

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

# Individual variation in Taste perception: Investigating sensory perception in Thermal Tasters and its impact on food choice behaviour

Thesis presented in partial fulfilment of the  
requirements for the degree of  
**Doctor of Philosophy in Food Technology**  
at Massey University, Palmerston North Campus,  
New Zealand.

Janita Jossie Botha (MSc)

May 2024

To my mother, Benita Botha.

I know you are proud of me, but if you could still read and fully understand this work, you would have been so immensely proud.

## Acknowledgements

First and foremost, thank you to my main supervisor, Prof Joanne Hort, for your support and guidance over this (perhaps needlessly long) journey. None of us could choose who we got to journey through the COVID-19 pandemic with, but I could not have made a better choice if I could. Thank you for being a safe place through many crises, and for being the kind of person who is willing to update their personal “algorithm” when presented with new information.

Thank you to my two relay co-supervisors; Dr Peter Cannon, for the first part of the journey and Dr Charles Diako for the second leg. Both of you kept me sane.

Ken Mercer and Malcolm Loudon deserve all the credit for building the thermode and developing the interface software for RapCoTT. Thank you for a fun and fruitful collaboration.

Thank you to the Feast and Riddet Institute staff for their support in the execution of this project, and to the Riddet Institute for the funding that made it all possible. Specifically, also thanks to Maheeka Weernawarna for her assistance with thermal taste and “wet” phenotyping of some of the participants. Thank you to all the participants who completed the multiple sessions of my study, of which at least one involved sticking out their tongues.

Thanks to my friends, including those I have yet to meet in person, including the members of the R4DS, r/GaylorSwift and ADHD friends communities. In particular, thanks to Adam, Annette, Danny, Hannah M, Hannah R, Jessica, Jill, Katie, Marco, Meekashni, Mikana, Petra, Quirina, Sara, Sina, Sinj, and Tara. Also thanks to Mia and Gerhard Uys (and Joey the doubting dog), for always having space for me in their home.

Thanks to my parents, Johan and Benita, for instilling a love of food, science, statistics and logical thinking in me. Growing up, I could never understand why “girls” supposedly don’t believe they can do science. Thank you for your support over the years, and for giving me a space to just be after two years of isolation. Also, thank you to Martha Dube and Fati Diaz for all the meals and practical care during my retreat in the Mpumalanga Lowveld.

Frieda Dehrmann: A decade ago, after having met you once, I sent you an email asking for your views on pursuing a PhD. I could not have imagined where the conversation and the journey of your mentorship and friendship would take us. Thank you for the advice, the support and the many times you told me truths I perhaps did not want to hear. You have given me the map, but also the compass to chart my own course.

Finally, thanks to my partner, Han. You loved me enough to let me go - to literally the other side of the planet from South Africa. Thank you for sticking by me through the longest distance and the many video calls across an 11-hour time difference, and for putting up with my spacey in-your-space-ness while getting this completed. Soon we will be Dr & Dr Booottha/Bottooha.

*“The only thing that’s left is the manuscript, one last souvenir from my trip to your shore”*- Taylor Swift (2024, “The Manuscript”, The Tortured Poets Department: The Anthology)

## Abstract

Thermal taste is a phenomenon whereby some people experience taste from thermal stimulation in the absence of a chemosensory stimulus. Thermal tasters (TT) differ from thermal non-tasters (TnT) in orosensory sensitivity but are also thought to differ from TT in behavioural aspects, including food choice. In addition, TT are hypothesised to differ as a group in their thermal taste experiences. However, traditional phenotyping methods have failed to establish differences in food choice due to their inefficiency (a large proportion of participants are unclassified). In addition, traditional methods are incapable of collecting detailed data on the thermal taste response, as responses are recorded after thermal exposure.

This study developed a new phenotyping method, RapCoTT, which comprises a singular thermal exposure protocol, combined with TCATA data collection. When validated compared to existing methods, this approach was found to be superior in effectiveness and efficiency, and to agree with existing approaches. From data collected through RapCoTT, TT clusters were established, each reporting different thermal tastes and reporting tastes at different temperatures. When modelling thermal taste response data, further evidence was found for the proposed link between TRPM5 and sweet thermal taste. In addition, several novel mechanistic links were discovered, including several potential pathways involving TRPA1.

A group of participants with known thermal taste response groupings were profiled for a range of demographic, food behavioural, personality and orosensory responsiveness variables. Initially, links were explored between TT Status and TT Cluster and these variables, and a machine-learning approach was used to establish the relative importance of predictors. Overall, links between TT Status and TT Cluster and dimensions of the Big Five Inventory (personality) were established, and TT and TnT

were found to differ in their salt and spice addition behaviours. Evidence was also found for several other factors impacting TT Status, including food neophobia, Private Body Consciousness, oral processing behaviour, and aspects of spicy food liking. TT Clusters differed in their hedonic response patterns to sucrose (Sweet liker status clusters). The Set & Disposal subscale of the Food Involvement Scale discriminated TT Clusters.

Finally, food choice behaviour in the form of food frequency data was modelled from thermal taste, demographic, food behavioural, personality and orosensory responsiveness variables. Overall, TT Cluster was the most important variable set related to thermal taste in determining food choice, indicating that not only TT Status, but also the temperature and nature of thermal taste perceived affects food consumption. TT Cluster also affected the specific consumption frequencies of the most items. When clustering food items based on the impact of TT Cluster, items did not group according to predetermined categories, but some textural patterns were apparent.

Overall, this work sets the scene for a paradigm shift around thermal taste, its interaction with other factors and its influence on food behaviour as a whole.

## Table of Contents

Acknowledgements .....	i
Abstract.....	iii
Preface .....	1
Presentations and Publications.....	3
Note on stylistic choices .....	4
Abbreviations used by theme .....	5
List of Figures .....	8
List of Tables .....	18
Chapter 1 Introduction .....	23
1.1 Introduction to thermal taste .....	23
1.2 Thermally induced tastes .....	25
1.2.1 Groups of thermal tasters .....	28
1.3 Sensory responses of Thermal Tasters .....	30
1.3.1 Sensitivity differences .....	30
1.3.2 Impact on food choice .....	36
1.4 Thermal taste phenotyping.....	39
1.4.1 Phenotyping methodologies .....	40
1.4.2 The “unclassifieds” .....	44
1.5 Potential relationships between thermal taste and demographic, food behaviour, personality and orosensory responsiveness.....	49
1.5.1 Demographics (gender, age, ethnicity) .....	49
1.5.2 Food Behaviour .....	50
1.5.3 Personality (BFI/PBC) .....	52

1.5.4 Other Taste phenotypes (orosensitivity).....	53
1.6 Research gaps.....	53
1.7 Thesis objectives and structure.....	54
1.7.1 Thesis structure.....	56
Chapter 2 <i>Fast and Furious</i> : Developing and validating a rapid (fast) portable temporal phenotyping method for phenotyping thermal taste (RapCoTT) with improved efficiency (furious).....	
2.1 Introduction .....	57
2.1.1 Aims .....	61
2.2 Materials & methods .....	61
2.2.1 Design of thermode and associated protocols .....	62
2.2.2 Method validation study.....	66
2.2.3 Data analysis .....	68
2.3 Results.....	72
2.3.1 Classification .....	72
2.3.2 Efficacy of classification.....	73
2.3.3 Consistency across methods .....	74
2.3.4 Temperature trial evaluation.....	77
2.4 Discussion.....	78
2.4.1 A new, more efficient protocol for phenotyping Thermal Taster Status.....	85
2.5 Limitations.....	87
2.6 Conclusion .....	88
Chapter 3 <i>Divide and conquer</i> : Exploring underlying mechanisms of thermal taste by clustering and describing thermal taste response collected via RapCoTT using Mixture Hidden Markov Models and functional data analysis .....	
	90

3.1 Introduction .....	90
3.1.1 Background on thermal taste clusters and mechanisms .....	90
3.1.2 Modelling thermal response data: key concepts .....	92
3.1.3 Aims.....	103
3.2 Materials and Methods.....	104
3.2.1 Participants and protocols .....	105
3.2.2 Data analysis .....	105
3.3 Results.....	115
3.3.1 Relationships between tastes and phases and intra-taste relationships....	116
3.3.2 Clustering of thermal taste response.....	118
3.3.3 Taste and time relationships .....	123
3.4 Discussion.....	128
3.4.1 General Patterns of taste response .....	129
3.4.2 Potential mechanistic links.....	135
3.4.3 Clusters of thermal taste responses.....	141
3.5 Limitations.....	146
3.6 Conclusion .....	147
 Chapter 4 <i>Seeing the forest and the trees</i> : Demystifying complex relationships between thermal taste response and demographic, personality, behavioural and orosensory responsiveness variables .....	 149
4.1 Introduction .....	149
4.1.1 Using machine learning to investigate interrelated factors.....	150
4.1.2 Interpreting machine learning models .....	153
4.1.3 Aims.....	156
4.2 Materials and Methods.....	157

4.2.1 Recruitment, data collection & data processing .....	157
4.2.2 Online survey .....	158
4.2.3 Taste phenotyping .....	165
4.2.4 Data analysis .....	167
4.3 Results.....	175
4.3.1 Participants .....	175
4.3.2 Models.....	175
4.3.3 Demographics .....	175
4.3.4 Food behaviour (Survey 1 +2) .....	176
4.3.1 Personality (Survey 3) .....	185
4.3.2 Orosensory response (Taste Phenotyping) .....	190
4.3.1 Results summary .....	195
4.4 Discussion .....	196
4.4.1 Modelling discussion.....	196
4.4.2 Relationships between variables of interest and thermal taste .....	198
4.5 Limitations.....	207
4.6 Conclusion .....	209
Chapter 5 <i>The path to enlightenment: The relative impact of factors mediating food choice behaviour, including Thermal Taste</i> .....	210
5.1 Introduction .....	210
5.1.1 Potential mechanisms of thermal taste's influence on food choice .....	211
5.1.2 Influence on food choice of factors potentially related to thermal taste... ..	212
5.1.3 Specific considerations around thermal taste.....	219
5.1.4 Modelling background .....	219
5.1.5 Aims .....	225

5.2 Materials and methods .....	226
5.2.1 Data Collection .....	227
5.2.2 Data analysis .....	231
5.3 Results.....	235
5.3.1 Overall model results .....	236
5.3.2 Focus on data blocks .....	241
5.3.3 Impact of aspects of TT on food consumption prediction .....	262
5.4 Discussion .....	271
5.4.1 Relative predictive effect of selected factors on food choice .....	272
5.4.2 Impact of thermal taster status (Cluster) on food choice .....	280
5.4.3 Model fitting.....	281
5.5 Limitations .....	281
5.6 Conclusion .....	282
Chapter 6 Discussion and future work .....	283
6.1 Introduction .....	283
6.2 Summary.....	284
6.2.1 Chapter 2: Fast and Furious .....	284
6.2.2 Chapter 3: Divide and conquer.....	285
6.2.3 Chapter 4: Seeing the forest and the trees .....	288
6.2.4 Chapter 5: Path to enlightenment .....	291
6.3 Strengths and Limitations .....	293
6.3.1 RapCoTT method .....	293
6.3.2 Data collection .....	295
6.3.3 Timing .....	296
6.3.4 Modelling .....	297

6.4 Future work .....	299
6.4.1 Better thermal taste study.....	299
6.4.2 Food behavioural studies.....	300
6.4.3 Genetic study on thermal taste.....	302
6.4.4 Further factors to assess .....	303
6.5 Conclusion .....	303
References .....	305
Appendix A. Ethics approval SOA 19/26 .....	332
Appendix B. Ethics approval SOA 20/40 .....	333
Appendix C. RapCoTT method design documentation .....	335
Appendix D. Additional tables RapCoTT.....	346
Appendix E. FDA derivatives .....	348
Appendix F. Mathematical details of Fatty acid detection calculation .....	350
Appendix G. Full distributions of relevant variables.....	351
Appendix H. ALE plots binary model .....	353
Appendix I. ALE plots multiclass model.....	360
Appendix J. Changes in explained variance ROSA model .....	376
Appendix K. Contributions of datablocks distinct model .....	377
Appendix L. Contributions of data blocks common component model .....	381
Appendix M. Variable name abbreviations .....	385
Appendix N. Variable importance values .....	390
Appendix O. Additional x-loadings plots.....	392

## Preface

Thermal taste is a phenomenon whereby a proportion of the population experiences taste from thermal stimuli in the absence of a chemosensory stimulus. Although the cause of thermal taste is yet to be established, temperature-sensitive taste channels are thought to be involved in the phenomenon. Thermal tasters (TT) are typically identified through two separate thermal exposures (cooling and warming), after which any perceived tastes are reported. However, the traditional approach is highly inefficient in phenotyping TT. TT differ from thermal non-tasters (TnT) in terms of orosensory sensitivity. In addition, there is evidence that TT differ as a group in their thermal taste experiences, and there is preliminary evidence that TT and TnT differ in personality and food attitudes. It is also expected that TT differ from TnT in their food choice behaviour due to different orosensory experiences, although limited sample sizes, resulting indirectly from inefficient phenotyping which has been a serious limitation of previous studies.

This PhD aimed to challenge the existing paradigm of thermal taste as a singular phenotype. This aim was achieved through two key enablers: developing a new phenotyping method and using advanced modelling. The new phenotyping method, RapCoTT, combined the two temperature trials into one dynamic exposure, combined with TCATA data collection. Data collected in this way was used to establish links between specific thermal tastes, temperatures and mechanisms. This data was also used to establish groups of TT with different response patterns. In addition, a group of participants phenotyped for thermal taste were profiled for demographic, food behaviour, personality and orosensory responsiveness. This data was used to establish differences between TT subgroups. Finally, links between thermal taste and food choice were established by modelling food frequency from thermal taste,

demographic, food behaviour, personality and orosensory responsiveness data.

