouré, H., Hélard, C., Andriot, I., Cordelle, S., Le Quéré, J. L., & Schlich, P. (2022). Automatic pretreatment and multiblock analysis of flavor release and sensory temporal data simultaneously collected in vivo. *Journal of Chemometrics*. <https://doi.org/10.1002/CEM.3450>
- Peltier, C., Visalli, M., Schlich, P., & Cardot, H. (2023). Analyzing Temporal Dominance of Sensations data with Categorical Functional Data Analysis. *Food Quality and Preference* (in Revision).
- Penders, B. (2018). Ten simple rules for responsible referencing. *PLOS Computational Biology*, 14(4), e1006036. <https://doi.org/10.1371/JOURNAL.PCBI.1006036>
- Pereira, C. T. M., Pereira, D. M., & Bolini, H. M. A. (2021). Dynamic sensory profile of mango skyr yoghurt added of prebiotic and natural sweeteners: multiple time-intensity analysis and temporal dominance of sensations. *International Journal of Food Science and Technology*, 56(8), 4159–4169. <https://doi.org/10.1111/IJFS.15045>
- Pereira, T. S., Batista, N. N., Santos Pimenta, L. P., Martinez, S. J., Ribeiro, L. S., Oliveira Naves, J. A., & Schwan, R. F. (2022). Self-induced anaerobiosis coffee fermentation: Impact on microbial communities, chemical composition and sensory quality of coffee. *Food Microbiology*, 103. <https://doi.org/10.1016/J.FM.2021.103962>
- Perrin, L., & Pagès, J. (2009). Construction of a product space from the ultra-flash profiling method: Application to 10 red wines from the loire valley. *Journal of Sensory Studies*, 24(3), 372–395. <https://doi.org/10.1111/J.1745-459X.2009.00216.X>
- Peryam, D. R., & Swartz, V. W. (1950). Measurement of sensory differences. *Food Technology*, 4, 390–395.
- Peute, L. W. P., de Keizer, N. F., & Jaspers, M. W. M. (2015). The value of Retrospective and Concurrent Think Aloud in formative usability testing of a physician data query tool. *Journal of Biomedical Informatics*, 55, 1–10. <https://doi.org/10.1016/J.JBI.2015.02.006>
- Peyron, M. A., Gierczynski, I., Hartmann, C., Loret, C., Dardevet, D., Martin, N., & Woda, A. (2011). Role of physical bolus properties as sensory inputs in the trigger of swallowing. *PLoS ONE*, 6(6). <https://doi.org/10.1371/JOURNAL.PONE.0021167>
- Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: advancing the approach and enhancing the consistency. *Research Synthesis Methods*, 5(4), 371–385. <https://doi.org/10.1002/JRSM.1123>

- Piedmont, R. L. (2014). Criterion Validity. *Encyclopedia of Quality of Life and Well-Being Research*, 1348–1348. https://doi.org/10.1007/978-94-007-0753-5_618
- Piepho, H. P. (2012). An Algorithm for a Letter-Based Representation of All-Pairwise Comparisons. *13(2)*, 456–466. <https://doi.org/10.1198/1061860043515>
- Pierguidi, L., Spinelli, S., Monteleone, E., & Dinnella, C. (2021). The combined use of temporal dominance of sensations (TDS) and discrete time-intensity (DTI) to describe the dynamic sensory profile of alcoholic cocktails. *Food Quality and Preference*, 93. <https://doi.org/10.1016/J.FOODQUAL.2021.104281>
- Pineau, N., Cordelle, S., & Schlich, P. (2003). Temporal Dominance of Sensations: A new technique to record several sensory attributes simultaneously over time. 5th Pangborn Symposium.
- Pineau, N., Schlich, P., Cordelle, S., Mathonnière, C., Issanchou, S., Imbert, A., Rogeaux, M., Etiévant, P., & Köster, E. (2009). Temporal Dominance of Sensations: Construction of the TDS curves and comparison with time–intensity. *Food Quality and Preference*, 20(6), 450–455.
- Pineau, N., de Bouillé, A. G., Lepage, M., Lenfant, F., Schlich, P., Martin, N., & Rytz, A. (2012). Temporal Dominance of Sensations: What is a good attribute list? *Food Quality and Preference*, 26(2), 159–165.
- Pinto da Rosa, P., Pio Ávila, B., Damé Veber Angelo, I., Garavaglia Chesini, R., Albandes Fernandes, T., da Silva Camacho, J., Bugoni, M., Roll, V. F. B., & Gularte, M. A. (2021). Impact of different chicken meat production systems on consumers' purchase perception. *British Poultry Science*, 62(3), 387–395. <https://doi.org/10.1080/00071668.2020.1857335>
- Pio Ávila, B., Cardozo, L. O., Alves, G. D., Pereira, A. M., Gularte, M. A., & de Oliveira, R. P. (2022). Targeted chemical and sensory profiling to guide consumption of blood orange. *Journal of Culinary Science and Technology*, 20(4), 350–365. <https://doi.org/10.1080/15428052.2020.1843581>
- Pionnier, E., Nicklaus, S., Chabanet, C., Mioche, L., Taylor, A. J., Le Quéré, J. L., & Salles, C. (2004). Flavor perception of a model cheese: Relationships with oral and physico-chemical parameters. *Food Quality and Preference*, 15(7-8 SPEC.ISS.), 843–852. <https://doi.org/10.1016/J.FOODQUAL.2004.04.011>
- Pittari, E., Piombino, P., Andriot, I., Cheynier, V., Cordelle, S., Feron, G., Gourrat, K., Le Quéré, J. L., Meudec, E., Moio, L., Neiers, F., Schlich, P., & Canon, F. (2022). Effects of oenological tannins on aroma release and perception of oxidized and non-oxidized red wine: A dynamic real-time in-vivo study coupling sensory evaluation and analytical chemistry. *Food Chemistry*, 372. <https://doi.org/10.1016/J.FOODCHEM.2021.131229>
- Plaehn, D. (2012). CATA penalty/reward. *Food Quality and Preference*, 24(1), 141–152. <https://doi.org/10.1016/J.FOODQUAL.2011.10.008>
- Popping the (PICO) question in research and evidence-based practice. (2002). *Applied Nursing Research*, 15(3), 197–198. <https://doi.org/10.1053/APNR.2002.34181>

- Portela, C. da S., Almeida, I. F. de, Reis, T. A. D. dos, Hickmann, B. R. B., & Benassi, M. de T. (2022). Effects of brewing conditions and coffee species on the physicochemical characteristics, preference and dynamics of sensory attributes perception in cold brews. *Food Research International*, 151. <https://doi.org/10.1016/J.FOODRES.2021.110860>
- Poveromo, A. R., & Hopfer, H. (2019). Temporal check-all-that-apply (TCATA) reveals matrix interaction effects on flavor perception in a model wine matrix. *Foods*, 8(12). <https://doi.org/10.3390/FOODS8120641>
- Pu, D., Duan, W., Huang, Y., Zhang, L., Zhang, Y., Sun, B., Ren, F., Zhang, H., & Tang, Y. (2021). Characterization of the dynamic texture perception and the impact factors on the bolus texture changes during oral processing. *Food Chemistry*, 339. <https://doi.org/10.1016/J.FOODCHEM.2020.128078>
- Pu, D., Shan, Y., Qiao, K., Zhang, L., Sun, B., & Zhang, Y. (2022). Development of an effective protocol for evaluating the saltiness intensity enhancement of umami compounds. *Journal of Agricultural and Food Chemistry*. <https://doi.org/10.1021/ACS.JAFC.2C06293>
- Pu, D., Zhang, H., Zhang, Y., Sun, B., Ren, F., Chen, H., & He, J. (2019). Characterization of the aroma release and perception of white bread during oral processing by gas chromatography-ion mobility spectrometry and temporal dominance of sensations analysis. *Food Research International*, 123, 612–622. <https://doi.org/10.1016/J.FOODRES.2019.05.016>
- Pu, D., Zhang, Y., Sun, B., Ren, F., Zhang, H., Chen, H., & Tang, Y. (2021). Characterization of the key taste compounds during bread oral processing by instrumental analysis and dynamic sensory evaluation. *LWT*, 138. <https://doi.org/10.1016/J.LWT.2020.110641>
- Puerta, P., Garzón, R., Rosell, C. M., Fiszman, S., Laguna, L., & Tárrega, A. (2021). Modifying gluten-free bread's structure using different baking conditions: Impact on oral processing and texture perception. *LWT*, 140, 110718. <https://doi.org/10.1016/J.LWT.2020.110718>
- Puerta, P., Laguna, L., Villegas, B., Rizo, A., Fiszman, S., & Tarrega, A. (2020). Oral processing and dynamics of texture perception in commercial gluten-free breads. *Food Research International*, 134. <https://doi.org/10.1016/J.FOODRES.2020.109233>
- Purkayastha, A., Palmaro, E., Falk-Krzesinski, H. J., & Baas, J. (2019). Comparison of two article-level, field-independent citation metrics: Field-Weighted Citation Impact (FWCI) and Relative Citation Ratio (RCR). *Journal of Informetrics*, 13(2), 635–642. <https://doi.org/10.1016/J.JOI.2019.03.012>
- R Core team. (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rahman, M. S., & McCarthy, O. J. (2009). A classification of food properties, 93–99. <https://doi.org/10.1080/10942919909524593>
- Ramírez-Rivera, E. de J., Rodríguez-Buenfil, I. M., Pérez-Robles, M., Oney-Montalvo, J. E., Prinyawiwatkul, W., Cabal-Prieto, A., Ramírez-Sucre, M. O., & Herrera-Corredor, J. A. (2021). Development and validation of a methodology for the sensometric characterisation of high-pungency peppers: a case study of habanero pepper (*Capsicum chinense* Jacq.). *International Journal of Food Science and Technology*, 56(2), 573–586. <https://doi.org/10.1111/IJFS.14735>