The thesis is structured as follows. **Chapter 1** reviews the literature on thermal taste, thermal taster phenotyping, and factors potentially related to thermal taste. **Chapter 2** outlines the development of a novel TT phenotyping method, RapCoTT, and its validation as a superior alternative to existing approaches. **Chapter 3** establishes subgroups of TT based on their thermal taste experiences and details the extraction of thermal taste and temperature patterns. **Chapter 4** details the profiling of the group of participants clustered in Chapter 3 for demographic, food behavioural, personality and orosensory responsiveness variables and establishes the relative importance of each variable in determining TT Status or TT Cluster memberships. **Chapter 5** models food choice, in the form of food frequency data from the data obtained in chapters 3 and 4. Finally, **Chapter 6** summarises the findings of the thesis and proposes avenues for future work.

## Presentations and Publications

### Conference presentations:

Botha, J. J. & Hort, J. (2019, February 3-5). *Challenges across the African continent: Developing robust approaches to sensory and consumer science research* [Conference Presentation]. 13th NZOZ Sensory Consumer Science Symposium, Dunedin, New Zealand

Botha, J. J., Cannon, P. & Hort, J. (2020, February 10-12). *An improved approach for phenotyping thermal taster status* [Conference Presentation]. 14th NZOZ Sensory Consumer Science Symposium, Melbourne, Australia

Botha, J. J., Cannon, P., Hort, J. (2021, February 9-11). *Impact of Thermal taste on food temperature preference* [Conference Presentation]. 15th NZOZ Sensory and Consumer Science Symposium, Online

Botha, J. J., Cannon, P., Hort, J. (2020, December 13–16). *RapCoTT: A new, more efficient method for Thermal Taste Phenotyping* [Poster Presentation]. 9th European Conference on Sensory and Consumer Research, Online

Botha, J. J., Cannon, P., Hort, J. (2021, August 9-12). *Do only some like if hot? Comparing food and beverage temperature preference between thermal tasters and thermal non tasters* [ECR Conference presentation]. 14th Pangborn Sensory Science Symposium, Online

Botha, J. J., Cannon, P., Hort, J. (2019, September 30). *For some, heat tastes sweet: Preliminary results from an improved protocol for studying Thermal Taste* [Conference Presentation]. Riddet Institute Student Colloquium, Rotarua, New Zealand

### Journal articles:

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>

## **Note on stylistic choices**

Several stylistic choices were made in this thesis to avoid confusion to the reader.

Thermal taste, as a phenomenon was not capitalised throughout, a choice contrary to the convention in thermal taste literature. This was to facilitate discussion of thermal taste in the same context as chemosensory taste, with similar validity. Similarly, it did not make sense to capitalise thermal taste when discussing a specific thermal taste. Therefore, to not have two different capitalisation conventions, the convention of not capitalising thermal taste was used.

Thermal taster status was abbreviated as TT Status, and not as TTS, as commonly used in literature. Here the choice was driven to facilitate discussion of different types of thermal taster groupings, including TT Clusters.

Three terms were used to discuss different groups of thermal tasters. TT groups refer to all groupings of thermal tasters (TT), including TT Status groups and different subgroups of TT. TT subgroups was used to refer to subgroups of TT conceptually. In contrast, TT Clusters was used when specifically referring to the clusters of TT obtained in Chapter 3.

Outside the realm of thermal taste, choices regarding abbreviations were made to avoid confusion to the reader. Due to the breadth of subject matter, abbreviations used were reintroduced in each chapter. In addition, unnecessary use of abbreviations was avoided in many cases for the sake of clarity.

Also note that all p-values were reported to two significant figures, to a maximum of three decimal places.

The researcher uses they/she as their personal pronouns, and chose to use the singular “they” throughout, not only when referring to themselves, but also when referring to all other researchers.

## Abbreviations used by theme

### Thermal taste

RapCoTT	Rapid Combined Phenotyping for Thermal Taste
TnT	Thermal non taster
TT	Thermal taster
TT Cluster	Thermal taster Cluster
TT Status	Thermal taster status
UC	Unclassified Thermal Tasters

### Taste Phenotyping

FA	Fatty Acids
HSL	High Sweet Liker
LSL	Low Sweet Liker
MSL	Medium Sweet Liker
PNT	PROP non Taster
PROP	6-n-propylthiouracil
PST	PROP strong Taster
SLS	Sweet Liker Status
UN	Unclassified Sweet likers

### Sensory Methods

gLMS	Generalised line magnitude scale
gVAS	Generalised visual analogue scale
LAM	Labelled Affective Magnitude Scale
TCATA	Temporal Check-All-That-Apply
TDS	Temporal Dominance of sensations

### **Food behaviour/personality**

BFI	Big Five Inventory
FFQ	Food Frequency Questionnaire
FIS	Food Involvement Scale
FNS	Food Neophobia Scale
P&E	Preparation and Eating
PBC	Private Body Consciousness
S&D	Set and Disposal

### **Taste Mechanisms**

ENaC	Epithelial Na <sup>+</sup> (Sodium) channel
OTOP1	Otopetrin1
TAS1R(2/3)	Taste Receptor type 1 member
TMC(4)	transmembrane channel-like
TRP	transient receptor potential
TRPA1	transient receptor potential akryn
TRPC(5)	transient receptor potential canonical
TRPM(4/5/8)	transient receptor potential melastatin
TRPV(1/4)	transient receptor potential vanilloid
ENaC	Epithelial Na <sup>+</sup> channel
OTOP1	Otopetrin1

### **Statistics**

AHC	Agglomerative hierarchical clustering
AIC	Aikeke Information criterion
ALE (plots)	Accumulated local effects (plots)
BIC	Bayesian Information Criterion

CI	Confidence interval
coeff	Coefficient
fda	Functional data analysis
fdr	False discovery rate
GCV	Generalised cross-validation
HMM	Hidden Markov model
MHMM	Mixture Hidden Markov model
OR	Odds ratio
PCA	Principle Component Analysis
PLS	Partial Least Squares
PMT	PROP medium Taster
RFM	Random forest model
ROSA	Response Oriented Sequential Alternation
VIP	Variable importance on projection

**General**

°C	Degrees Celsius
BMI	Body Mass Index
MSG	Monosodium glutamate
NZ	New Zealand
s	seconds
SNP	Single nucleotide polymorphism
SOA	Massey Human Ethics Southern A
Tbsp	Tablespoon
tsp	Teaspoon
yrs	Years

## List of Figures

- Figure 1-1 Visual representation of the temperature exposures of the Green and George (2004), Bajec and Pickering (2008) and Skinner et al. (2018) protocols. In areas greyed out, participants were instructed to attend only to the temperature change..... 26
- Figure 1-2 Summary of tastes reported by thermal tasters during cooling or warming of the tongue. Note that in the studies which note “all”, the authors made no distinction between the cooling and the warming trials. .... 27
- Figure 1-3 An example of the General Linear Magnitude Scale (gLMS). Participants are asked to first rank a number of familiar sensations on the scale (for example the brightness of the sun) prior to ranking tastes. As the markings on the scale are quasi-logarithmic data is often log transformed. The labels on the scale and their relative positions are: barely detectable, 1.4; weak, 6.1; moderate, 17.2; strong, 35.4; very strong, 53.3; strongest imaginable, 100. (Green et al., 1996)..... 43
- Figure 1-4. The expanding definition of Thermal Taste a) Green and George (2004) Perceive tastes between 15°C and 35°C plus reports taste consistently b) Bajec and Pickering (2008) Perceive tastes between 5°C and 40°C plus has no reports taste consistently c) Skinner et al. (2018) Perceive tastes between 5°C and 40°C but may be inconsistent in reporting taste d) Potentially expanded definition of Thermal Taste. .... 45
- Figure 2-1 Overview of the system, connected via Bluetooth a) Laptop computer running joint thermode control/data collection software with primary display for researcher b) Secondary display and mouse (participant interface) c) Thermode consisting of a control box, the thermode itself and power supply d) detailed picture of the tip of the thermode showing the Peltier device sandwiched between a copper plate and the aluminium rod (covered with tasteless plastic film during phenotyping). .... 63
- Figure 2-2 Summary of a) and b) the traditional protocol and c) the combined protocol. The traditional protocol consists of two trials, a) warming (35°C – 15°C - 40°C) and b) cooling (35°C – 5°C held for 5s). c) The combined protocol only consists of one temperature trial, which has a cooling portion (35°C -5°C) a warming portion (5°C -40°C) and return or end portion (40°C -35°C). .... 65

Figure 2-3 Agreement plots comparing each pair of methods using the full dataset (a-c) and subset (d-i) according to which protocol was performed first. In each plot, the axes are split into three categories (TnT, UC, TT) according to the number of participants classified in that category. The grey areas show the overlap between methods. Black squares are equal to the number of participants classified the same by both methods. Dark grey shading represents those classified in an adjacent category. The diagonal line is an interpretation guide for directional bias..... 75

Figure 2-4 The RapCoTT method, showing three training steps namely taste training, TCATA training and a practice run with the protocol. This is followed by two replicates of a temperature exposure (35°C, 5°C, 40°C 1°C/s) with concurrent TCATA data collection..... 85

Figure 3-1 Overview of the steps taken to determine the number of clusters in thermal taste responses, showing the data processing steps (standardisation, binning and creation of state sequences) and the clustering steps (creation of a distance matrix and hierarchal clustering)..... 109

Figure 3-2 Modelling process, which shows how the final MHMM was obtained. The left shows the iterative process using the genetic algorithm, which determines the number of states per cluster. On the right, the process of specifying (building) and fitting the model is shown, resulting in cluster membership and hidden states..... 112

Figure 3-3 Relationships between tastes and trial portions when only duplicate tastes are considered for Thermal Taster classification, labelled with the odds ratio (OR). Line thickness is proportional to the strength of association in the form of odds ratio or inverse of odds ratio in the case of negative association ( $OR < 1$ ). Negative associations ( $OR < 1$ ) are indicated with a dashed line, and positive associations ( $OR > 1$ ) with a solid line. Light grey lines are less significant. ( $0.01 < p < 0.05$ ) and black lines are more significant ( $p < 0.01$ ). Figure after Thibodeau et al. (2020)..... 116

Figure 3-4 Relationships between tastes and trial portions when a) responses are pooled per participant b) responses are considered individually. Thickness of lines are proportional to the strength of association in the form of odds ratio or inverse of odds ratio in the case of negative association ( $OR < 1$ ). Negative associations are indicated with a dashed line, and positive associations with a solid line. Light grey lines are less significant. ( $0.01 < p < 0.05$ ) and black lines are more significant ( $p < 0.01$ ). Figure after Thibodeau et al. (2020). ..... 117

Figure 3-5 Dendrogram obtained from the exploratory hierarchical clustering of the state sequence data, using the dissimilarity matrix obtained using Hamming distances, highlighting the three clusters..... 119

Figure 3-6 Distributions of cluster probabilities of individuals considering two approaches to final clustering; a) Assignment to cluster with highest mean across two replicates b) Assignment to highest affinity across two replicates..... 120

Figure 3-7 Emission probabilities of each of the hidden states for a) Cluster 1, b) Cluster 2 and c) Cluster 3. The size of the bubbles is proportional to the emission probabilities and are also coloured accordingly. A taste was considered to characterise a state if the emission probability exceeded 0.7. Note that the numbers do not add up to 1, as there is an accompanying (complementary) emission probability for a taste not being reported. .... 121

Figure 3-8 TCATA difference plot showing the citation proportion difference between Cluster 1 and the entire sample. Significance levels were obtained using bootstrapped confidence intervals. Only areas of the plot where a taste was reported more frequently than the overall sample are shown. Shaded numbers highlight start and end temperatures of significant portions. .... 124

Figure 3-9 Functional curves obtained from Metallic and Salty reported frequencies for responses contained in Cluster 1. Numbers in bold indicate temperatures where local maxima or minima are reached (first derivative zero), and boxed numbers indicate “trigger points” where the first derivative reaches a local minimum/indicating (second derivative zero) indicating that the taste is being frequently clicked or unclicked. Solid lines indicate temperatures where the taste in question was more frequently reported by Cluster 1 (Figure 3-8) ..... 124

Figure 3-10 TCATA difference plot showing the citation proportion difference between Cluster 2 and the entire sample. Significance levels were obtained using bootstrapped confidence intervals. Only areas of the plot where a taste was reported more frequently than the overall sample are shown. Shaded numbers highlight start and end temperatures of significant portions. .... 125

Figure 3-11 Functional curves obtained from Savoury and Sour reported frequencies for responses contained in Cluster 2. Numbers in bold indicate temperatures where local maxima or minima are reached, and boxed numbers indicate “trigger points” where the first

derivative reaches a local minimum/indicating points where the taste is either being clicked or unclicked. Solid lines indicate temperatures where the taste in question was more frequently reported by Cluster 2 (Figure 3-10).....	126
Figure 3-12 TCATA difference plot showing the citation proportion difference between Cluster 3 and the entire sample. Significance levels were obtained using bootstrapped confidence intervals. Only areas of the plot where a taste was reported more frequently than the overall sample are shown. Shaded numbers highlight start and end temperatures of significant portions. ....	127
Figure 3-13 Functional curves obtained from Bitter and Sweet reported frequencies for responses contained in Cluster 3. Numbers in bold indicate temperatures where local maxima or minima are reached, and boxed numbers indicate “trigger points” where the first derivative reaches a local minimum, indicating points where the taste is either being clicked or unclicked. Solid lines indicate temperatures where the taste in question was more frequently reported by Cluster 3 (Figure 3-12).....	128
Figure 4-1 Flow chart showing the modelling approach for ensemble RFM fitting. Shapes are used to group similar steps or concepts, with all variables and parameters denoted as diamond shapes.....	170
Figure 4-2 Flowchart explaining the approach taken for hyperparameter tuning during RFM fitting. ....	172
Figure 4-3 Flowchart showing decision process for inclusion of variables in subsequent RF models	174
Figure 4-4 Distribution of the Age data according to TT Status showing boxplots and means, as well as summary statistics for the Kruskal-Wallis test.....	176
Figure 4-5 Accumulated local effects (ALE) plot for oral processing behaviour included in the final binary (TT Status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept +/- 0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.....	177
Figure 4-6 Accumulated local effects (ALE) plots for items from the food involvement scale (FIS) impacting the final multiclass (TT Cluster) RFM, showing chance probability ( 0.25; baseline cluster probability), four guides (chance +/- 0.01 and chance +/- 0.005), and the	

effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability..... 179

Figure 4-7 Accumulated local effects (ALE) plots for items related to the food neophobia scale (FNS) included in the final binary (TT Status) RFM which impact TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability..... 180