- Ramos-Díaz, J. M., Rinnan, Å., & Jouppila, K. (2019). Application of NIR imaging to the study of expanded snacks containing amaranth, quinoa and kañiwa. *LWT*, 102, 8–14. <https://doi.org/10.1016/J.LWT.2018.12.029>
- Ramos-Gabriel, S. U., Herrera-Corredor, J. A., Gamboa-Alvarado, J. G., & Ramírez-Rivera, E. de J. (2019). Impact of fermented whey addition on resulting sensory characteristics and consumer preference of ripened cheeses. *Emirates Journal of Food and Agriculture*, 31(6), 449–458. <https://doi.org/10.9755/EJFA.2019.V31.I6.1959>
- Ramsey, I., Dinu, V., Linforth, R., Yakubov, G. E., Harding, S. E., Yang, Q., Ford, R., & Fisk, I. (2020). Understanding the lost functionality of ethanol in non-alcoholic beer using sensory evaluation, aroma release and molecular hydrodynamics. *Scientific Reports*, 10(1). <https://doi.org/10.1038/S41598-020-77697-5>
- Ramsey, I., Ross, C., Ford, R., Fisk, I., Yang, Q., Gomez-Lopez, J., & Hort, J. (2018). Using a combined temporal approach to evaluate the influence of ethanol concentration on liking and sensory attributes of lager beer. *Food Quality and Preference*, 68, 292–303. <https://doi.org/10.1016/J.FOODQUAL.2018.03.019>
- Rébénaque, P., Rawlyer, A., Boldi, M. O., & Deneulin, P. (2015). Comparison between sensory and nephelometric evaluations of tannin fractions obtained by ultrafiltration of red wines. *Chemosensory Perception*, 8(1), 33–43. <https://doi.org/10.1007/S12078-015-9175-X>
- Reyes, M. M., Castura, J. C., & Hayes, J. E. (2017). Characterizing dynamic sensory properties of nutritive and nonnutritive sweeteners with temporal check-all-that-apply. *Journal of Sensory Studies*, 32(3). <https://doi.org/10.1111/JOSS.12270>
- Ribeiro, L. S., Evangelista, S. R., da Cruz Pedrozo Miguel, M. G., van Mullem, J., Silva, C. F., & Schwan, R. F. (2018). Microbiological and chemical-sensory characteristics of three coffee varieties processed by wet fermentation. *Annals of Microbiology*, 68(10), 705–716. <https://doi.org/10.1007/S13213-018-1377-4>
- Ribeiro, L. S., Ribeiro, D. E., Evangelista, S. R., Miguel, M. G. da C. P., Pinheiro, A. C. M., Borém, F. M., & Schwan, R. F. (2017). Controlled fermentation of semi-dry coffee (*Coffea arabica*) using starter cultures: A sensory perspective. *LWT - Food Science and Technology*, 82, 32–38. <https://doi.org/10.1016/J.LWT.2017.04.008>
- Ribeiro, M. N., Rodrigues, D. M., Rocha, R. A. R., Silveira, L. R., Condino, J. P. F., Júnior, A. C., de Souza, V. R., Nunes, C. A., & Pinheiro, A. C. M. (2020). Optimising a stevia mix by mixture design and napping: A case study with high protein plain yoghurt. *International Dairy Journal*, 110. <https://doi.org/10.1016/J.IDAIRYJ.2020.104802>
- Rios-Mera, J. D., Saldaña, E., Cruzado-Bravo, M. L. M., Martins, M. M., Patinho, I., Selani, M. M., Valentin, D., & Contreras-Castillo, C. J. (2020). Impact of the content and size of NaCl on dynamic sensory profile and instrumental texture of beef burgers. *Meat Science*, 161. <https://doi.org/10.1016/J.MEATSCI.2019.107992>
- Rizo, A., Jimenez-Pérez, I., Camacho-García, A., Fiszman, S., Pérez-Soriano, P., & Tarrega, A. (2019). Impact of texture TDS and flavour TDS tasks and of chocolate-chip biscuit

- characteristics on oral processing features. *Food Quality and Preference*, 76, 109–117. <https://doi.org/10.1016/J.FOODQUAL.2019.04.005>
- Rizo, A., Peña, E., Alarcon-Rojo, A. D., Fiszman, S., & Tarrega, A. (2019). Relating texture perception of cooked ham to the bolus evolution in the mouth. *Food Research International*, 118, 4–12. <https://doi.org/10.1016/J.FOODRES.2018.02.073>
- Rizo, A., Vidák, K., Fiszman, S., & Tarrega, A. (2020). Influence of fading duration on TCATA evaluation. *Food Quality and Preference*, 79. <https://doi.org/10.1016/J.FOODQUAL.2018.12.004>
- Robert, P., & Escoufier, Y. (1976). A unifying tool for linear multivariate statistical methods: The RV-coefficient. *Applied Statistics*, 25(3), 257. <https://doi.org/10.2307/2347233>
- Rocha, R. A. R., Ribeiro, M. N., Silva, G. A., Rocha, L. C. R., Pinheiro, A. C. M., Nunes, C. A., & Carneiro, J. de D. S. (2020). Temporal profile of flavor enhancers MAG, MSG, GMP, and IMP, and their ability to enhance salty taste, in different reductions of sodium chloride. *Journal of Food Science*, 85(5), 1565–1575. <https://doi.org/10.1111/1750-3841.15121>
- Rodrigues, D. M., de Souza, V. R., Mendes, J. F., Nunes, C. A., & Pinheiro, A. C. M. (2016). Microparticulated salts mix: An alternative to reducing sodium in shoestring potatoes. *LWT*, 69, 390–399. <https://doi.org/10.1016/J.LWT.2016.01.056>
- Rodrigues, D. M., Veríssimo, B. V. E., Pinheiro, A. C. M., & de Souza, V. R. (2018). Drivers of liking by TDS and acceptance of orange juice subject to different preservation processes. *Journal of Food Processing and Preservation*, 42(6). <https://doi.org/10.1111/JFPP.13639>
- Rodrigues, H., & Parr, W. V. (2019). Contribution of cross-cultural studies to understanding wine appreciation: A review. *Food Research International*, 115, 251–258. <https://doi.org/10.1016/J.FOODRES.2018.09.008>
- Rodrigues, J. F., Gonçalves, C. S., Pereira, R. C., Carneiro, J. D. S., & Pinheiro, A. C. M. (2014). Utilization of temporal dominance of sensations and time intensity methodology for development of low-sodium Mozzarella cheese using a mixture of salts. *Journal of Dairy Science*, 97(8), 4733–4744. <https://doi.org/10.3168/JDS.2014-7913>
- Rodrigues, J. F., Junqueira, G., Gonçalves, C. S., Carneiro, J. D. S., Pinheiro, A. C. M., & Nunes, C. A. (2014). Elaboration of garlic and salt spice with reduced sodium intake. *Anais Da Academia Brasileira de Ciências*, 86(4), 2065–2075. <https://doi.org/10.1590/0001-3765201420130464>
- Rodrigues, J. F., Andrade, R. da S., Bastos, S. C., Coelho, S. B., & Pinheiro, A. C. M. (2016). Miracle fruit: An alternative sugar substitute in sour beverages. *Appetite*, 107, 645–653. <https://doi.org/10.1016/J.APPET.2016.09.014>
- Rodrigues, J. F., Andrade, R. da S., Souza, V. R. de, Abreu, L. R. de, Barcelos, A. de F., Cruz, A. G. da, Esmerino, E. A., & Pinheiro, A. C. M. (2022). Drivers of linking of Prato cheeses: An evaluation using the check all that apply (CATA) and temporal dominance of sensations (TDS) tools. *Food Science and Technology International*, 28(5), 379–387. <https://doi.org/10.1177/10820132211018037>

- Rodrigues, J. F., Condino, J. P. F., Pinheiro, A. C. M., & Nunes, C. A. (2016). Temporal dominance of sensations of chocolate bars with different cocoa contents: Multivariate approaches to assess TDS profiles. *Food Quality and Preference*, 47, 91–96. <https://doi.org/10.1016/J.FOODQUAL.2015.06.020>
- Rodrigues, J. F., de Souza, V. R., Lima, R. R., Carneiro, J. de D. S., Nunes, C. A., & Pinheiro, A. C. M. (2016). Temporal dominance of sensations (TDS) panel behavior: A preliminary study with chocolate. *Food Quality and Preference*, 54, 51–57.
- Rodrigues, J. F., Resende, L. M. B., da Silva, L. F. de O., Pedroso, M. P., Pinheiro, A. C. M., & Nunes, C. A. (2019). Quality of olive oils from southeastern Brazil. *Bragantia*, 78(4), 479–489. <https://doi.org/10.1590/1678-4499.20180294>
- Rodrigues, J. F., Souza, V. R. de, Lima, R. R., Cruz, A. G. da, & Pinheiro, A. C. M. (2018). Tds of cheese: Implications of analyzing texture and taste simultaneously. *Food Research International*, 106, 1–10. <https://doi.org/10.1016/J.FOODRES.2017.12.048>
- Rogério Tavares Filho, E., Almeida Esmerino, E., de Almeida Santos-Junior, V., Cazzelato Lins da Silva, A., & Maria André Bolini, H. (2020). Dynamic aspects of salt reduction in tomato sauce by use of flavor enhancers and a bitter blocker. *Food Science and Technology International*, 26(6), 549–559. <https://doi.org/10.1177/1082013220913361>
- Rosenthal, A. J., & Share, C. (2014). Temporal Dominance of Sensations of peanuts and peanut products in relation to Hutchings and Lillford’s “breakdown path.” *Food Quality and Preference*, 32(PC), 311–316. <https://doi.org/10.1016/J.FOODQUAL.2013.09.004>
- Ross, P. T., & Bibler Zaidi, N. L. (2019). Limited by our limitations. *Perspectives on Medical Education*, 8(4), 261. <https://doi.org/10.1007/S40037-019-00530-X>
- Rothman, L., & Parker, M. J. (2009). Just-about-right (jar) scales: Design, usage, benefits, and risks. ASTM Committee E-18 on Sensory Evaluation of Materials and Products, & ASTM International. ASTM Manual MNL63, Conshohocken, PA: ASTM International.
- Runnebaum, R. C., Boulton, R. B., Powell, R. L., & Heymann, H. (2011). Key constituents affecting wine body – an exploratory study. *Journal of Sensory Studies*, 26(1), 62–70. <https://doi.org/10.1111/J.1745-459X.2010.00322.X>
- Rusticus, S. (2014). Content Validity. *Encyclopedia of quality of life and well-being research*, 1261–1262. https://doi.org/10.1007/978-94-007-0753-5_553
- Ryan, B., & Haslegrave, C. M. (2007). Use of concurrent and retrospective verbal protocols to investigate workers’ thoughts during a manual-handling task. *Applied Ergonomics*, 38(2), 177–190. <https://doi.org/10.1016/J.APERGO.2006.03.003>
- Sáenz-Navajas, M. P., Alegre, Y., de-la-Fuente, A., Ferreira, V., García, D., Eizaguirre, S., Razquin, I., & Hernández-Orte, P. (2016). Rapid sensory-directed methodology for the selection of high-quality aroma wines. *Journal of the Science of Food and Agriculture*, 96(12), 4250–4262. <https://doi.org/10.1002/JSFA.7636>

- Sáenz-Navajas, M. P., Ballester, J., Pêcher, C., Peyron, D., & Valentin, D. (2013). Sensory drivers of intrinsic quality of red wines: Effect of culture and level of expertise. *Food Research International*, 54(2), 1506–1518. <https://doi.org/10.1016/J.FOODRES.2013.09.048>
- Saint-Eve, A., Délérís, I., Panouillé, M., Dakowski, F., Cordelle, S., Schlich, P., & Souchon, I. (2011). How texture influences aroma and taste perception over time in candies. *Chemosensory Perception*, 4(1–2), 32–41. <https://doi.org/10.1007/S12078-011-9086-4>
- Saint-Eve, A., Panouillé, M., Capitaine, C., Délérís, I., & Souchon, I. (2015). Dynamic aspects of texture perception during cheese consumption and relationship with bolus properties. *Food Hydrocolloids*, 46, 144–152. <https://doi.org/10.1016/J.FOODHYD.2014.12.015>
- Saita, A., Yamamoto, K., Raevskiy, A., Takei, R., Washio, H., Shioiri, S., & Sakai, N. (2021). Crispness, the key for the palatability of “kakinotane”: A sensory study with onomatopoeic words. *Foods*, 10(8). <https://doi.org/10.3390/FOODS10081724>
- Saldaña, E., Martins, M. M., Menegali, B. S., Selani, M. M., & Contreras-Castillo, C. J. (2019). Effect of smoking and fat content on the dynamic perception of bacon by Brazilian consumers. *Scientia Agropecuaria*, 10(2), 169–173. <https://doi.org/10.17268/SCI.AGROPECU.2019.02.01>
- Saldaña, E., Soletti, I., Martins, M. M., Menegali, B. S., Merlo, T. C., Selani, M. M., Teixeira, A. C. B., da Silva, F. G., & Contreras-Castillo, C. J. (2019). Understanding consumers’ dynamic sensory perception for bacon smoked with different Brazilian woods. *Meat Science*, 154, 46–53. <https://doi.org/10.1016/J.MEATSCI.2019.04.006>
- Santagiuliana, M., Broers, L., Marigómez, I. S., Stieger, M., Piqueras-Fiszman, B., & Scholten, E. (2020). Strategies to compensate for undesired gritty sensations in foods. *Food Quality and Preference*, 81. <https://doi.org/10.1016/J.FOODQUAL.2019.103842>
- Santos Gonçalves, G. A., Resende, N. S., Gonçalves, C. S., de Alcântara, E. M., Nunes Carvalho, E. E., de Resende, J. V., Cirillo, M. Â., & de Barros Vilas Boas, E. V. (2017). Temporal dominance of sensations for characterization of strawberry pulp subjected to pasteurization and different freezing methods. *LWT*, 77, 413–421. <https://doi.org/10.1016/J.LWT.2016.11.077>
- Saporta, G., & Hatabian, G. (1986). Régions de confiance en analyse factorielle. E. D., et Al. (Eds.), *Data Analysis and Informatics*, Elsevier Science Publisher B.V. (1986). https://scholar.google.com/scholar_lookup?title=Régions
- Sato, K., & Kinugasa, H. (2019). Influence of Japanese green tea on the Koku attributes of bonito stock: Proposed basic rules of pairing Japanese green tea with Washoku. *Journal of Sensory Studies*, 34(6). <https://doi.org/10.1111/JOSS.12539>
- Sauvageot, F., Urdapilleta, I., & Peyron, D. (2006). Within and between variations of texts elicited from nine wine experts. *Food Quality and Preference*, 17(6), 429–444. <https://doi.org/10.1016/J.FOODQUAL.2005.05.007>
- Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (2011). *Psychology* (2nd Edition). Worth.
- Schiffman, S. S., Reynolds, M. L., Young, F. W., Kruskal, J. B., & Carroll, J. D. (1981). Introduction to multidimensional scaling : theory, methods, and applications. 413.