Figure 4-8 Accumulated local effects (ALE) plots for items and dimensions from the food neophobia scale (FNS) included in the final multiclass (TT Cluster) RFM which impact cluster probability. These plots show chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. .... 181

Figure 4-9 Distribution of the food modification raw data (indicated by prefix FMRave) regarding a) salt addition and b) heat addition according to TT Status, showing boxplots, means, summary statistics for the Kruskal-Wallis test. .... 182

Figure 4-10 Accumulated local effects (ALE) plots for calculated indexes from the food modification questionnaire in the final binary (TT Status) RFM which impact TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. The prefix "FMR" indicates raw values (scale items not concatenated)..... 183

Figure 4-11 Accumulated local effects (ALE) plot for calculated raw (not concatenated) indexes from the food modification questionnaire included in the final multiclass (TT Cluster) RFM which impact TT Cluster probability. This plot shows chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. .... 184

Figure 4-12 Accumulated local effects (ALE) plots for the items from the spicy food questionnaire included in the final binary (TT Status) RFM which impact TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability..... 185

Figure 4-13 Distribution of five factors of the BFI (BFI\_A = Agreeableness, BFI\_E = Extraversion, BFI\_C = Conscientiousness, BFI\_N = Neuroticism, BFI\_O = Openness to experience) according to a) TT Status showing boxplots and means, as well as summary statistics for the Kruskal-Wallis test. .... 186

Figure 4-14 Distribution of five factors of the BFI (BFI\_A = Agreeableness, BFI\_E = Extraversion, BFI\_C = Conscientiousness, BFI\_N = Neuroticism, BFI\_O = Openness to experience) according to ) TT Cluster showing boxplots and means, as well as summary statistics for the Kruskal-Wallis test. and pairwise comparisons (results of the Wilcox-test) corrected for the false discovery rate (fdr) ..... 186

Figure 4-15 Accumulated local effects (ALE) plots for personality factors and items from the Big Five Inventory (BFI) included in the final binary (TT Status) RFM which impact TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. Note that individual items are prefixed by the factors to which they contribute. .... 188

Figure 4-16 Accumulated local effects (ALE) plots for personality items and factors from the Big Five Inventory (BFI) included in the final multiclass (TT Cluster) RFM impacting cluster probability. These plots show chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that individual items are prefixed by the factors to which they contribute. .... 189

Figure 4-17 Accumulated local effects (ALE) plots for Private body consciousness (PBC) and the items from the PBC questionnaire included in the final binary (TT Status) RFM impacting TT

probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. .... 190

Figure 4-18 Accumulated local effects (ALE) plots for sweetness response variables (liking, intensity, SLS cluster) included in the final binary (TT Status) RFM impacting TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. Note that the prefix “sls” indicates liking on the LAM scale. .... 192

Figure 4-19 Accumulated local effects (ALE) plots for sweetness response variables (liking, intensity, SLS cluster) included in the final multiclass (TT Cluster) RFM impacting cluster probability. This plot shows chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that the prefix “sls” indicates liking on the LAM scale. .... 193

Figure 4-20 Accumulated local effects (ALE) plots for PROP responsiveness variables included in the final binary (TT Status) RFM, showing the model intercept ( 0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. .... 193

Figure 4-21 Accumulated local effects (ALE) plots for Fatty acid response variables included in the final binary (TT Status) RFM, impacting TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. .... 194

Figure 4-22 Accumulated local effects (ALE) plots for Fatty acid response variables included in the final multiclass (TT Cluster) RFM impacting cluster probability. These plots show chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$

0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. .... 195

Figure 5-1 Graphical representation of the a) distinct component and b) common component ROSA models. Each grey line indicates the prediction trajectory (increased explained variance with each fitted component) of a single food item. The black line indicates total explained variance of the X space (blocks used for prediction). The final fitted component prior to destabilisation is highlighted with a vertical line, as are the items with the highest and lowest proportion of explained variance at that point. For figures detailing the prediction trajectory of each item, see Appendix K and Appendix L..... 237

Figure 5-2 Plot of unadjusted  $R^2$  (black line) and cross-validated  $R^2$  (red line) for all items using the distinct component model, with the final number of components (20) chosen highlighted with a vertical dashed grey line. The higher the unadjusted  $R^2$ , the more variation in the data explained, whereas the cross-validated  $R^2$  indicates model fit. Negative cross-validated  $R^2$  indicates an overfitted model. .... 238

Figure 5-3 Plot showing a) loading weights and b) x-loadings for the components extracted from the TT Cluster data block. Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M ..... 244

Figure 5-4 Plot showing x-loadings for Component 2 (Demographics) and Component 2 (Orosensory response). Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M ..... 248

Figure 5-5 Plot showing x-loadings for Component 3 (Spicy Food Liking), Component 4 (Spicy Food Liking) and Component 5 (Oral processing). Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M ..... 251

Figure 5-6 Plot showing x-loadings for Component 7 (food modification), Component 8 (food involvement) and Component 9 (Oral Processing). Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M ..... 254

Figure 5-7 Plot showing x-loadings for Components 10, 11 and 12, extracted from the Sweet Liking data block. Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M ..... 257

Figure 5-8 Clustering of different food items based on the impact of thermal taste on the consumption. Items are colour-coded according to their categories as described in section 5.2.1.3 ..... 263

Figure 5-9 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster A1, b) subcluster A2, c) subcluster A3, d) subcluster A4, e) subcluster A5 and f) subcluster A5. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2 ..... 265

Figure 5-10 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster A7 and b) subcluster A8. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2 ..... 266

Figure 5-11 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster B1, b) subcluster B2, c) subcluster B3, d) subcluster B4, e) subcluster B5 and f) subcluster B6. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2. .... 268

Figure 5-12 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster C1, b) subcluster C2 and c) subcluster C3. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2. .... 270

Figure 5-13 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster C7 and b) subcluster C8. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients.

The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section  
5.3.2..... 271

## List of Tables

Table 1-1 Chronological summary of studies on TT Status and orosensory responsiveness in non-food matrices. All results reported are significant for TT Status unless otherwise specified. ....	31
Table 1-2 Chronological summary of studies on TT Status and orosensory responsiveness in food and beverage matrices. All results reported are significant for TT status unless otherwise specified.....	33
Table 1-3. Chronological summary of studies exploring TT status and food choice. All findings are significant for TT Status unless stated otherwise.....	37
Table 1-4 Summary of methodologies, devices and strategies used for different Thermal Taste classification methods. Adapted from Thibodeau et al. (2019). ....	41
Table 1-5 Summary of classification strategies and percentages reported for different Thermal Taste classification methods. Adapted from Thibodeau et al. (2019). ....	42
Table 1-6. Hypothetical experiences and situations that may cause individuals to be unclassified under some classification schemes, as well as their potential classifications under classification schemes proposed by Bajec and Pickering (2008), Skinner et al. (2018) and Thibodeau et al. (2019). ....	47
Table 2-1 Tastes and concentrations used for taste familiarisations. Adapted from Skinner (2017) ..	67
Table 2-2 Summary of percentages (numbers of participants) classified as either TT or TnT by two schemes (Scheme A + Scheme B) under the traditional protocol as well as the combined protocol (Scheme C) .....	72
Table 2-3. Agreement of each pair of methods and classification category using Cohen's $\kappa$ (weighted and unweighted) as well as interpretation. Weighted Cohen's $\kappa$ allows for misclassification of one category (e.g. TnT to UC, UC to TT). Cohen's K, diagnostic categories: poor ( $\kappa < 0.200$ ), fair ( $\kappa = 0.201-0.400$ ), moderate ( $\kappa = 0.401-0.600$ ), good ( $\kappa = 0.601-0.800$ ), or very good ( $\kappa = 0.801-1.0$ ).....	74
Table 2-4 Summary of approximated times for the traditional protocol, RapCoTT as applied during the present study, as well as proposed changes to RapCoTT .....	78

Table 3-1 Potential mechanisms and related temperatures for thermal taste and related sensations .....	92
Table 3-2 Parameters of Mixture Hidden Markov Models. Abbreviations: Spec = specified. Res = Results .....	99
Table 3-3 Parameters used to input into the MHMM, showing an example for a four-state cluster. Note that the initial probabilities indicate a constrained model, which “forces” the process to start in State 1. The emission probabilities are unconstrained, but the starting values for State 1 are set to indicate that it is expected that State 1 would be dominated by “no sensation”. Transition probabilities were set to 1/n, which means that these had to be modelled from scratch. Each state refers to a distinct experience, linked to a likely taste mechanism. ....	110
Table 3-4 Summary of the number of tastes reported per participant and contained per response. All the tastes reported in both replicates are considered when participants are counted. Each replicate x participant is treated as an experimental unit when observations are counted. .....	116
Table 3-5 Contingency table of classification of participants in the same cluster across the two different replicates .....	120
Table 4-1 Items on the food involvement scale (FIS) (Bell & Marshall, 2003). Reversal (R) is indicated, as well as the existing subscales (Bell & Marshall, 2003) (Preparation & Eating and Set & Disposal). and the additional subscales tested in this study (Cooking, Talking and Planning).....	160
Table 4-2 Items of the Food Neophobia Scale (FNS) used to assess Food Neophobia. Reversal (R) is indicated, as well as loading on the Approach and Avoidance subscales (Nezlek & Forestell, 2019). Note that by default, items load to the approach and avoidance subscales as is, except where reversal is specifically indicated. ....	161
Table 4-3 Questions asked for food modification. Questions were phrased in the format “how often do you?” .....	162
Table 4-4 Items used to assess Private body Consciousness (PBC), as detailed in Miller et al. (1981). Note that none of the items are reversed, and no subscale is present. ....	164

Table 4-5 Items of the Big Five Inventory (BFI) (John et al., 2008), used to assess personality. Reversal (R) is indicated, as well as factor to which each item loaded (Agreeableness, Conscientiousness, Extraversion, Neuroticism, Openness to experience). Each item was presented in the format: “I am someone who” .....	165
Table 4-6 Maximum entropy hypercube sizes used for hyperparameter tuning.....	171
Table 4-7 Variable importance measures used for variable selection for subsequent RFM. (AOC = Area Under Curve).....	173
Table 4-8 Variables removed from the final RFM due to high correlation.....	175
Table 4-9 Variables and constructs from the Food Involvement questionnaire included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs. Note that the Talking, Cooking and Planning constructs all form part of the Preparation & Eating subscale. ....	178
Table 4-10 Variables and constructs from the Food Neophobia scale included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.....	180
Table 4-11 Calculated mean Food Modification responses included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs. ....	183
Table 4-12 Responses from the Spicy Food Liking questionnaire included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.....	184
Table 4-13 Variables and constructs from the Big Five Inventory personality questionnaire included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs. ....	187
Table 4-14 Variables from the Private Body Consciousness questionnaire included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs. ....	189
Table 4-15 Variables from Sweet Liker Status phenotyping (including hedonic and intensity responses) included in the final binary (TT Status) and multiclass (TT Cluster) RFMs.	

Variables were conservatively selected based on their importance in the two previous RFMs.....	191
Table 4-16 Variables (including responses to specific oleic acid concentrations) related to fatty acid responsiveness included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.....	194
Table 5-1 Steps in each iteration of the ROSA algorithm, focussed on component extraction and component/block selection.....	222
Table 5-2 Data blocks for input into food choice prediction model. Size refers to the dimensions of the data block, which refers to either the number of recoded categories or the number of variables in the data block. The category (Cat) refers to the data block categorisation in this chapter; Thermal Taste (T), Unrelated to thermal taste or food behaviour (N), Direct measures of food behaviour (D) or Indirect measures of food behaviour (I). .....	228
Table 5-3 Dairy, Fruit, Vegetable and Protein items included in the Food Frequency Questionnaire, including their serving sizes and examples as presented to participants .....	229
Table 5-4 Cereals, Fast Food and Sweets, Baked Goods and Miscellaneous items included in the Food Frequency Questionnaire, including their serving sizes and examples as presented to participants.....	230
Table 5-5 Component selection order by the distinct and common component model. Component number indicates the position in the model that the component was selected in, with lower numbers (earlier selection) indicating higher importance to the model. In the common component model, where two data blocks are listed, this indicates the selection of a component from the joint space of these two data blocks. Category (Cat) refers to whether the data block is related to thermal taste (T), Direct measures of food behaviour (D), indirect measures of food behaviour (I) or unrelated to either thermal taste or food behaviour (N).....	240
Table 5-6 Ten variables from the distinct component ROSA model with the highest variable importance in projection scores.....	242

Table 5-7 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the TT Cluster data block. The sign of the coefficient (+/-) relates the impact on prediction (positive/negative)..... 246

Table 5-8 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Demographics data block. The sign of the coefficient (+-) relates the impact on prediction (positive/negative) ..... 247

Table 5-9 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Orosensory sensitivity data block. The sign of the coefficient (+-) relates the impact on prediction (positive/negative).)..... 249

Table 5-10 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Spicy Food Liking, Oral Processing, Food Modification and Food involvement data blocks. The sign of the coefficient (+-) relates the impact on prediction (positive/negative). ..... 252

Table 5-11 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Sweet Liking data block. The sign of the coefficient (+-) relates the impact on prediction (positive/negative)..... 259

Table 5-12 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Big Five Inventory data block. The sign of the coefficient (+-) relates the impact on prediction (positive/negative). ..... 260

Table 5-13 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Food Neophobia data block. The sign of the coefficient (+-) relates the impact on prediction (positive/negative). ..... 262

## **Chapter 1 Introduction**

This review starts with a general overview of thermal taste, followed by specifics of thermally induced tastes. Subsequently, the impact of TT Status on perception sensitivity is explored, and the development of phenotyping methodologies and their impact is discussed. Identified gaps and resulting research questions are then presented.

### **1.1 Introduction to thermal taste**

The concept that different individuals have different taste experiences was first discovered in the 1930s when Blakeslee and Fox (1932) found that people differ in their sensitivity to PTC (phenylthiocarbamide) or PROP (6-n-propylthiouracil). PROP sensitivity has been well-studied regarding its impact on food choice and health (see Hayes & Keast, 2011; Reed, 2008), but PROP sensitivity is not the only individual difference in taste experiences. For example, a group of people experience oral irritants as bitter (Green & Schullery, 2003), and there are differences in sweetness perception (Looy et al., 1992). Different taste experiences likely lead to different food choices, resulting in different health outcomes (Chamoun et al., 2018).

In addition to differences in orosensory sensitivity, differences in food preference have been estimated. Individuals can be segmented based on how they interact with food texture (oral processing behaviour)– and their related texture preferences (Jeltema et al., 2015). There are also clear groups of optimal sweetness liking (Pangborn, 1970). These differences have impacted how sensory scientists think about working with sensory panels and consumers.

One known but lesser-studied area of difference is differences in how people respond to temperature, which is particularly relevant as temperature is an inherent property

of all food. Temperature influences taste perception; a phenomenon studied since the 1860's (Pangborn et al., 1970) and comprehensively reviewed in Lemon (2017). However, the direct study of the influence of temperature on taste perception of the tongue (Green & Frankmann, 1987) led to the discovery of thermal taste – the phenomenon that, in some individuals, taste can be induced by a change in applied temperature (Cruz & Green, 2000). Typically, thermal taste is induced by applying a temperature-changing device (called a thermode) to the tongue but has also been induced in consumption situations (Yang, 2015) and through changing water temperature (Nachtigal & Green, 2020). Thermal tasters, individuals who experience a thermally induced taste, are estimated to make up between 20% (Yang et al., 2014) and 50% (Pickering & Klodnicki, 2016) of the population, although proportions as high as 80% have been reported (Green & George, 2004). (See section 1.4 for a full review).

Green and George (2004) initially observed that thermal tasters (TT) showed higher sensitivity to PROP than Thermal non Tasters (TnT), but this enhancement presents independently to PROP taster status (Bajec & Pickering, 2008; Thibodeau et al., 2019; Yang et al., 2014) and presents independently from the genes connected to PROP sensitivity (Bering et al., 2013). In addition, TT Status is unrelated to other physiological taste differences like fungiform papillae density (Bajec & Pickering, 2008; Eldeghaidy et al., 2018; Yang, 2015) and saliva flow rate (Bajec & Pickering, 2008). No relationship has been found between smoker status (Bajec & Pickering, 2008) or BMI (Bajec & Pickering, 2010). There is also strong evidence that TT Status is unrelated to sex/gender (Bajec & Pickering, 2008; Pickering & Kvas, 2016; Pickering et al., 2016; Thibodeau et al., 2019; Yang, 2015; Yang et al., 2020). (The term sex/gender is used to denote the overlapping physiological and psycho-social-cultural constructs (see Spence, 2019)). One study (Thibodeau et al., 2019) has found a

relationship between age and TT Status, with TT being slightly younger. In addition Yang et al. (2020) found a relationship between TT Status and ethnicity, with TT being more likely to be of Asian ethnicity.

The mechanisms behind thermal taste have yet to be clarified. Green and George (2004) initially hypothesised that thermal taste occurs through a “central gain” mechanism and proposed an integration in TT between taste and nociceptive response, but their study with capsaicin indicated otherwise (Green et al., 2005). Evidence was, however, found of central gain in a future study which used functional magnetic resonance imaging (fMRI) to examine the neural correlates of TT when exposed to CO<sub>2</sub> (Hort et al., 2016). A further hypothesis is that thermal taste is under genetic control and related to the temperature-dependency of taste responses. Temperature-activated taste channels have been linked to the thermal taste phenomenon, for example, TRPM5 for sweet (Talavera et al., 2005) and TRPV1 for salty (Talavera et al., 2007), and are commonly cited as a potential mechanism (for example Bajec & Pickering, 2008; Skinner et al., 2018; Thibodeau & Pickering, 2019). Skinner et al. (2018) found that thermal sweet taste frequently occurs between 22°C and 35°C during warming, which is in the same temperature range that activates the TRMP5 channel (15-35°C) (Talavera et al., 2005). However, this finding fails to fully explain the wider variety of tastes, produced during temperature stimulation. Therefore, thermal taste likely occurs due to a spectrum of different responses with several underlying mechanisms (Skinner et al., 2018).

## **1.2 Thermally induced tastes**

TT Status is identified by applying a temperature-changing device (called a thermode) to the tip of the tongue (Bajec & Pickering, 2008; Cruz & Green, 2000). Traditional

phenotyping consists of two trials. In the warming trial, the tongue is cooled and subsequently warmed up, with attention on tastes, if perceived, being reported for the warming phase. In the cooling trial, the tongue is cooled, and the process is repeated.

Figure 1-1 summarises the different phenotyping trials.

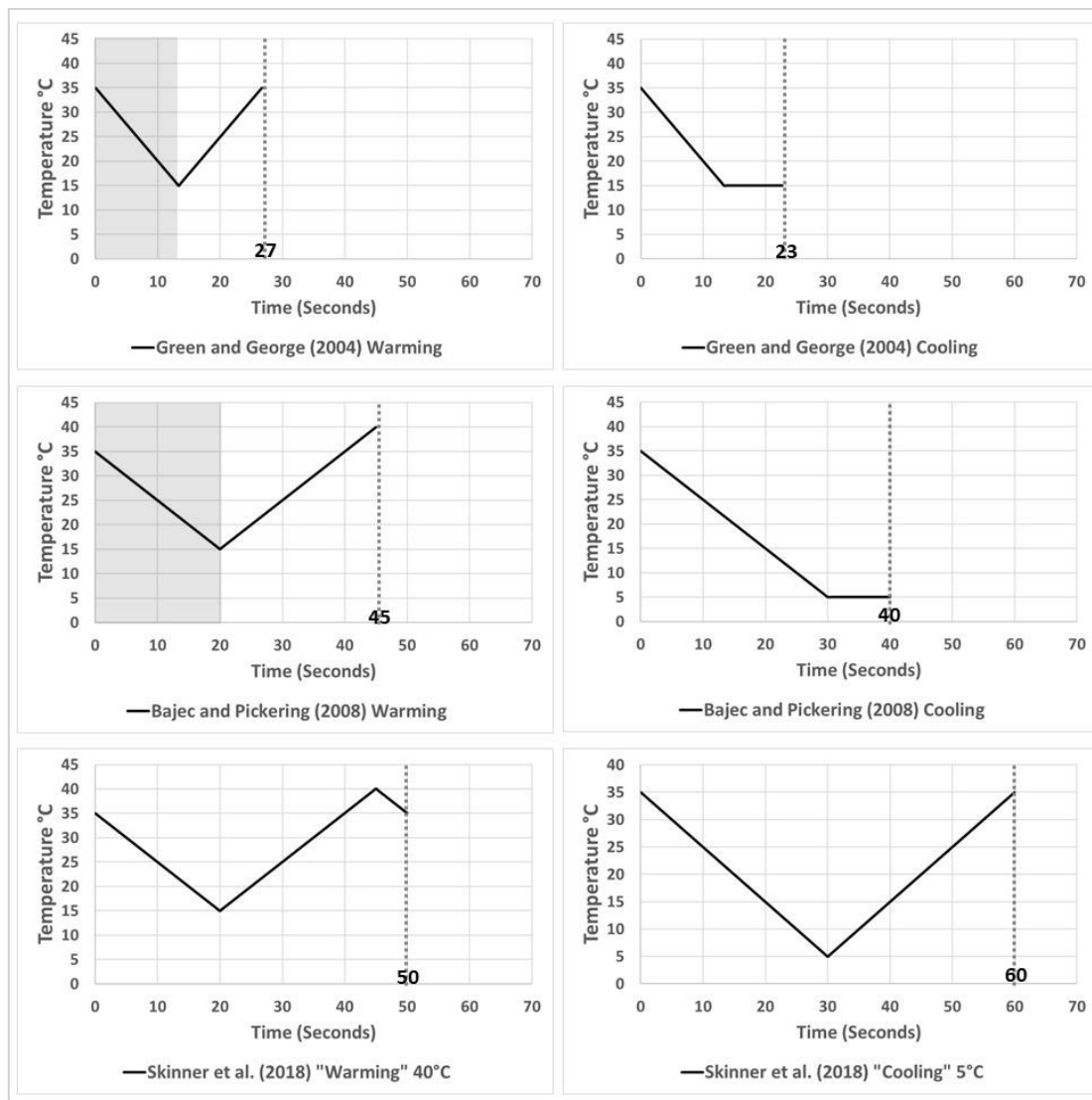


Figure 1-1 Visual representation of the temperature exposures of the Green and George (2004), Bajec and Pickering (2008) and Skinner et al. (2018) protocols. In areas greyed out, participants were instructed to attend only to the temperature change.

Cruz and Green (2000) initially reported that warming invoked sweetness and cooling evoked either sourness or saltiness, but subsequent studies have reported several different sensations during both trials (Pickering & Kvas, 2016; Yang et al., 2014), summarised in Figure 1-2.

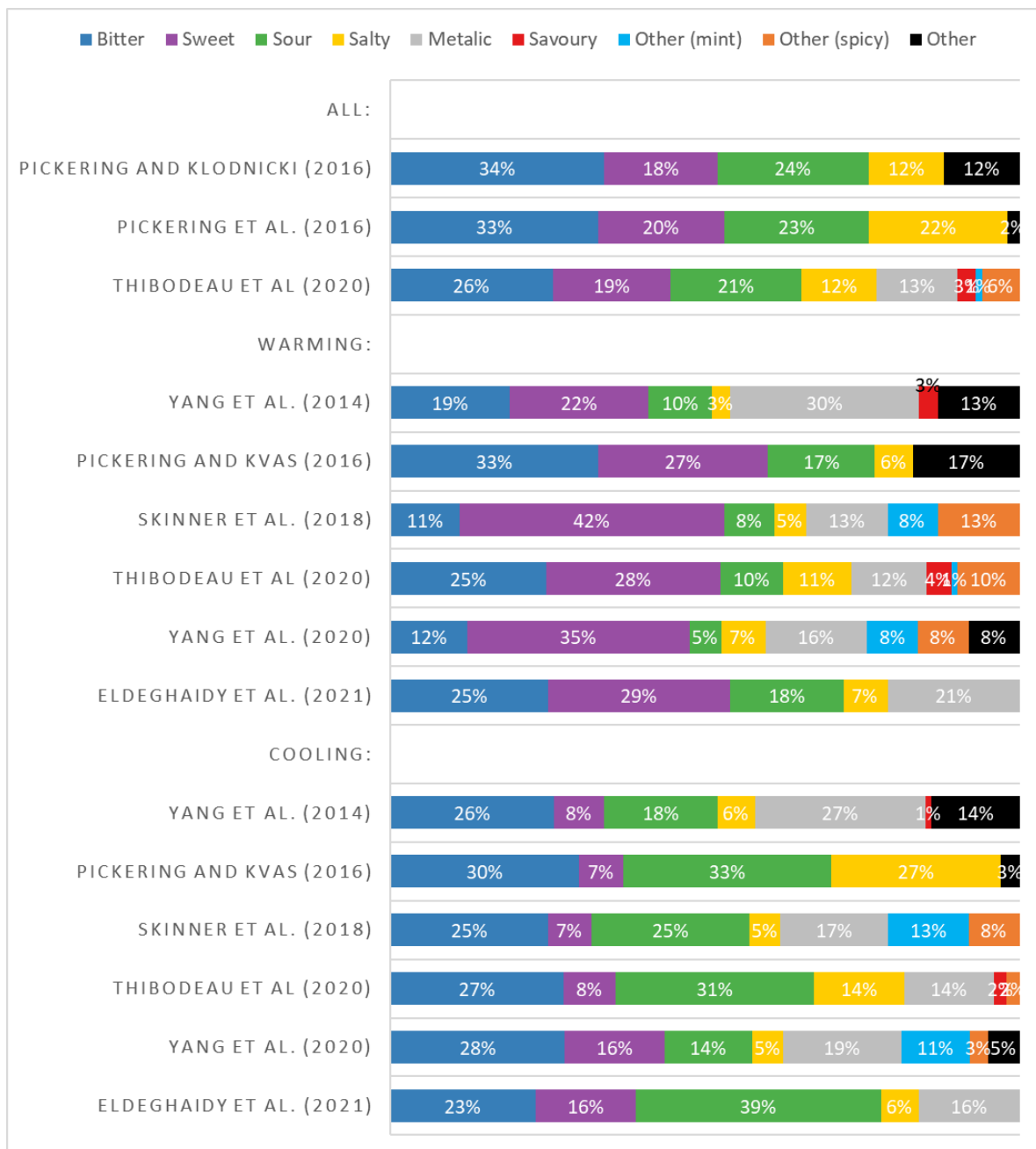


Figure 1-2 Summary of tastes reported by thermal tasters during cooling or warming of the tongue. Note that in the studies which note “all”, the authors made no distinction between the cooling and the warming trials.