- Schlich, P., Medel Maraboli, M., Urbano, C., & Parr, W. V. (2015). Perceived complexity in Sauvignon blanc wines: Influence of domain-specific expertise. *Australian Journal of Grape and Wine Research*, 21(2), 168–178. <https://doi.org/10.1111/AJGW.12129>
- Schlich, P. (2017). Temporal Dominance of Sensations (TDS): a new deal for temporal sensory analysis. *Current Opinion in Food Science*, 15, 38–42. <https://doi.org/10.1016/J.COFS.2017.05.003>
- Schumaker, M. R., Diako, C., Castura, J. C., Edwards, C. G., & Ross, C. F. (2019). Influence of wine composition on consumer perception and acceptance of *Brettanomyces* metabolites using temporal check-all-that-apply methodology. *Food Research International*, 116, 963–972. <https://doi.org/10.1016/J.FOODRES.2018.09.034>
- Schüttler, A., Friedel, M., Jung, R., Rauhut, D., & Darriet, P. (2015). Characterizing aromatic typicality of Riesling wines: merging volatile compositional and sensory aspects. *Food Research International*, 69, 26–37. <https://doi.org/10.1016/J.FOODRES.2014.12.010>
- Schwarz, N. (2012). Retrospective and concurrent self-reports: the rationale for real-time data capture. In Mehl & Connor (Eds.), *Handbook of Research Methods for Studying Daily Life*. Guilford.
- Sethupathy, P., Priyadarshini, S. R., Moses, J. A., & Anandharamakrishnan, C. (2021). Matrix-dependent oral processing, oro-sensory perception, and glycemic index of chocolate bars. *Journal of Food Processing and Preservation*, 45(12). <https://doi.org/10.1111/JFPP.16067>
- Shahbazi, M., Jäger, H., Chen, J., & Ettelaie, R. (2021). Construction of 3D printed reduced-fat meat analogue by emulsion gels. Part II: Printing performance, thermal, tribological, and dynamic sensory characterization of printed objects. *Food Hydrocolloids*, 121. <https://doi.org/10.1016/J.FOODHYD.2021.107054>
- Shahbazi, M., Jäger, H., & Ettelaie, R. (2021). Application of Pickering emulsions in 3D printing of personalized nutrition. Part II: Functional properties of reduced-fat 3D printed cheese analogues. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 624. <https://doi.org/10.1016/J.COLSURFA.2021.126760>
- Shahbazi, M., Jäger, H., & Ettelaie, R. (2022). A promising therapeutic soy-based pickering emulsion gel stabilized by a multifunctional microcrystalline cellulose: application in 3D food printing. *Journal of Agricultural and Food Chemistry*, 70(7), 2374–2388. <https://doi.org/10.1021/ACS.JAFC.1C05644>
- Shanthamma, S., Priyanka, S., Priyanga, S., Moses, J. A., & Anandharamakrishnan, C. (2021). Production of low glycemic index chocolates with natural sugar substitutes. *Journal of Culinary Science and Technology*. <https://doi.org/10.1080/15428052.2021.1978364>
- Shapin, S. (2016). A taste of science: Making the subjective objective in the California wine world. *Social Studies of Science*, 46(3), 436–460. <https://doi.org/10.1177/0306312716651346>
- Shapiro, L., & Hubert, A. (2005). Something from the oven: reinventing dinner in 1950s America. [Http://Journals.Openedition.Org/Aof](http://Journals.Openedition.Org/Aof), 4. <https://doi.org/10.4000/AOF.233>

- Sharma, M., & Duizer, L. (2019). Characterizing the dynamic textural properties of hydrocolloids in pureed foods - A comparison between TDS and TCATA. *Foods*, 8(6). <https://doi.org/10.3390/FOODS8060184>
- Shimaoka, N., Okamoto, S., Akiyama, Y., & Yamada, Y. (2022). Linking Temporal Dominance of Sensations for primary-sensory and multi-sensory attributes using Canonical Correlation Analysis. *Foods* 2022, Vol. 11, Page 781, 11(6), 781. <https://doi.org/10.3390/FOODS11060781>
- Sidel, J. L., & Stone, H. (1993). The role of sensory evaluation in the food industry. *Food Quality and Preference*, 4(1–2), 65–73. [https://doi.org/10.1016/0950-3293\(93\)90314-V](https://doi.org/10.1016/0950-3293(93)90314-V)
- Silva, A. P., Voss, H. P., van Zyl, H., Hogg, T., de Graaf, C., Pintado, M., & Jager, G. (2018). Temporal dominance of sensations, emotions, and temporal liking measured in a bar for two similar wines using a multi-sip approach. *Journal of Sensory Studies*, 33(5), e12459. <https://doi.org/10.1111/JOSS.12459>
- Silva, A. P., Voss, H. P., van Zyl, H., Hogg, T., de Graaf, C., Pintado, M., & Jager, G. (2019). Effect of adding hop aroma in beer analysed by temporal dominance of sensations and emotions coupled with temporal liking. *Food Quality and Preference*, 75, 54–63. <https://doi.org/10.1016/J.FOODQUAL.2019.02.001>
- Silva, H. L. A., Balthazar, C. F., Silva, R., Vieira, A. H., Costa, R. G. B., Esmerino, E. A., Freitas, M. Q., & Cruz, A. G. (2018). Sodium reduction and flavor enhancer addition in probiotic prato cheese: Contributions of quantitative descriptive analysis and temporal dominance of sensations for sensory profiling. *Journal of Dairy Science*, 101(10), 8837–8846. <https://doi.org/10.3168/JDS.2018-14819>
- Silva, R., Rocha, R. S., Guimarães, J. T., Balthazar, C. F., Hugo Scudino, Ramos, G. L. P. A., Pimentel, T. C., Silva, M. C., Henrique F. Silva, P., Duarte, M. C. K. H., Freitas, M. Q., Cruz, A. G., & Esmerino, E. A. (2020). Dulce de leche submitted to ohmic heating treatment: Consumer sensory profile using preferred attribute elicitation (PAE) and temporal check-all-that-apply (TCATA). *Food Research International*, 134. <https://doi.org/10.1016/J.FOODRES.2020.109217>
- Simioni, S. C. C., Tovar, D. M., Rodrigues, J. F., de Souza, V. R., Nunes, C. A., Vietoris, V., & Pinheiro, A. C. M. (2018). Temporal dominance of sensations and preferences of Brazilians and Slovaks: A cross-cultural study of cachaças stored with woods from the Amazon rainforest. *Journal of the Science of Food and Agriculture*, 98(11), 4058–4064. <https://doi.org/10.1002/JSFA.8922>
- Simunovic, S., Đorđević, V., Lakićević, B., Djekic, I., Lorenzo, J. M., Barba, F. J., & Tomasevic, I. (2022). Digital evaluation of nitrite-reduced “kulen” fermented sausage Quality. *Journal of Food Quality*, 2022. <https://doi.org/10.1155/2022/2480746>
- Sinding, C., Thibault, H., Hummel, T., & Thomas-Danguin, T. (2021). Odor-induced saltiness enhancement: insights into the brain chronometry of flavor perception. *Neuroscience*, 452, 126–137. <https://doi.org/10.1016/J.NEUROSCIENCE.2020.10.029>
- Sizo, A., Lino, A., Reis, L. P., & Rocha, Á. (2019). An overview of assessing the quality of peer review reports of scientific articles. *International Journal of Information Management*, 46, 286–293. <https://doi.org/10.1016/J.IJINFOMGT.2018.07.002>