In studies that pooled data from both trials (cooling and warming) (Pickering & Klodnicki, 2016; Pickering et al., 2016), a large proportion of tasters (~34%) reported bitterness, followed by sourness (~24%). A similar proportion of tasters reported sweetness (~20%), but there is some discrepancy (12% vs 20%) in the proportion of tasters who report saltiness (Pickering & Klodnicki, 2016; Pickering et al., 2016). In another study (Pickering & Kvas, 2016) that used the same approach (descriptors and

temperature trials), a larger proportion of tasters (27%) reported saltiness on the cooling trial, compared to a smaller proportion (6%) on the warming trial. This discrepancy is likely an artefact of the loss of detail from pooling the data of both trials.

In contrast, saltiness was reported at a relatively low frequency (3-6%) by the participants phenotyped by Yang et al. (2014) and Skinner et al. (2018). This low frequency could be a consequence of the fact that in both these studies, researchers supplied participants with additional options for reporting taste (savory for both studies, metallic by Yang et al., 2014), and that previously such responses may have been “dumped” under salty as more relevant options were not available. In other words, in studies where only “salty” was provided as an option, participants may have either selected the taste descriptor that fits more closely to their experience, or their perception may have been influenced by the available choices (Clark & Lawless, 1994). Thibodeau et al. (2020) presented some evidence that including training with an umami stimulus and presenting participants with the option to include “umami” as a response increases the proportion of Savoury/Umami responses.

Typically, more participants report sweet thermal taste during the warming trial than during the cooling trial. Between 22% (Yang et al., 2014) and 42% (Skinner et al., 2018) of participants report sweet thermal taste during warming, whereas only 8% (Yang et al., 2014) to 16% (Eldeghaidy et al., 2021; Yang et al., 2020) report sweet thermal taste during the cooling trial (Figure 1-2).

### **1.2.1 Groups of thermal tasters**

Previous research has tried to group TT according to their responses. Bajec et al. (2012) subdivided TT into four groups. The first two groups reported bitter thermal taste but were split by whether they reported bitterness during the warming or cooling

trial. The third and fourth groups reported sweet thermal taste on the warming and sour thermal taste on the cooling trials, respectively. This grouping is, however, an oversimplification as some tasters report tastes during both trials (Pickering & Kvas, 2016; Skinner et al., 2018), and sweetness and sourness have been reported on both cooling and warming trials (Pickering & Kvas, 2016; Yang et al., 2014). The groupings also do not take the variety of responses on both treatments, such as saltiness and metallic sensation, into account (Pickering & Kvas, 2016; Skinner et al., 2018; Yang et al., 2014).

Skinner et al. (2018) collected more detailed information on thermal taste responses using a rollerball to report the intensity of tastes perceived during a modified temperature trial. Respondents reported the identity of tastes perceived after the trial. These responses were collected continuously throughout two cycles, one starting at 35°C, cooling to 5°C and returning to 35°C (5°C trial), and the other starting at 35°C, cooling to 15°C heating to 40°C and returning to 35°C (40°C trial)(see Figure 1-1). They observed four different response patterns for both trials. Different subgroups experienced different peaks of sensation at different time points – and the associated temperature. For example, a considerable proportion (50%) of the sample reported taste during the cooling section of the 5°C trial, which peaked and persisted during the warming section. Others (21%) only reported taste at the end of the warming section. TT also differed in the reproducibility of their experience. Some participants reported a consistent thermal taste experience, but others did not report taste on all replicates or at the same temperatures. These results point to the involvement of several mechanisms in thermal taste, further suggesting that thermal taste occurs on a spectrum – highlighting the need to consider responses during phenotyping in more detail.

More recently, Thibodeau et al. (2020) grouped a large sample of thermal tasters based on the tastes they reported, as well as treatments (cooling/warming) and tongue locations inducing reported tastes, to determine whether any differences existed between groups in terms of orosensory responsiveness. In addition, they investigated relationships and overlaps between different groupings. Overall, no differences in orosensory responsiveness were found. However, Thibodeau et al. (2020) did find that Savoury/Umami TT were more responsive to sweet stimuli, although this finding was not significant after adjustment for multiple comparisons.

### **1.3 Sensory responses of Thermal Tasters**

Although the temperature-induced taste perceived and reported by TT is an interesting phenomenon, it becomes even more interesting and relevant to sensory scientists when considering how this may affect taste perception and variation in hedonic responses.

#### **1.3.1 Sensitivity differences**

Not only do TT perceive taste from temperature stimulation, they also perceive traditional tastes differently in simple matrices and some foods. In the literature, there is general agreement that TT exhibit increased sensitivity to some orosensory stimuli. This phenomenon was first described by Green and George (2004), but has been shown more extensively by several other studies, summarised in Table 1-1 and Table 1-2.

Different types of sensory methods have been used to investigate sensitivity to stimuli, including response magnitude using line scales (like thegLMS, gVAS or gLS) (Bajec & Pickering, 2008; Green & George, 2004; Pickering et al., 2016; Pickering et al., 2009; Yang et al., 2014), detection thresholds (Yang et al., 2014), difference thresholds

Table 1-1 Chronological summary of studies on TT Status and orosensory responsiveness in non-food matrices. All results reported are significant for TT Status unless otherwise specified.

Study	Method used	Matrix	Significant Findings
Green and George (2004)	Intensity data, gLMS	Water – tested sucrose, saccharin and NaCl QSO <sub>4</sub> , citric acid, PROP, MSG	• TT rated all stimuli higher
	Intensity data	Water, Sucrose, sucrose + vanillin, vanillin, citric acid	• Enhancement of sweetness of sucrose by vanillin, orthonasal + retronasal perception of vanillin (retronasal shows a larger effect).
Green et al. (2005)	Intensity data, gLMS	Applied Capsaicin and menthol	• No differences in intensity of burning, stinging, prickling
	Intensity data, gLMS	Water - Sucrose at a range of concentrations	• TT rated sweetness intensity 2x higher
Bajec and Pickering (2008)	Intensity data, gLMS	Alum, NaCl, sucrose, tartaric acid, quinine, iron sulphate	• Low astringency, high metallic significant. High astringency, bitter, sweet, and salty significant when data was log transformed.
Bajec et al. (2012)	Time-intensity, tested at 2 temperatures	Water with sucrose, citric acid, quinine, alum, MSG	• No significant difference on any parameters extracted from TI curves (TMax, I <sub>max</sub> , DUR, AUC, IAng, IArea, Dang, DArea, IDelay, IInt)
Yang et al. (2014)	Detection thresholds (ASTM)	Water + sucrose, caffeine, WS3 (cooling), capsaicin, ethyl butyrate, Isoamyl acetate	• Significantly lower threshold for sugar
	Supra-threshold gLMS	Sucrose, citric acid, NaCl, Caffeine, Capsaicin, ethyl butyrate at 2 concentrations	• Approached significance for salty and caffeine. When modalities are grouped together, TT perceived tastes and trigeminal stimuli more intensely, but not aroma.
Thibodeau et al. (2019)	gLMS	Aqueous solutions used for descriptor training for sweet (sucrose), Salty (NaCl – not full cohort), Sour (citric or tartaric acid), Bitter (Quinine), Umami (MSG), Metallic (Cupric sulphate or Iron sulphate – not full cohort) and astringent (Aluminium sulphate – not full cohort)	• TT rated all tastes except Umami significantly higher. • Umami responsiveness higher in TT than TnT but not significant <sup>1</sup>
Yang et al. (2020)	gLMS	Aqueous solutions representing prototypical tastes: (sucrose), Salty (NaCl), Sour (citric acid), Bitter (Quinine), Umami (MSG), Metallic (Ferrous sulphate)	• Caucasian participants: no differences between TT and TnT. • Asian participants: TT scored metallic intensity higher than NC. • TT Status predictive of metallic intensity, but less important than ethnicity and PROP taster status.
Eldeghaidy et al. (2021)	Sweetness intensity ranked on four-point category scale	Sucrose solutions delivered to participants at cold (5± 2 °C) and ambient (20 ± 2 °C) temperatures.	• Trend of increased sweetness perception in TT for cold sweet sample across all trials. • Cortical responses to Cold sweet samples increased in TT

<sup>1</sup> Note that in this study, not all participants were tested for umami responsiveness, and the lack of significance may be an artefact

Table 1.1 (Cont.) Chronological summary of studies on TT Status and orosensory responsiveness in non-food matrices. All results reported are significant for TT Status unless otherwise specified.

Study	Method used	Matrix	Significant Findings
Thibodeau and Pickering (2021)	Ranking task (three concentrations)	Aqueous solutions of fructose, aluminium sulphate, tartaric acid, and quinine	<ul style="list-style-type: none"> <li>• No performance difference between TT and TnT</li> </ul>
	Identification task	Aqueous solutions of fructose, aluminium sulphate, tartaric acid, and quinine	<ul style="list-style-type: none"> <li>• No performance difference between TT and TnT</li> </ul>
	Multiple gLMS (sweet, sour, bitter, astringency, burning/tingling)	Ethanol, fructose, aluminium sulphate, tartaric acid, and quinine. (three concentrations)	<ul style="list-style-type: none"> <li>• Ethanol: TT more responsive to sweetness and astringency, and more response (not significant) to bitterness and burning/tingling.</li> <li>• Fructose: TT were more responsive to sweetness</li> <li>• Aluminium sulphate: TT more responsive to sourness. TT more discriminant to sourness intensity (low vs high concentration)</li> </ul>
	Multiple gLMS (sweet, sour, bitter, astringency, burning/tingling)	Binary mixtures: Ethanol in combination with fructose, aluminium sulphate, tartaric acid, and quinine. (three concentrations)	<ul style="list-style-type: none"> <li>• TT more responsive to increase in astringency in ethanol + fructose between 5-13% ethanol.</li> <li>• TT more responsive to burning/tingling of 23% ethanol.</li> <li>• TT had higher mean orosensory ratings than TnT for binary mixtures of ethanol and fructose (sweetness, astringency, burning/tingling), quinine (sweet), tartaric acid (bitter, sour) and aluminium sulphate (sweet, astringent)</li> </ul>
	Analysis of number of scales used	Ethanol, fructose, aluminium sulphate, tartaric acid, and quinine. (three concentrations)	<ul style="list-style-type: none"> <li>• TT used significantly more scales than TnT to describe water, 5% ethanol, 13% ethanol, 140 mM fructose, 2.05 mM aluminium sulphate and 5.43 mM aluminium sulphate.</li> </ul>
	Analysis of number of scales used	Binary mixtures: Ethanol in combination with fructose, aluminium sulphate, tartaric acid, and quinine. (three concentrations)	<ul style="list-style-type: none"> <li>• TT used mores scales to describe some binary solutions; 5% ethanol/960 mM fructose, 13% ethanol/0.73 mM aluminium sulphate, 5% ethanol/5.43 mM aluminium sulphate and 23% ethanol/6.91 mM tartaric acid.</li> </ul>

Table 1-2 Chronological summary of studies on TT Status and orosensory responsiveness in food and beverage matrices. All results reported are significant for TT status unless otherwise specified.

Study	Method used	Matrix	Significant Findings
Pickering et al. (2009)	Intensity data, gVAS	Wine (red and white)	Higher ratings of sweetness and sourness in white wine, and sourness, astringency, and bitterness in red wines
Pickering et al. (2010)	Intensity data gLS	Beer (several styles)	Higher rating of bitterness in lager, higher rating of sourness in NAB, trend to rate all attributes higher
Yang (2015)	gLMS	Strawberry-flavoured drink at frozen, cold, ambient, and warm temperatures	TT found a significantly greater difference in temperature between cold (5°C)/ambient (20°C) and frozen (-4°C)/ambient (20°C) than TnT.
Hort et al. (2016)	Discrimination (% correctly identified samples) during concurrent fMRI scanning	Levels of CO <sub>2</sub>	Significantly better discrimination of CO <sub>2</sub> level at high CO <sub>2</sub> levels
Pickering and Klodnicki (2016)	Profiling on gVAS	Various food products	Rated “crispiness” higher in crisps/chips and crispbread. (Results approach significance)
Pickering and Kvas (2016)	Difference thresholds (2AFC)	Wine – tested sucrose, quinine + tartaric acid	Trend of TT having lower difference threshold, only significant for tartaric acid.
Pickering et al. (2016)	gVAS	Various food products	More intense tastes in “bitter” orosensory group
Small-Kelly (2018)	gLMS	Beer and cider spiked for range of EtOH and iso-alpha acid concentrations (hop bitterness)	TT rated sourness (beer and cider), overall taste intensity (beer and cider), bitterness (beer) and astringency (cider) higher. TT more responsive to bitterness of ethanol.
Mitchell et al. (2019)	TCATA	Dealcoholised beer	Increased perception of astringent, bitter, sour and malty, temperature and sweet interaction significant.