- Sjöström, L. B. (1954). The descriptive analysis of flavor. *Food Acceptance Testing Methodology*.
- Sokolowsky, M., & Fischer, U. (2012). Evaluation of bitterness in white wine applying descriptive analysis, time-intensity analysis, and temporal dominance of sensations analysis. *Analytica Chimica Acta*, 732, 46–52. <https://doi.org/10.1016/J.ACA.2011.12.024>
- Sokolowsky, M., Rosenberger, A., & Fischer, U. (2013). Sensory impact of skin contact on white wines characterized by descriptive analysis, time-intensity analysis and temporal dominance of sensations analysis. *Food Quality and Preference*, 39, 285–297. <https://doi.org/10.1016/J.FOODQUAL.2014.07.002>
- Solomando, J. C., Antequera, T., Ventanas, S., & Perez-Palacios, T. (2021). Sensory profile and consumer perception of meat products enriched with EPA and DHA using fish oil microcapsules. *International Journal of Food Science and Technology*, 56(6), 2926–2937. <https://doi.org/10.1111/IJFS.14932>
- Sørensen, L. B., Møller, P., Flint, A., Martens, M., & Raben, A. (2003). Effect of sensory perception of foods on appetite and food intake: a review of studies on humans. *International Journal of Obesity and Related Metabolic Disorders : Journal of the International Association for the Study of Obesity*, 27(10), 1152–1166. <https://doi.org/10.1038/SJ.IJO.0802391>
- Sousa Lima, R., Cazelat de Medeiros, A., & André Bolini, H. M. (2020). Does the indoor thermal environment influence the dominant sensation in a functional beverage attribute? *Journal of Food Science*, 85(10), 3536–3542. <https://doi.org/10.1111/1750-3841.15427>
- Souza Olegario, L., González-Mohino, A., Estévez, M., Madruga, M. S., & Ventanas, S. (2022). Impact of absence of lactose on the dynamic sensory profile of yogurt: A multiple-intake TDS approach. *LWT*, 162. <https://doi.org/10.1016/J.LWT.2022.113430>
- Sparvoli, F., Giofré, S., Cominelli, E., Avite, E., Giuberti, G., Luongo, D., Gatti, E., Cianciabella, M., Daniele, G. M., Rossi, M., & Predieri, S. (2021). Sensory characteristics and nutritional quality of food products made with a biofortified and lectin free common bean (*Phaseolus vulgaris* L.) Flour. *Nutrients* 2021, Vol. 13, Page 4517, 13(12), 4517. <https://doi.org/10.3390/NU13124517>
- Spence, C., & Wang, Q. J. (2018). What does the term ‘complexity’ mean in the world of wine? *International Journal of Gastronomy and Food Science*, 14, 45–54.
- Sroka, C. J., & Nagaraja, H. N. (2018). Odds ratios from logistic, geometric, Poisson, and negative binomial regression models. *BMC Medical Research Methodology*, 18(1), 1–11. <https://doi.org/10.1186/S12874-018-0568-9/FIGURES/1>
- Steiner, J. E., Glaser, D., Hawilo, M. E., & Berridge, K. C. (2001). Comparative expression of hedonic impact: affective reactions to taste by human infants and other primates. *Neuroscience & Biobehavioral Reviews*, 25(1), 53–74. [https://doi.org/10.1016/S0149-7634\(00\)00051-8](https://doi.org/10.1016/S0149-7634(00)00051-8)
- Stelick, A., & Dando, R. (2018). Thinking outside the booth — the eating environment, context and ecological validity in sensory and consumer research. *Current Opinion in Food Science*, 21, 26–31. <https://doi.org/10.1016/J.COFS.2018.05.005>

- Stevens, S. S. (1969). Sensory scales of taste intensity. *Perception & Psychophysics*, 6(5), 302–308. <https://doi.org/10.3758/BF03210101/METRICS>
- Stolle, T., Grondinger, F., Dunkel, A., Meng, C., Médard, G., Kuster, B., & Hofmann, T. (2017). Salivary proteome patterns affecting human salt taste sensitivity. *Journal of Agricultural and Food Chemistry*, 65(42), 9275–9286. <https://doi.org/10.1021/ACS.JAFC.7B03862>
- Stone, A. A., Shiffman, S. S., & DeVries, M. W. (1999). Ecological momentary assessment. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology*.
- Stone, A., Shiffman, S., Atienza, A., & Nebeling, L. (2007). *The science of real-time data capture: self-reports in health research* (Google eBook). 416.
- Stone, H. (2018). Example food: What are its sensory properties and why is that important? *Npj Science of Food* 2018 2:1, 2(1), 1–3. <https://doi.org/10.1038/s41538-018-0019-3>
- Stone, H., Bleibaum, R. N., & Thomas, H. A. (2012a). Sensory evaluation practices. In *Sensory Evaluation Practices* (4th ed.). Elsevier Inc. <https://doi.org/10.1016/C2009-0-63404-8>
- Stone, H., Bleibaum, R. N., & Thomas, H. A. (2012b). Test strategy and the design of experiments. *Sensory Evaluation Practices*, 117–165. <https://doi.org/10.1016/B978-0-12-382086-0.00004-2>
- Stone, H., Sidel, J., Oliver, S., Woolsey, A., & Singleton, R. C. (1974). Sensory evaluation by Quantitative Descriptive Analysis. In *Descriptive Sensory Analysis in Practice* (pp. 23–34). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470385036.CH1C>
- Takagaki, M., & Asakura, Y. (1984). Taste time-intensity measurements using a new computerized system. *Proceedings of the 18th Japanese Symposium on Taste and Smell*, Ed. K. Ueda. JASTS., 105–108.
- Tan, V. W. K., Wee, M. S. M., Tomic, O., & Forde, C. G. (2019). Temporal sweetness and side tastes profiles of 16 sweeteners using temporal check-all-that-apply (TCATA). *Food Research International*, 121, 39–47. <https://doi.org/10.1016/J.FOODRES.2019.03.019>
- Tanaya, H., & Nagai, Y. (2022). Investigating the relationship between geometric features of voids and textures in aerated foods. *Journal of Advanced Mechanical Design, Systems and Manufacturing*, 16(6). <https://doi.org/10.1299/JAMDSM.2022JAMDSM0062>
- Tang, C. S., Tan, V. W. K., Teo, P. S., & Forde, C. G. (2020). Savoury and kokumi enhancement increases perceived calories and expectations of fullness in equicaloric beef broths. *Food Quality and Preference*, 83. <https://doi.org/10.1016/J.FOODQUAL.2020.103897>
- Tang, J., Larsen, D. S., Ferguson, L., & James, B. J. (2017). Textural complexity model foods assessed with instrumental and sensory measurements. *Journal of Texture Studies*, 48(1), 9–22. <https://doi.org/10.1111/JTXS.12188>
- Teillet, E., Schlich, P., Urbano, C., Cordelle, S., & Guichard, E. (2010). Sensory methodologies and the taste of water. *Food Quality and Preference*, 21(8), 967–976. <https://doi.org/10.1016/J.FOODQUAL.2010.04.012>

- ten Kleij, F., & Musters, P. A. D. (2003). Text analysis of open-ended survey responses: a complementary method to preference mapping. *Food Quality and Preference*, 14(1), 43–52. [https://doi.org/10.1016/S0950-3293\(02\)00011-3](https://doi.org/10.1016/S0950-3293(02)00011-3)
- Thach, L., & Chang, K. (2015). 2015 Survey of American wine consumer preferences. *Wine Business*. <https://www.winebusiness.com/news/article/160722/>
- Thach, Liz, & Olsen, J. (2015). Profiling the high frequency wine consumer by price segmentation in the US market. *Wine Economics and Policy*, 4(1), 53–59. <https://doi.org/10.1016/J.WEP.2015.04.001>
- Thomas, A., Chambault, M., Dreyfuss, L., Gilbert, C. C., Hegyi, A., Henneberg, S., Knippertz, A., Kostyra, E., Kremer, S., Silva, A. P., & Schlich, P. (2017). Measuring temporal liking simultaneously to Temporal Dominance of Sensations in several intakes. An application to Gouda cheeses in 6 Europeans countries. *Food Research International*, 99, 426–434. <https://doi.org/10.1016/J.FOODRES.2017.05.035>
- Thomas, A., van der Stelt, A. J., Prokop, J., Lawlor, J. B., & Schlich, P. (2016). Alternating temporal dominance of sensations and liking scales during the intake of a full portion of an oral nutritional supplement. *Food Quality and Preference*, 53, 159–167. <https://doi.org/10.1016/J.FOODQUAL.2016.06.008>
- Thomas, A., van der Stelt, A. J., Schlich, P., & Lawlor, J. Ben. (2018). Temporal drivers of liking for oral nutritional supplements for older adults throughout the day with monitoring of hunger and thirst status. *Food Quality and Preference*, 70, 40–48. <https://doi.org/10.1016/J.FOODQUAL.2017.05.001>
- Thomas, Arnaud, Visalli, M., Cordelle, S., & Schlich, P. (2015). Temporal drivers of liking. *Food Quality and Preference*, 40, 365–375.
- Thomas, D., & Hodges, I. (2010). Designing and managing your research project: core skills for social and health research. <https://doi.org/10.4135/9781446289044>
- Toko, K. (2022). Visualization of changes in taste of food during chewing process. *Sensors and Materials*, 34(10), 3739–3750. <https://doi.org/10.18494/SAM4012>
- Tomadoni, B., Fiszman, S., Moreira, M. R., & Tarrega, A. (2018). The role of the dynamic sensory perception in the reformulation of shakes: use of TDS for studying the effect of milk, fiber, and flavor addition. *Journal of Food Science*, 83(1), 198–204. <https://doi.org/10.1111/1750-3841.14008>
- Torres-Moreno, M., Hort, J., Tarrega, A. (2016). The challenges of evaluating temporal sensations in beverages: Performance of TDS, TCATA and AEF in multiple-sip experiments. Presented at 7th European Conference on Sensory and Consumer Research (Eurosense), Dijon, France, September 2016.
- Torrìco, D. D. (2021). Novel techniques to measure the sensory, emotional, and physiological responses of consumers toward foods. *Foods* 2021, Vol. 10, Page 2620, 10(11), 2620. <https://doi.org/10.3390/FOODS10112620>

- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- Urbano, C., Mahieu, B., Thomas, A., Schlich, P., Visalli, M. (2018). Recording facial mimics and temporal dominance of sensations and emotions. 8th European Conference on Sensory and Consumer Research, Verona, Italy, September 2018.
- Urdapilleta, I., Parr, W., Dacremont, C., & Green, J. (2011). Semantic and perceptive organisation of Sauvignon blanc wine characteristics: Influence of expertise. *Food Quality and Preference*, 22(1), 119–128. <https://doi.org/10.1016/J.FOODQUAL.2010.08.005>
- Valentin, D., Chollet, S., Lelièvre, M., & Abdi, H. (2012). Quick and dirty but still pretty good: a review of new descriptive methods in food science. *International Journal of Food Science & Technology*, 47(8), 1563–1578. <https://doi.org/10.1111/J.1365-2621.2012.03022.X>
- Valentin, D., Valente, C., Ballester, J., Symoneaux, R., Smith, I., Bauer, F. F., & Nieuwoudt, H. (2021). The impact of “wine country of origin” on the perception of wines by South African and French wine consumers: a cross-cultural comparison. *Foods* 2021, Vol. 10, Page 1710, 10(8), 1710. <https://doi.org/10.3390/FOODS10081710>
- van Bommel, R., Stieger, M., Boelee, N., Schlich, P., & Jager, G. (2019). From first to last bite: Temporal dynamics of sensory and hedonic perceptions using a multiple-intake approach. *Food Quality and Preference*, 78. <https://doi.org/10.1016/J.FOODQUAL.2019.103748>
- van Bommel, R., Stieger, M., Schlich, P., & Jager, G. (2019). Dutch consumers do not hesitate: Capturing implicit ‘no dominance’ durations using Hold-down Temporal Dominance methodologies for Sensations (TDS) and Emotions (TDE).’ *Food Quality and Preference*, 71, 332–342.
- Van Den Haak, M. J., De Jong, M. D. T., & Schellens, P. J. (2003). Retrospective vs. concurrent think-aloud protocols: Testing the usability of an online library catalogue. *Behaviour and Information Technology*, 22(5), 339–351. <https://doi.org/10.1080/0044929031000>
- van Eck, A., Fogliano, V., Galindo-Cuspinera, V., Scholten, E., & Stieger, M. (2019). Adding condiments to foods: How does static and dynamic sensory perception change when bread and carrots are consumed with mayonnaise? *Food Quality and Preference*, 73, 154–170. <https://doi.org/10.1016/J.FOODQUAL.2018.11.013>
- van Eck, A., Hardeman, N., Karatza, N., Fogliano, V., Scholten, E., & Stieger, M. (2019). Oral processing behavior and dynamic sensory perception of composite foods: Toppings assist saliva in bolus formation. *Food Quality and Preference*, 71, 497–509. <https://doi.org/10.1016/J.FOODQUAL.2018.05.009>
- Vandeputte, A., Romans, J., Lenfant, F., Pineau, N. (2011). Innovative methods to assess the evolution of the sensory characteristics during the tasting of a full product portion (several bites). 9th Pangborn Sensory Science Symposium, Toronto, Canada, 4-8th September 2011.

- Varela, P., Antúnez, L., Carlehög, M., Alcaire, F., Castura, J. C., Berget, I., Giménez, A., Næs, T., & Ares, G. (2018). What is dominance? An exploration of the concept in TDS tests with trained assessors and consumers. *Food Quality and Preference*, 64, 72–81.
- Varela, P., & Ares, G. (2012). Sensory profiling, the blurred line between sensory and consumer science. A review of novel methods for product characterization. *Food Research International*, 48(2), 893–908. <https://doi.org/10.1016/J.FOODRES.2012.06.037>
- Varela, P., Mosca, A. C., Nguyen, Q. C., McEwan, J. A., & Berget, I. (2021). Individual differences underlying food intake and liking in semisolid foods. *Food Quality and Preference*, 87, 104023. <https://doi.org/10.1016/J.FOODQUAL.2020.104023>
- Varela, P., Pintor, A., & Fiszman, S. (2014). How hydrocolloids affect the temporal oral perception of ice cream. *Food Hydrocolloids*, 36, 220–228. <https://doi.org/10.1016/J.FOODHYD.2013.10.005>
- Vazallo-Valleumbrocio, G., Medel-Marabolí, M., Peña-Neira, Á., López-Solís, R., & Obrequeslier, E. (2017). Commercial enological tannins: Characterization and their relative impact on the phenolic and sensory composition of Carménère wine during bottle aging. *LWT*, 83, 172–183. <https://doi.org/10.1016/J.LWT.2017.05.022>
- Vázquez-Araújo, L., Parker, D., & Woods, E. (2013). Comparison of temporal-sensory methods for beer flavor evaluation. *Journal of Sensory Studies*, 28(5), 387–395. <https://doi.org/10.1111/JOSS.12064>
- Velázquez, A. L., Vidal, L., Varela, P., & Ares, G. (2020a). Cross-modal interactions as a strategy for sugar reduction in products targeted at children: Case study with vanilla milk desserts. *Food Research International*, 130. <https://doi.org/10.1016/J.FOODRES.2019.108920>
- Velázquez, A. L., Vidal, L., Varela, P., & Ares, G. (2020b). Can children use temporal sensory methods to describe visual and food stimuli? *Food Quality and Preference*, 86, 104002. <https://doi.org/10.1016/J.FOODQUAL.2020.104002>
- Veldhuizen, M. G., Shepard, T. G., Wang, M. F., & Marks, L. E. (2010). Coactivation of gustatory and olfactory signals in flavor perception. *Chemical Senses*, 35(2), 121. <https://doi.org/10.1093/CHEMSE/BJP089>
- Veldhuizen, M. G., Siddique, A., Rosenthal, S., & Marks, L. E. (2018). Interactions of lemon, sucrose and citric acid in enhancing citrus, sweet and sour flavors. *Chemical Senses*, 43(1), 17–26. <https://doi.org/10.1093/CHEMSE/BJX063>
- Verhage, A., & Boels, D. (2017). Critical appraisal of mixed methods research studies in a systematic scoping review on plural policing: assessing the impact of excluding inadequately reported studies by means of a sensitivity analysis. *Quality and Quantity*, 51(4), 1449–1468. <https://doi.org/10.1007/S11135-016-0345-Y>
- Vidal, L., Antúnez, L., Ares, G., Cuffia, F., Lee, P. Y., Le Blond, M., & Jaeger, S. R. (2019). Sensory product characterisations based on check-all-that-apply questions: Further insights on how the static (CATA) and dynamic (TCATA) approaches perform. *Food Research International*, 125, 108510. <https://doi.org/10.1016/J.FOODRES.2019.108510>