(Pickering & Kvas, 2016) and temporal methods like time-intensity measures (Bajec et al., 2012) and temporal check-all-that-apply (TCATA) (Mitchell et al., 2019)(Table 1-1 and Table 1-2). Several studies demonstrated an increased sensitivity to basic tastes (Bajec & Pickering, 2008; Green et al., 2005; Green & George, 2004; Thibodeau & Pickering, 2021; Thibodeau et al., 2019; Yang et al., 2014), and astringent stimuli, (Bajec & Pickering, 2008; Thibodeau & Pickering, 2021; Thibodeau et al., 2019). There is also some evidence that TT experience orosensations in foods and beverages more intensely (Mitchell et al., 2019; Pickering et al., 2010; Pickering et al., 2016; Pickering et al., 2009; Small-Kelly, 2018) and experience more complex sensations when different stimuli are combined (Thibodeau & Pickering, 2021). TT are also more descriptive of orosensations, as they use, on average, more scales (descriptors) to describe simple and complex combined orosensations than TnT (Thibodeau & Pickering, 2021).

One less studied area where there has been disagreement is the sensitivity of TT to odourants. Green and George (2004) reported a greater sensitivity in TT for vanillin and vanillin-sucrose mixtures. This effect was more pronounced for retronasal stimuli than for olfactory stimuli. On the other hand, Yang et al. (2014) investigated different odorants – namely isoamyl acetate and ethyl butyrate – and found no response difference between TT and TnT. The difference observed by Green and George (2004) may be because vanillin activates the temperature-sensitive TRPV1, TRPV3 and TRPA1 taste channels (Lubbert et al., 2013). The contradictory findings between these two studies (Green & George, 2004; Yang et al., 2014) point to differences in perception mechanisms between the two classes of odourants. Temperature-sensitive channels (TRPM5 and TRPV1) have been potentially linked as a mechanism behind thermal taste (Talavera et al., 2007; Talavera et al., 2005), and the selective advantage

for vanillin could be related.

There have also been contradictory reports regarding trigeminal stimuli. Green et al. (2005) found no significant differences in the intensity ratings of TT and TnT for menthol and capsaicin, but Yang et al. (2014) subsequently found an overall significant difference in intensity rating when pooling the intensity ratings of two concentrations of capsaicin. Additionally, TT have demonstrated better discrimination ability of CO<sub>2</sub> level than TnT (Hort et al., 2016), and have rated other orosensations (metallic and astringency) higher (Bajec & Pickering, 2008; Pickering et al., 2009; Small-Kelly & Pickering, 2019). Many of these stimuli have been shown to activate temperature-sensitive taste channels. Menthol activates TRPM8, capsaicin TRPV1 (Latorre et al., 2009), divalent salts which induce metallic taste activate TRPV1 and TRPM5 (Riera et al., 2009), and TRPA 1 has been linked to the nociceptive response to CO<sub>2</sub> (Wang et al., 2010).

Although earlier studies looking into the temperature sensitivity of TT had mixed results (Bajec & Pickering, 2008; Green & George, 2004; Yang et al., 2014), in the meta-study by Thibodeau et al. (2019), TT showed significantly higher temperature responses on both the tongue, lip and hands. TT have also reported higher intensity differences between beverages at different temperatures (cold vs ambient and frozen vs ambient) (Yang, 2015). When completing a TCATA task, there was a significant interaction between TT Status, temperature and sweetness response (Mitchell et al., 2019). This effect, however, was not demonstrated with time-intensity measures at different temperatures (Bajec et al., 2012), as no response differences in any of the measures of time-intensity curves were found between TT and TnT. It is also important to note that the differences in temperature sensitivity have been found in real beverages (beer and a strawberry-flavoured beverage) (Mitchell et al., 2019; Yang,

2015), instead of model solutions (Bajec et al., 2012), or applied temperatures (Bajec & Pickering, 2008; Green & George, 2004; Yang et al., 2014). These results point to the significance of temperature as a potential driver of food choice for TT.

### **1.3.2 Impact on food choice**

As TT experience oral sensations differently from TnT, TT could also be expected to have different hedonic experiences from TnT, resulting in different food and beverage choices. Table 1-3 summarises studies investigating the effect of TT Status on hedonic responses and food choices. Most of these studies were limited by their relatively small sample sizes of less than 50 (Hort et al., 2016; Pickering et al., 2010; Pickering & Klodnicki, 2016; Pickering et al., 2016; Pickering et al., 2009; Yang, 2015) with the largest reported study still involving only 76 participants (Bajec & Pickering, 2008) (Effect sizes in terms of liking that are typically considered consumer-relevant means that it is recommended to test with between 75-150 consumers (Lawless & Heymann, 2010).) Non-standard methodological approaches for testing food preference have also limited other studies, as some included an affective task with a descriptive task (Pickering et al., 2010; Pickering et al., 2009), others tested only recalled liking of foodstuffs (Bajec & Pickering, 2008), and serving products at a different temperature than normal consumption (Pickering et al., 2010).

Studies using alcoholic beverages have not found significant differences in liking between TT and TnT (Pickering et al., 2010; Pickering et al., 2009; Q. Yang et al., 2018). However, two of these studies were limited by the completion of a descriptive and an affective task in the same session (Pickering et al., 2010; Pickering et al., 2009), shown to reduce discrimination (Popper et al., 2004). It has, however, been suggested that TT are more discriminant in liking for CO<sub>2</sub> levels (Hort et al., 2016) with TT exhibiting a clear preference direction towards lower CO<sub>2</sub> vs high. In their study, TnT did not favour

Table 1-3. Chronological summary of studies exploring TT status and food choice. All findings are significant for TT Status unless stated otherwise.

Reference	Pickering et. al, (2010)	Bajec and Pickering (2010)	Pickering et al. (2010)	Yang (2015)	Pickering and Klodnicki (2016)	Hort et al. (2016)	Pickering et al. (2016)	Yang et al. (2018)
<b>Method</b>	<ul style="list-style-type: none"> <li>Liking on 7-point hedonic scale.</li> <li>Combined in session with orosensations/ profiling</li> </ul>	<ul style="list-style-type: none"> <li>Reported liking of 332 food items on 7-point Likert scale</li> </ul>	<ul style="list-style-type: none"> <li>Liking on 9-point hedonic scale.</li> <li>Combined in session with orosensations/ profiling</li> </ul>	<ul style="list-style-type: none"> <li>Overall liking on LAM scale.</li> <li>Combined with intensity scaling.</li> <li>LAM done in separate session and always before intensity</li> </ul>	<ul style="list-style-type: none"> <li>Liking on 9-point hedonic scale.</li> <li>Combined in study with orosensations/ profiling</li> </ul>	<ul style="list-style-type: none"> <li>Preference sorting (3 samples)</li> </ul>	<ul style="list-style-type: none"> <li>Liking on gDOL scale.</li> <li>Combined in study with orosensations/ profiling.</li> </ul>	<ul style="list-style-type: none"> <li>Overall liking on LAM scale</li> <li>Emotional responses</li> </ul>
<b>Samples</b>	<ul style="list-style-type: none"> <li>4 wine styles (2 red Valpolicella &amp; Cabernet Sauvignon 2 white: Pinot Grigio &amp; Gewurztraminer)</li> </ul>	<ul style="list-style-type: none"> <li>No samples tasted. 332 food items listed.</li> <li>Grouped into food groups, orosensory groups, correlation groups</li> </ul>	<ul style="list-style-type: none"> <li>Different beer styles (Lager, NAB, Weiss, Brown ale, Stout).</li> <li>All served at room temperature and expectorated</li> </ul>	<ul style="list-style-type: none"> <li>Strawberry flavoured drink (SFD) served frozen, cold, ambient, and warm</li> </ul>	<ul style="list-style-type: none"> <li>22 food and beverage items grouped according to dominant orosensation</li> </ul>	<ul style="list-style-type: none"> <li>Carbonated model beverage at 3 different levels of carbonation</li> </ul>	<ul style="list-style-type: none"> <li>20 everyday foods that encompass a range of orosensory sensations.</li> <li>Grouped into food category groups and orosensory groups (groups which share predominant orosensory quality)</li> </ul>	<ul style="list-style-type: none"> <li>Beer presented at different carbonation and temperature levels</li> </ul>
<b>Sample size (TT/TnT)</b>	<ul style="list-style-type: none"> <li>40</li> </ul>	<ul style="list-style-type: none"> <li>76 (26/50)</li> </ul>	<ul style="list-style-type: none"> <li>40 (20/20)</li> </ul>	<ul style="list-style-type: none"> <li>47 (22/25)</li> </ul>	<ul style="list-style-type: none"> <li>44 (23/21)</li> </ul>	<ul style="list-style-type: none"> <li>24 (12/40)</li> </ul>	<ul style="list-style-type: none"> <li>44 (25/19)</li> </ul>	<ul style="list-style-type: none"> <li>60 (30/30)</li> </ul>
<b>Significant findings</b>	<ul style="list-style-type: none"> <li>No difference between TT and TnT</li> </ul>	<ul style="list-style-type: none"> <li>TT gave lower liking scores for: Cooked fruit food group, Cooked fruit/Vegetable &amp; bitter correlation group.</li> <li>Approach significance on Fruit, seafood, Pasta, Mushy orosensory group and Raw fruit correlation group.</li> </ul>	<ul style="list-style-type: none"> <li>No difference between TT and TnT</li> </ul>	<ul style="list-style-type: none"> <li>TT showed significantly lower liking of SFD at frozen, cold, and warm temperatures. Ambient is the same.</li> </ul>	<ul style="list-style-type: none"> <li>TT gave lower liking for “creamy orosensory groups (dairy products).</li> <li>TT gave lower liking (approach significance) for “aversive” food group (asparagus, broccoli, dark chocolate, walnut, pomegranate, cranberry juice)</li> </ul>	<ul style="list-style-type: none"> <li>TT prefer lower carbonation, no preference difference in TnT</li> </ul>	<ul style="list-style-type: none"> <li>Applesauce (making up the grainy orosensory group) liked more by TT than TnT</li> </ul>	<ul style="list-style-type: none"> <li>No differences in overall liking between TT and TnT</li> <li>TT felt more tame/safe, curious, underwhelmed, shocked, bored, disconfirmed, and disgusted</li> </ul>

carbonation levels in either direction. TT also exhibited significantly lower liking ratings than TnT at extreme temperatures in strawberry-flavoured beverages (Yang, 2015). In light of these two findings, the lack of significant differences in liking by Q. Yang et al. (2018) is somewhat surprising, as the study investigated the interaction between serving temperature and carbonation on liking. However, the study did find that TT rated six out of 10 emotions higher than TnT. This discrepancy in results may result from emotional measures being a more discriminating tool for measuring food choice than traditional hedonic ratings (Ng et al., 2013).

In contrast to results for alcoholic beverages, there has been some evidence of differences between TT and TnT when making liking judgments about complex food matrices. In a study investigating recalled liking of 322 food items, TT reported significantly lower remembered likings for cooked fruit and a group of bitter foodstuffs (Bajec & Pickering, 2010). Similarly Pickering et al. (2016) found that TT had a significantly lower liking for applesauce than TnT. Pickering and Klodnicki (2016) also reported TT to exhibit a lower liking for the “creamy” orosensory group comprised of dairy products.

Studies investigating TT Status and Neophobia have yielded mixed results. Bajec and Pickering (2010) reported marginally higher (but non-significant:  $p = 0.25$ ,  $n=76$ ) food neophobia scores in TT, whereas Yang (2015) found a tendency of lower food neophobia ( $p=0.2$ ,  $n=116$ ). The same study also reported that TT are significantly more likely to sample new and different foods and score higher in food involvement. In this study, TT also scored significantly higher in the Openness dimension of the Five-factor personality model, linked to lower Neophobia in other studies (Knaapila et al., 2011; Nezlek & Forestell, 2019). The contrast in the desire to try new foods and overall Neophobia could mean that TT scored higher in the “approach tendency” factor of food

Neophobia but not necessarily the “avoidance tendency” (Nezlek & Forestell, 2019). This may be linked to the fact that the differences in liking have been found based on texture (Bajec & Pickering, 2010; Pickering & Klodnicki, 2016), and aversion to a range of textures has been linked to overall food neophobia (Nederkoorn et al., 2019).

In summary, despite research into the topic being limited by low sample sizes and low-discriminant methodology, there is some indication that TT exhibit different food choice behaviour than TnT. However, these differences seem to be more related to temperature (Yang, 2015), trigeminal sensations (Hort et al., 2016) and texture (Bajec & Pickering, 2010; Pickering & Klodnicki, 2016). More discriminant methods have also picked up differences (Q. Yang et al., 2018), suggesting that there may not be large differences in hedonic experiences of TT, but rather minor differences related to the holistic consumption experience.

#### **1.4 Thermal taste phenotyping**

As highlighted in section 1.2, TT Status is phenotyped by applying a changing temperature stimulus to the tongue, typically using a temperature-changing Peltier device, changing temperature at a specific rate. Typically, participants report taste responses after temperature exposure. Generally, individuals who perceive taste induced through changing temperature are considered TT, and those who do not perceive taste are considered TnT (Green & George, 2004). However, different studies have applied different, specific classification methods. In some classification methods, it is possible to be neither classified as a TT nor a TnT (Bajec & Pickering, 2008). This third, unclassified group (UC/NC) can sometimes constitute up to 50% of the sample (Thibodeau et al., 2019).

#### 1.4.1 Phenotyping methodologies

Cruz and Green (2000) initially used a protocol that only involved exposing the tongue to heating or cooling but was refined for future work (Green et al., 2005; Green & George, 2004). The updated trial applied a temperature-changing Peltier device to three locations on the tongue: the tongue tip and 10mm to both the left and right-hand side from the midline. As summarised in Table 1-4 and previously shown in Figure 1-1, the updated warming trial started at 35°C, cooled to 15°C and then heated to 35°C, with participants instructed to only attend to the temperature during the warming phase (Table 1-4 and Figure 1-1). The cooling trial also started at 35°C, cooled to 15°C and then kept the tongue at this temperature for 10 seconds. All cooling and warming happened at a constant rate of approximately 1.5°C/second. After each trial, participants rated the intensity of sweet, sour, bitter or salty taste that they perceived using a quasi-logarithmic intensity scale (gLMS, as shown in Figure 1-3). If no taste was perceived, participants rated zero or “no sensation” for all four tastes. If participants reported a taste more intense than “weak” on the gLMS, the phenotyping was repeated to confirm the response. All other participants were classified as TnT (Table 1-5). Warming trials always preceded cooling trials, as keeping the tongue at a low temperature induced a numbing effect.