- Vidal, L., Antúnez, L., Giménez, A., Medina, K., Boido, E., & Ares, G. (2016). Dynamic characterization of red wine astringency: Case study with Uruguayan Tannat wines. *Food Research International*, 82, 128–135. <https://doi.org/10.1016/J.FOODRES.2016.02.002>
- Vidal, L., Castura, J. C., Coste, B., Picallo, A., Jaeger, S. R., & Ares, G. (2017). Analysis of TCATA Fading data: Imputation of gaps in temporal profiles. *Food Quality and Preference*, 59, 114–122. <https://doi.org/10.1016/J.FOODQUAL.2017.02.008>
- Vidal, L., Giménez, A., Medina, K., Boido, E., & Ares, G. (2015). How do consumers describe wine astringency? *Food Research International*, 78, 321–326. <https://doi.org/10.1016/J.FOODRES.2015.09.025>
- Vietoris, V., Zajác, P., Čapla, J., Mendelová, A., Křižanová, K., & Benešová, L. (2015). Comparison of coffee species by sensory panel and electronic nose. *Journal of Microbiology, Biotechnology and Food Sciences*, 5(3), 234–237. <https://doi.org/10.15414/JMBFS.20115/16.5.3.234-237>
- Vilela, A. (2021). SENSORY EVALUATION OF FOOD AND ITS EVOLUTION OVERTIME. *História: Consensos e Dissensos Engendrados*, 45–58. <https://doi.org/10.22533/AT.ED.3162128065>
- Vimercati, W. C., Araújo, C. da S., Macedo, L. L., Fonseca, H. C., Guimarães, J. S., Abreu, L. R. de, & Pinto, S. M. (2020). Physicochemical, rheological, microbiological and sensory properties of newly developed coffee flavored kefir. *LWT*, 123. <https://doi.org/10.1016/J.LWT.2020.109069>
- Visalli, M., Lange, C., Mallet, L., Cordelle, S., & Schlich, P. (2016). Should I use touchscreen tablets rather than computers and mice in TDS trials? *Food Quality and Preference*, 52, 11–16. <https://doi.org/10.1016/j.foodqual.2016.03.007>
- Visalli, M., Mahieu, B., Thomas, A., & Schlich, P. (2020). Concurrent vs. retrospective temporal data collection: Attack-evolution-finish as a simplification of Temporal Dominance of Sensations? *Food Quality and Preference*, 85, 103956. <https://doi.org/10.1016/j.foodqual.2020.103956>
- Visalli, M., Béno, N., Nicolle, L., & Schlich, P. (2023). Assessment of the validity and reliability of temporal sensory evaluation methods used with consumers on controlled stimuli delivered by a gustometer. *Food Quality and Preference*, 110, 104942. <https://doi.org/10.1016/J.FOODQUAL.2023.104942>
- Visalli, M., Cordelle, S., Béno, N., & Schlich, P. (2023). Product discrimination ability of temporal sensory evaluation methods used with consumers. *Food Quality and Preference*, 115, 105123. <https://doi.org/10.1016/j.foodqual.2024.105123>
- Visalli, M., Cordelle, S., Mahieu, B., Pedron, C., Hoffarth, B., Praudel, M., Coutière, M., & Schlich, P. (2022). A dataset of sensory perception of chocolates, guacamoles, ice teas and crisps collected with consumers using six temporal methods. *Data in Brief*, 45, 108708. <https://doi.org/10.1016/j.dib.2022.108708>
- Visalli, M., Dubois, M., Schlich, P., Ric, F., Cardebat, J. M., & Georgantzis, N. (2023a). A dataset on the sensory and affective perception of Bordeaux and Rioja red wines collected from

French and Spanish consumers at home and international wine students in the lab. *Data in Brief*, 46. <https://doi.org/10.1016/j.dib.2022.108873>

Visalli, M., Dubois, M., Schlich, P., Ric, F., Cardebat, J. M., & Georgantzis, N. (2023b). Relevance of free-comment to describe wine temporal sensory perception: An application with panels varying in culture and expertise. *Food Quality and Preference*, 105, 104785. <https://doi.org/10.1016/j.foodqual.2022.104785>

Visalli, M., & Galmarini, M. V. (2022). Multi-attribute temporal descriptive methods in sensory analysis applied in food science: Protocol for a scoping review. *PLoS ONE*, 17(7 July). <https://doi.org/10.1371/journal.pone.0270969>

Visalli, M., & Galmarini, M. V. (2023a). A template to share sensory evaluation data. <https://doi.org/10.57745/B35XCS>

Visalli, M., & Galmarini, M. V. (2023b). Multi-attribute temporal descriptive methods in sensory analysis applied in food science: a systematic scoping review. *Comprehensive Reviews in Food Science and Food Safety*, 23, 1–45. <https://doi.org/10.1111/1541-4337.13294>

Visalli, M., Mahieu, B., Peltier, C., Cordelle, S., & Schlich, P. (2023). Temporal precision and resolution of TDS and TCATA used with consumers. *Food Quality and Preference* (Submitted).

Visalli, M., Mahieu, B., & Schlich, P. (2021). Time periods segmentation in TDS and TCATA. *AgroStat*. <https://hal.archives-ouvertes.fr/hal-03356891/>

Visalli, M., Mahieu, B., & Schlich, P. (2023). A dataset on concurrent and immediate retrospective measures of sensory perception and preferences of dark chocolates. *Data in Brief*, 49, 109314. <https://doi.org/10.1016/j.dib.2023.109314>

Visalli, M., Schlich, P., Mahieu, B., Thomas, A., Weber, M., & Guichard, E. (2023). First steps towards FAIRization of product-focused sensory data. *Food Quality and Preference*, 104, 104765. <https://doi.org/10.1016/j.foodqual.2022.104765>

Visalli, M., Wakihira, T., & Schlich, P. (2022a). A dataset on concurrent and immediate retrospective methods for measuring sensory perception and preferences of lemon-flavoured carbonated alcoholic drinks. *Data in Brief*, 43. <https://doi.org/10.1016/j.dib.2022.108346>

Visalli, M., Wakihira, T., & Schlich, P. (2022b). Concurrent vs. immediate retrospective temporal sensory data collection: A case study on lemon-flavoured carbonated alcoholic drinks. *Food Quality and Preference*, 101, 104629. <https://doi.org/10.1016/j.foodqual.2022.104629>

Visalli, M., & Galmarini, M. V. (2023). Checklist for reporting research involving descriptive sensory evaluation methods - Sensory evaluation. <https://doi.org/10.57745/JUJRTJ>

Wagoner, T. B., McCain, H. R., Foegeding, E. A., & Drake, M. A. (2018). Food texture and sweetener type modify sweetness perception in whey protein-based model foods. *Journal of Sensory Studies*, 33(4), e12333. <https://doi.org/10.1111/JOSS.12333>

Wagoner, Ty B., Çakır-Fuller, E., Shingleton, R., Drake, M. A., & Foegeding, E. A. (2020). Viscosity drives texture perception of protein beverages more than hydrocolloid type. *Journal of Texture Studies*, 51(1), 78–91. <https://doi.org/10.1111/JTXS.12471>

- Wakihira, T., Miyashita, S., Kobayashi, M., Uemura, K., & Schlich, P. (2020). Temporal Dominance of Sensations paired with dynamic wanting in an ad libitum setting: A new method of sensory evaluation with consumers for a better understanding of beer drinkability. *Food Quality and Preference*, 86. <https://doi.org/10.1016/J.FOODQUAL.2020.103992>
- Wakihira, T., Visalli, M., & Schlich, P. (2023). Temporal drivers of liking by period: A case study on lemon-flavored carbonated alcoholic drinks with consumers in natural settings. *Food Quality and Preference*, 106, 104793. <https://doi.org/10.1016/J.FOODQUAL.2022.104793>
- Wal, R. C. van der, & Dillen, L. F. van. (2013). Leaving a flat taste in your mouth: task load reduces taste perception. *https://Doi.Org/10.1177/0956797612471953*, 24(7), 1277–1284. <https://doi.org/10.1177/0956797612471953>
- Wang, Q. J., Mesz, B., Riera, P., Trevisan, M., Sigman, M., Guha, A., & Spence, C. (2019). Analysing the impact of music on the perception of red wine via temporal dominance of sensations. *Multisensory Research*, 32(4–5), 455–472. <https://doi.org/10.1163/22134808-20191401>
- Wang, Q. J., Niaura, T., & Kantono, K. (2021). How does wine ageing influence perceived complexity? Temporal-Choose-All-That-Apply (TCATA) reveals temporal drivers of complexity in experts and novices. *Food Quality and Preference*, 92. <https://doi.org/10.1016/J.FOODQUAL.2021.104230>
- Ward, T., Bernier, R., Mukerji, C., Perszyk, D., McPartland, J. C., Johnson, E., Faja, S., Nevers, M., Frazier, T., Howlin, P., Savage, S., Zane, T., Lanner, T., Myers, M., VanBergeijk, E., Huestis, S., Bauminger-Zviely, N., Doehring, P., Voorst, G., Perszyk, D. (2013). Face validity. *Encyclopedia of Autism Spectrum Disorders*, 1226–1227. https://doi.org/10.1007/978-1-4419-1698-3_308
- Warton, D. I., Lyons, M., Stoklosa, J., & Ives, A. R. (2016). Three points to consider when choosing a LM or GLM test for count data. *Methods in Ecology and Evolution*, 7(8), 882–890. <https://doi.org/10.1111/2041-210X.12552>
- Watanabe, G., Motoyama, M., Orita, K., Takita, K., Aonuma, T., Nakajima, I., Tajima, A., Abe, A., & Sasaki, K. (2019). Assessment of the dynamics of sensory perception of Wagyu beef strip loin prepared with different cooking methods and fattening periods using the temporal dominance of sensations. *Food Science and Nutrition*, 7(11), 3538–3548. <https://doi.org/10.1002/FSN3.1205>
- Watanabe, G., Ohmori, H., Tajima, K., Sasaki, Y., Wakiya, Y., Motoyama, M., Nakajima, I., & Sasaki, K. (2019). Relative contribution of sensory characteristics for different types of pork loin, assessed by temporal dominance of sensations. *Journal of the Science of Food and Agriculture*, 99(12), 5516–5525. <https://doi.org/10.1002/JSFA.9813>
- Weerawarna N. R. P., M., Godfrey, A. J. R., Ellis, A., & Hort, J. (2021). Comparing temporal sensory product profile data obtained from expert and consumer panels and evaluating the value of a multiple sip TCATA approach. *Food Quality and Preference*, 89, 104141. <https://doi.org/10.1016/J.FOODQUAL.2020.104141>