Bajec and Pickering (2008) adapted the protocol by expanding the temperature range of both trials. The warming trial was extended from 35°C to 40°C, and the minimum temperature of the cooling trial was changed from 15°C to 5°C (Figure 1-1, section 1.2). In addition, a second replicate was added for all participants, not only those who initially reported taste, and the rate of temperature change changed from 1.5°C/s to 1°C/s. Those who reported none of the accepted tastes (sweet, bitter, sour, or salty) on either of the 12 trials were considered TnT, as were any participants who only reported

Table 1-4 Summary of methodologies, devices and strategies used for different Thermal Taste classification methods. Adapted from Thibodeau et al. (2019).

Source first reported	Subsequent articles using the same scheme	Location	Methodological differences			
			Temperatures		Reps	Probe Size
			Warming	Cooling		
Green and George (2004)	Green et al. (2005)	3: tip, 1-cm to left/right of midline	15°C to 35°C	35°C to 15°C	2: 1 full, 1 as needed	64mm <sup>2</sup> (Shape not reported)
Bajec and Pickering (2008)	All papers in research group <sup>1</sup>	3: tip, 1-cm to left/right of midline	15°C to 40°C	35°C to 5°C	2	6 mm <sup>2</sup> (Square)
Yang et al. (2014)	None	1: tip	15°C to 40°C	35°C to 5°C	2	28.26mm <sup>2</sup> (Truncated Cone)
Hort et al. (2016)	Q. Yang et al. (2018) <sup>2</sup>	1: tip	15°C to 40°C	35°C to 5°C	2	28.26mm <sup>2</sup> (Truncated Cone)
Skinner et al. (2018)	None	1: tip	15°C to 40°C	35°C to 5°C	3: 2 full, 1 as needed	256mm <sup>2</sup> (Square)

<sup>1</sup> Bajec and Pickering (2010); Bajec et al. (2012); Bering et al. (2013); Mitchell et al. (2019); Pickering et al. (2010); Pickering and Klodnicki (2016); Pickering and Kvas (2016); Pickering et al. (2016); Pickering et al. (2009)

<sup>2</sup>Participants recruited by Yang et al. (2018) were tested using the same methodology as Hort et al. (2016) except that a larger probe was used (256 mm<sup>2</sup> square)

Table 1-5 Summary of classification strategies and percentages reported for different Thermal Taste classification methods. Adapted from Thibodeau et al. (2019).

Source first reported	Other work using scheme	Valid thermal tastes	Phenotypes:		
			TT	TnT	UC
Green and George (2004)	Green et al. (2005) <sup>1</sup>	Sweet, sour, salty and bitter.	Definition: Reports a taste sensation above 'weak' on first rep that can be confirmed when the same location and temperature regime is re-tested. Reported: 46–54%	Any participants not classified as a TT.	Not included in this scheme.
Bajec and Pickering (2008)	All papers in research group <sup>2</sup>	Sweet, sour, salty, bitter, Umami (savory) and metallic <sup>3</sup> .	Definition: Reports the same taste sensation, at the same location, during the same temperature cycle above "weak". Reported: 20–38%	Reports no taste sensations (including "other") during all trials <sup>4</sup> 24–40%	Any participants not classified as a TT or TnT. 24–50%
Yang et al. (2014)	None	Sweet, sour, salty, bitter, savoury, metallic and other.	Definition: Reports any taste sensations above 'weak' during all trials. Taste sensations do not need to be the same across trials. Reported: 27%	Reports no taste sensations during all trials. 30%	Any participants not classified as a TT or TnT. 43%
Hort et al. (2016)	Q. Yang et al. (2018)	Sweet, sour, salty, bitter, Umami, other (minty).	Definition: Reports any taste sensations above "weak" both warming and/or both cooling trials. Taste sensations do not need to be the same across trials. Reported: 23%	Reports no taste sensations during all trials. 77%	None reported under this scheme. Not reported
Skinner et al. (2018)	None	Sweet, sour, salty, bitter, Umami and other (spicy, metallic, minty).	Definition: Reports the same taste sensation in two or more replicates of the same warming and/or cooling trial regardless of the intensity. Reported: 28%	Reports no taste sensations during all trials. 51%	Any participants not classified as a TT or TnT. 21%

<sup>1</sup>Participants recruited by Green et al. (2005) were tested using the same methodology as Green and George (2004) but only participants reporting sweet thermal taste during warming were retained.

<sup>2</sup>See Table 1-4

<sup>3</sup>Metallic and Umami were only included in more recent works: Mitchell et al. (2019); Small-Kelly and Pickering (2019); Thibodeau et al. (2019)

<sup>4</sup>Participants were classified as TnT if they reported experiencing another taste sensation that when described was heat related (e.g., spicy, hot peppers), cold related (e.g., minty, menthol), or a mouthfeel (e.g., drying, tingling)

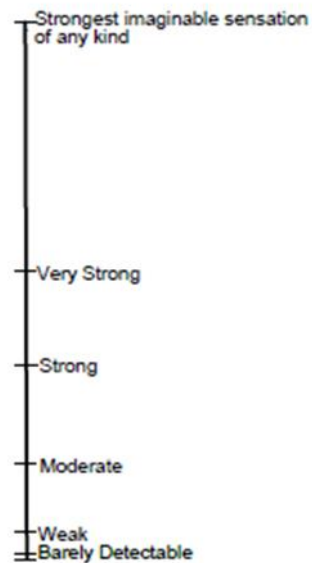


Figure 1-3 An example of the General Linear Magnitude Scale (gLMS). Participants are asked to first rank a number of familiar sensations on the scale (for example the brightness of the sun) prior to ranking tastes. As the markings on the scale are quasi-logarithmic data is often log transformed. The labels on the scale are their relative positions are: barely detectable, 1.4; weak, 6.1; moderate, 17.2; strong, 35.4; very strong, 53.3; strongest imaginable, 100. (Green et al., 1996)

“other” taste sensations that were temperature related. TT were all participants who reported the same taste above weak on the gLMS, at the same tongue location during the same temperature trial. All participants who reported taste inconsistently, at low intensity or a combination of both were considered unclassified (UC/NC). The minimum intensity and repeatability of thermal taste required was included to control for response bias (or demand bias) .This protocol has been adopted widely by both that research group and others (Bajec & Pickering, 2010; Bajec et al., 2012; Bering et al., 2013; Pickering et al., 2010; Pickering & Klodnicki, 2016; Pickering & Kvas, 2016; Pickering et al., 2016; Pickering et al., 2009) but the list of accepted thermal tastes has been expanded to include metallic and Savoury/Umami as valid tastes in recent works (Mitchell et al., 2019; Small-Kelly & Pickering, 2019; Thibodeau et al., 2019).

Clark (2011) (published as Hort et al., 2016) only tested one tongue location and expanded the list of accepted tastes to include “minty”, and removed the requirement

for the reported taste(s) to be the same across both trials. Yang et al. (2014) adapted this to include “other” as a valid thermally induced taste but also required TT to report taste on both trials (warming and cooling). Subsequently, Skinner et al. (2018) argued that the requirement for TT to report tastes above “weak” is arbitrary and removed this requirement to classify participants as TT. They also introduced a third replicate of the temperature trials, administered if participants reported taste inconsistently during the first two replicates. This adaptation was also driven by the argument that some participants were unclassified because of behavioural factors rather than factors related to TT Status (Skinner et al., 2018).

In their work aiming to describe thermally induced tastes, Skinner et al. (2018) adopted a continuous protocol and introduced the idea that a single temperature trial spanning from 40°C to 5°C with TCATA could be used as an alternative phenotyping methodology.

Although the definition of TnT has stayed constant, each subsequent phenotyping methodology and its corresponding classification scheme has resulted in a slightly changed definition of thermal taste, with implications for the portions of the population that are unclassified in specific schemes.

#### **1.4.2 The “unclassifieds”**

A characteristic that all the above-mentioned classification methods share is that TT Status classification is conditional beyond simply reporting thermal taste. TnT are consistently considered to be those who do not report taste, meaning that the traditional definitions of TT and TnT are non-overlapping and inevitably result in an unclassified group. This group varies between 21% (Skinner et al., 2018) and 42% (Yang et al., 2014) of individuals, with many studies reporting this number around

30% (Pickering & Kvas, 2016; Pickering et al., 2016), and other studies not reporting this number at all (for example Green & George, 2004; Mitchell et al., 2019; Pickering & Klodnicki, 2016).

This inconsistency in unclassified individuals is likely a result of the different classification methods used in different studies and the different implied definitions of TT Status. TT classification is defined along four dimensions: the temperature range in which the tastes occur, the sensations included as tastes, the intensity at which the tastes occur, and the repeatability of participant responses. As the understanding of thermal and chemosensory taste has developed, the definition of TT Status matured concurrently. Figure 1-4 shows a selection of definitions of TT Status by visualising differences in temperature exposures and the requirement of repeatable thermal taste reporting.

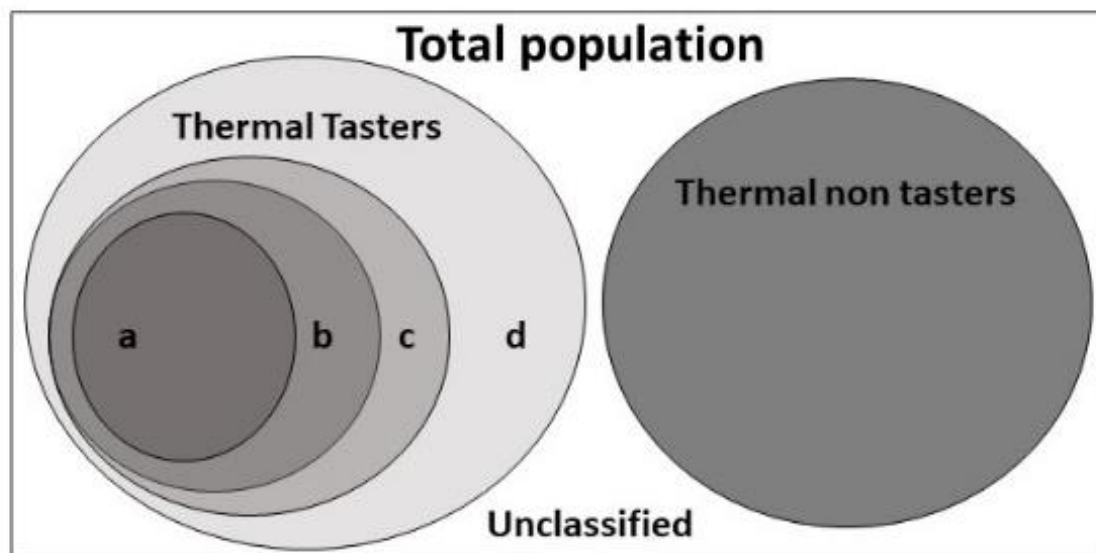


Figure 1-4. The expanding definition of Thermal Taste a) Green and George (2004) Perceive tastes between 15°C and 35°C plus reports taste consistently b) Bajec and Pickering (2008) Perceive tastes between 5°C and 40°C plus has no reports taste consistently c) Skinner et al. (2018) Perceive tastes between 5°C and 40°C but may be inconsistent in reporting taste d) Potentially expanded definition of Thermal Taste.

The first aspect of the definition of TT Status to expand was the temperature range. Green and George (2004) considered TT to be those who perceive thermally induced

tastes between 35°C and 15°C, whereas Bajec and Pickering (2008) expanded this definition to include those who perceive thermal tastes between 5°C and 40°C. Valid tastes for inclusion have also expanded from only including sweet, sour, bitter and salty (Green & George, 2004) to the inclusion of Savoury (Umami), Metallic and “Other” (Minty and Spicy) (Yang et al., 2014). Similarly, Bajec and Pickering (2008) considered TT to be those that repeatedly report the same taste over different replicates, whereas others only required TT to report any taste over two replicates (Hort et al., 2016). The minimum required intensity for reporting tastes has also been lifted by Skinner et al. (2018) to include individuals who reported tastes below “weak” as thermal tasters. They also gave participants a third chance to report taste if they were inconsistent in their initial responses. These schemes have adapted to attempt to include thermal tasters in studies who may have previously been excluded based on criteria unrelated to thermal taste and an acceptance that thermal taste may involve different phenomena (Skinner et al., 2018).

Existing classification schemes have likely classified participants as unclassified based on response and behavioural factors unrelated to thermal taste as a phenomenon (Skinner et al., 2018). First, participants may have trouble accurately and consistently naming tastes despite familiarisation with these tastes. Second, they may perceive taste weakly or have a conservative response strategy when using the gLMS or the gVAS. Third, they may be unsure of what they experienced on the first trial and not report taste. Fourth, they may experience something that is not a taste (like a minty sensation or chemesthesis) and either do not report it or may not have been classified by a specific scheme, depending on whether the scheme considered the orosensation they experienced as thermal taste. Fifth, despite not perceiving taste, they falsely reported taste on the first replicate due to other behavioural factors such as response

bias or wanting to please the experimenter. Depending on the classification scheme, these experiences would have different classification outcomes (Table 1.6). It is therefore vital to consider this when designing future classification methods, especially given the diversity of response within TT that has been shown (Skinner et al., 2018).