- Weerawarna N. R. P., M., Godfrey, A. J. R., Ellis, A., & Hort, J. (2023). Identifying temporal drivers of product acceptance and rejection across sips during whole product consumption. *Journal of Sensory Studies*, e12855. <https://doi.org/10.1111/JOSS.12855>
- Weerawarna N.R.P, M., Godfrey, A. J. R., Ellis, A., & Hort, J. (2022). Effect of sipping method on sensory response to single and multiple sips of vanilla milkshake using temporal-check-all-that-apply. *Journal of Sensory Studies*, 37(5). <https://doi.org/10.1111/JOSS.12778>
- Whyte IV, J., Cormier, E., & Pickett-Hauber, R. (2010). Cognitions associated with nurse performance: A comparison of concurrent and retrospective verbal reports of nurse performance in a simulated task environment. *International Journal of Nursing Studies*, 47(4), 446–451. <https://doi.org/10.1016/J.IJNURSTU.2009.09.001>
- Wilkin, J. D., Ross, K., Alric, T., Hooper, M., Grigor, J. V., & Chu, B. S. (2021). Optimisation of concentration of *Undaria pinnatifida* (Wakame) and *Himathalia elongate* (Sea spaghetti) varieties to effect digestibility, texture and consumer attribute preference. *Journal of Aquatic Food Product Technology*, 30(8), 932–943. <https://doi.org/10.1080/10498850.2021.1958114>
- Wirtz, D., Kruger, J., Scollon, C. N., & Diener, E. (2003). What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science*, 14(5), 520–524. <https://doi.org/10.1111/1467-9280.03455>
- Withers, C. A., Lewis, M. J., Gosney, M. A., & Methven, L. (2014). Potential sources of mouth drying in beverages fortified with dairy proteins: A comparison of casein- and whey-rich ingredients. *Journal of Dairy Science*, 97(3), 1233–1247. <https://doi.org/10.3168/JDS.2013-7273>
- Withers, C., Barnagaud, C., Mehring, P., Ferris, S., & Thomson, D. M. H. (2016). Adapting and enhancing sequential profiling to understand the effects of successive ingestion, using the sensory characteristics of high intensity sweeteners as a case study. *Food Quality and Preference*, 47, 139–147. <https://doi.org/10.1016/J.FOODQUAL.2015.03.012>
- Wooding, S. P., Ramirez, V. A., & Behrens, M. (2021). Bitter taste receptors Genes, evolution and health. *Evolution, Medicine, and Public Health*, 9(1), 431–447. <https://doi.org/10.1093/EMPH/EOAB031>
- Wu, A. Z., Lee, R. W., Calvé, B. Le, & Cayeux, I. (2019). Temporal profiling of simplified lemonade using temporal dominance of sensations and temporal check-all-that-apply. *Journal of Sensory Studies*, 34(6), e12531. <https://doi.org/10.1111/JOSS.12531>
- Xiong, R., & Meullenet, J. F. (2006). A PLS dummy variable approach to assess the impact of jar attributes on liking. *Food Quality and Preference*, 17(3–4), 188–198. <https://doi.org/10.1016/J.FOODQUAL.2005.03.006>
- Xu, Y., Hamid, N., Shepherd, D., Kantono, K., Reay, S., Martinez, G., & Spence, C. (2019). Background soundscapes influence the perception of ice-cream as indexed by electrophysiological measures. *Food Research International*, 125. <https://doi.org/10.1016/J.FOODRES.2019.108564>
- Xu, Y., Hamid, N., Shepherd, D., Kantono, K., & Spence, C. (2019). Changes in flavour, emotion, and electrophysiological measurements when consuming chocolate ice cream in different

- eating environments. *Food Quality and Preference*, 77, 191–205. <https://doi.org/10.1016/J.FOODQUAL.2019.05.002>
- Yang, N., Yang, Q., Chen, J., & Fisk, I. (2021). Impact of capsaicin on aroma release and perception from flavoured solutions. *LWT*, 138. <https://doi.org/10.1016/J.LWT.2020.110613>
- Yates, F. (1934). Contingency tables involving small numbers and the χ^2 test. Supplement to the *Journal of the Royal Statistical Society*, 1(2), 217. <https://doi.org/10.2307/2983604>
- Young, A. K., Cheong, J. N., Foster, K. D., Hedderley, D. I., Morgenstern, M. P., & James, B. J. (2016). Exploring the links between texture perception and bolus properties throughout oral processing. Part 1: Breakdown paths. *Journal of Texture Studies*, 47(6), 461–473. <https://doi.org/10.1111/JTXS.12185>
- Young, A. K., Cheong, J. N., Hedderley, D. I., Morgenstern, M. P., & James, B. J. (2013). Understanding the link between bolus properties and perceived texture. *Journal of Texture Studies*, 44(5), 376–386. <https://doi.org/10.1111/JTXS.12025>
- Zhang, L. L., Xu, S. S., Shi, B. L., Wang, H. Y., Liu, L. Y., Zhong, K., Zhao, L., & Chen, Z. X. (2018). Evaluation of the pungency intensity and time-related aspects of Chinese *Zanthoxylum bungeanum* based on human sensation. *Journal of Sensory Studies*, 33(6), e12465. <https://doi.org/10.1111/JOSS.12465>
- Zhang, L., Shi, B., Wang, H., Zhao, L., & Chen, Z. (2017). Pungency evaluation of hydroxyl-sanshool compounds after dissolution in taste carriers per time-related characteristics. *Chemical Senses*, 42(7), 575–584. <https://doi.org/10.1093/CHEMSE/BJX038>
- Zhou, H., Zhao, Y., Fan, D., Shen, Q., Liu, C., & Luo, J. (2022). Effect of solid fat content in fat droplets on creamy mouthfeel of acid milk gels. *Foods*, 11(19). <https://doi.org/10.3390/FOODS11192932>
- Zhu, Y., Bhandari, B., & Prakash, S. (2020). Relating the tribo-rheological properties of chocolate flavoured milk to temporal aspects of texture. *International Dairy Journal*, 110. <https://doi.org/10.1016/J.IDAIRYJ.2020.104794>
- Zimoch, J., & Findlay, C. J. (1998). Effective discrimination of meat tenderness using dual attribute time intensity. *Journal of Food Science*, 63(6), 940–944. <https://doi.org/10.1111/J.1365-2621.1998.TB15828.X>
- Zorn, S., Alcaire, F., Vidal, L., Giménez, A., & Ares, G. (2014). Application of multiple-sip temporal dominance of sensations to the evaluation of sweeteners. *Food Quality and Preference*, 36, 135–143. <https://doi.org/10.1016/J.FOODQUAL.2014.04.003>
- Zwillinger, S. A., & Halpern, B. P. (1991). Time-quality tracking of monosodium glutamate, sodium saccharin, and a citric acid-saccharin mixture. *Physiology & Behavior*, 49(5), 855–862. [https://doi.org/10.1016/0031-9384\(91\)90194-S](https://doi.org/10.1016/0031-9384(91)90194-S)