*Table 1-6. Hypothetical experiences and situations that may cause individuals to be unclassified under some classification schemes, as well as their potential classifications under classification schemes proposed by Bajec and Pickering (2008), Skinner et al. (2018) and Thibodeau et al. (2019).*

Experience	Response	Classification according:		
		Bajec and Pickering (2008)	Skinner et al. (2018)	Thibodeau et al. (2019)
Has trouble naming tastes	Names tastes inconsistently between repetitions	UC/NC	TT	NC-NoREP
Are conservative in scale use	Reports same (or different) tastes below “weak”	UC/NC	TT	NC-BWO
Unsure whether they experience taste on the first replicate	Responds with “No” followed with “Yes” (followed with Yes)	UC/NC	TT	NC-NoREP
Taste something which is not a “traditional” taste (like metallic or minty)	Responds with “other”, or “minty”	TnT	TT	TnT
	Names tastes inconsistently between repetitions	UC/NC	TT	NC-NoREP
	“Yes” followed by “No” or “No” followed by “Yes”	UC/NC	UC/NC	NC-NoREP
Falsely responds “Yes” on the first trial in spite of not truly perceiving taste	Responds “No” for both replicates	TnT	TnT	TnT
	“Yes” followed by “No” (followed by “No”)	UC/NC	TnT	NC-NoREP

Key: NC = Not classified, NC-NoREP = Not classified, because of unrepeatability, NC-BWO = Not Classified Below Weak Only,

Thibodeau et al. (2019) further sub-classified their group of unclassified tasters into four groups based on the reasons for their failure to be classified. They based this grouping on whether participants responded with a rating above weak for at least one taste (NC-AW) or with only ratings below weak (NC-BWO). The total group of NC tasters were then further sub-classified based on the reproducibility of the response, with individuals who reported the same taste at the same location on both replicates placed in one group (NC-REP) and those who did not show this consistency in another group (NC-NoREP). These groups’ intersections represent 9-12% of the total population (AW/REP = 12%, AW/NoREP = 12%, BWO/NoREP = 10% and BWO/REP = 9%). Although it is impossible to directly reclassify these subgroups using the classification criteria in Skinner et al. (2018), Skinner et al. (2018) would have

classified the BWO/REP group (representing 9%) as TT. In Skinner et al. (2018)'s data, the equivalent of the BWO/REP group constituted 5% of participants, in the same order of magnitude as that reported in Thibodeau et al. (2019).

When comparing the orosensory responsiveness of the different subgroups to those of TT and TnT, Thibodeau et al. (2019) found similarities between the NC-AW group and TT in their response to tastes (significant for sweet, sour and bitter), indicating that this group may constitute unclassified TT. Further, the NC-BWO group also generally responded with lower ratings on the gLMS on temperature, which in some cases were lower than those of TnT, pointing to conservative scale use, confirmed by comparing ratings for “brightness of the sun” for this group. This further underlines that a threshold for taste response is a critical factor in potentially misclassifying TT and helps to justify its removal as a requirement. Repeating the same analysis using the REP and NoREP groups showed less clear trends, indicating a more complex situation. Introducing additional replicates (Skinner et al., 2018), as well as removing the requirement to report the same taste on both replicates (Hort et al., 2016; Yang, 2015) aids in classifying TT – but also accurately classifies TnT who incorrectly reported taste on the first replicate. This strategy improves the overall efficacy of the phenotyping process.

There are several reasons why studying “Unclassified” TT or attempting to classify them is vital. Phenotyping of TT Status is a complex process requiring specialised equipment, which makes removing 20-40% of responses inefficient. Some NC individuals respond similarly to either TT or TnT, indicating that they were, in fact, incorrectly classified TT or TnT (Thibodeau et al., 2019). Thermal taste also occurs on a spectrum of responses (Skinner et al., 2018), making classification according to existing “methods” somewhat arbitrary. In addition, several response scenarios can

cause a non-classification (Skinner et al., 2018). As the understanding of thermal taste continues to evolve, classification methodologies should continue to adapt alongside this understanding.

### **1.5 Potential relationships between thermal taste and demographic, food behaviour, personality and orosensory responsiveness**

Although some of these factors are mentioned above, they are laid out below as a theoretical framework for the experimental approach of the PhD.

#### **1.5.1 Demographics (gender, age, ethnicity)**

As mentioned, no studies evaluating sex/gender and TT Status have found any relationship between TT Status and sex/gender (Bajec & Pickering, 2008; Pickering & Kvas, 2016; Pickering et al., 2016; Thibodeau et al., 2019; Yang, 2015; Yang et al., 2020). However, gender is related to food involvement (Rodrigues et al., 2020; Yang, 2015), some aspects of which have been directly (Yang, 2015) and indirectly (Pickering & Pickering, 2022) linked to TT Status.

Thus far, only Thibodeau et al. (2019) has found an age relationship with TT Status; in their cohort (aged 17-30), TT were slightly younger than TnT. Taste sensitivity (Jilani et al., 2022; Mojet et al., 2001; Mojet et al., 2003; Sato et al., 2022; Yang et al., 2022) and taste recognition ability (Doty et al., 2017) decreases with advanced age, although trigeminal sensitivity is less affected (Yang et al., 2022), all three factors involved in traditional thermal taster phenotyping. In addition, the two studies investigating TT Status and ethnicity had contrasting findings; in a Canadian population, Thibodeau et al. (2019) found no relationship between TT Status and ethnicity, whereas in a UK-based cohort Yang et al. (2020) found a relationship between TT Status and Asian ethnicity.

Overall, this points to a gap in knowledge regarding the effect of thermal taste in a larger population, using an improved phenotyping method which does not rely on taste sensitivity or recognition.

## **1.5.2 Food Behaviour**

### *1.5.2.1 Food behaviours linked to orosensitivity*

As a result of the overall increased orosensitivity (section 1.3.1), TT are expected to be more food neophobic, as measured by the food neophobia scale (FNS) (Pliner & Hobden, 1992). However, studies examining the relationship between TT Status and food neophobia have had contradictory results. Bajec and Pickering (2008) found TT to have higher (but non-significant) food neophobia than TnT, whereas Yang (2015) found TT to have lower food neophobia – reporting TT to be significantly more likely to sample new and different foods. These differences may point to yet unexplored differences in the "approach" and "avoidance" tendencies in Neophobia as detailed by Nezlek and Forestell (2019).

Orosensitivity can also impact food modification behaviour. Puputti, Hoppu, et al. (2019) found more orosensitive participants to be more likely to add ketchup to a meal and less likely to consume pungent foods. In addition, their study found no differences in overall recalled pleasantness between orosensitivity groups. However, whether these differences in food behaviour extend to TT and TnT is still unknown.

### *1.5.2.2 Food behaviours with potential mechanistic relationships with TT*

Two aspects of orosensation, namely texture sensation and chemesthesis, may share common underlying causes with thermal taste, resulting in differences in food behaviour. Previous studies have proposed the TRP channels, TRPM5 (Sweet) (Nachtigal & Green, 2020; Skinner et al., 2018; Talavera et al., 2005) and TRPV1

(Salty) (Talavera et al., 2007) as mechanisms for thermal taste. TRP channels are also responsible for the chemesthetic orosensations (burning, tingling, cooling, numbing) of herbs and spices such as capsaicin (TRPV1), garlic (TRPA1& TRPV1), black pepper (TRPA1), ginger (TRPA1), wasabi (TRPA1) and menthol (TRPM8) (Latorre et al., 2009). The TRP channels TRPA1 (Castiglioni & García-Añoveros, 2007) and TRPV4 (Suzuki et al., 2003) are involved in mechanosensation, including texture sensation. Differences in sensory experiences could lead to food behavioural differences between different groups.

Some evidence points to texture and spice-related perceptual and behavioural differences between TT and TnT. TT are more sensitive than TnT to chemesthetic stimuli, including capsaicin (Yang et al., 2014), CO<sub>2</sub> (Hort et al., 2016) and ethanol (Small-Kelly, 2018) and report different sensory experiences from ethanol than TnT (Thibodeau & Pickering, 2021). However, how these sensory differences relate to spicy food liking and consumption frequency is unclear. In addition, previous studies investigating food-related behaviours in TT and TnT (Bajec & Pickering, 2010; Pickering & Klodnicki, 2016) have pointed towards texture-driven preference differences. However, no data currently exists on the oral interaction of TT and TnT with different textures or their typical orosensory processing preference grouping as detailed by Jeltama et al. (2015) (“Smoocher”, “Cruncher”, “Chewer”, and “Sucker”).

### *1.5.2.3 Other food behaviours*

Previous research (Yang, 2015) found TT to score higher in specific aspects of the food involvement scale (Bell & Marshall, 2003) than TnT. (TT scored higher than TnT on five questions; “I don’t think much about food”, “Cooking and BBQ is not much fun”, “My food choices are not very important” – all reversed) (Yang, 2015). The findings of Pickering and Pickering (2022) that TT are more likely to be “foodies” can also be

considered connected to food involvement. Differences between TT and TnT in food involvement may indirectly result from an enhanced ability to describe orosensations, as correctly describing chemesthetic stimuli has been linked to food involvement (Byrnes et al., 2015). TT demonstrate this enhanced ability in complex aqueous solutions (Thibodeau & Pickering, 2021). However, the relationship may be an artefact of participants' responses during TT phenotyping, as those with better orosensory descriptive abilities are likely to describe their experiences during phenotyping more accurately and consistently – resulting in successful phenotyping (TT classification rather than UC/NC).

### **1.5.3 Personality (BFI/PBC)**

Personality differences, measured by the Big Five Inventory (BFI) (John & Srivastava, 1999), were linked to TT Status by Yang (2015). Although there were no significant differences in the five dimensions of the BFI (Openness to experience, Extraversion, Conscientiousness, Neuroticism, Agreeableness), TT tended to score higher in the Openness dimension, with TT being more imaginative, curious, inventive, deep thinking and like to reflect different ideas. In addition, Pickering and Pickering (2022) found TT more likely to be foodies, but also found foodies to score higher in Extraversion and Openness to experience, further suggesting a link between personality and TT Status. Note that Nezlek and Forestell (2019) linked the BFI to the FNS where a subfactor of the BFI (Approach) strongly correlated to the Openness dimension (Nezlek & Forestell, 2019).

Private body consciousness (PBC) (Miller et al., 1981) is a personality construct not yet studied across thermal taste phenotypes. PBC measures a person's confidence in their ability to perceive and interpret bodily sensations. Participants with higher PBC tend to rely more on sensory characteristics than external cues to make food choices (Deliza

& MacFie, 1996). Chronically ill participants with higher PBC exhibited higher processing speed (van der Werf et al., 2002), better symptom reporting capabilities (van der Werf et al., 2002), but also higher intensities of pain (Ferguson & Ahles, 1998). PBC might affect the ability to perceive thermal taste– or recognise thermal taste as a taste experience.

#### **1.5.4 Other Taste phenotypes (orosensitivity)**

Although PROP taster status is independent of TT Status (Bajec & Pickering, 2008; Bering et al., 2013; Thibodeau et al., 2019; Yang et al., 2014) two additional taste phenotypes may be linked to thermal taste due to common mechanistic underpinnings via TRMP5. There is evidence of an interaction between thermal taste and sweet (chemosensory) taste; Thibodeau et al. (2020) found Umami/Savoury TT to perceive sweet taste more intensely. Although this finding was insignificant after correcting for multiple comparisons, it points towards differences between thermal taster subgroups in sweet taste sensitivity, which also differs between Sweet Liker subgroups (Kavaliauskaite et al., 2023). Although a single study investigating TT Status and Sweet Liking found the two phenotypes to be independent (Yang et al., 2020), an interaction may exist. In addition, there is a potential unexplored link between fatty acid responsiveness (oleogustus) and thermal taste, as TRMP5 is involved in fatty acid perception (Liu et al., 2011; Sclafani & Ackroff, 2022).

#### **1.6 Research gaps**

A first gap is the absence of large studies investigating the effect of TT Status on food choice, partially due to the difficulty and immobility of Thermal Taster phenotyping. The first step to addressing this gap is developing a phenotyping method to phenotype participants more rapidly and reliably (Pickering et al., 2016). There is also an

opportunity to try and reduce the proportion of uncharacterised participants or to study those participants as a separate group (Thibodeau et al., 2019), as most studies have only studied TT and TnT. Beyond this, there is an opportunity to further study subgroups of thermal tasters as indicated by Skinner et al. (2018).

A second gap is the absence of information on temperature-taste relationships in thermal tasters, limited by the phenotyping protocol typically employed, which only allows participants to report thermal taste after their thermal treatment. Using a time-intensity protocol Skinner et al. (2018) showed a link between Sweet thermal taste and the activation temperature range of TRPM5. Nachtigal and Green (2020) presented further evidence of this connection. However, collecting more descriptive temporal data could enable understanding mechanisms for additional thermal tastes.

A third gap is further exploring the relationship between TT Status, taste phenotypes, personality factors, and other food behaviours. The relationship between TT Status and PROP has been well-studied, but only one study on TT Status and Sweet liker status (Yang et al., 2020), and none on oleogustus/fat taste sensitivity. Preliminary evidence also exists that TT differ from TnT in terms of personality, food neophobia and aspects of food involvement (Yang, 2015). These aspects need to be explored with a larger sample to confirm the findings. In addition, no studies to date have explored how TT engage with spicy food, their oral processing behaviour (Jeltema et al., 2015) or whether they tend to modify their foods through taste-masking and taste-enhancing additions (Puputti, Hoppu, et al., 2019) differently than TnT.

### **1.7 Thesis objectives and structure**

The identified research gaps led to four overall research objectives, explored in four research chapters with the following objectives:

**Chapter 2:** To improve current thermal taste phenotyping. Specifically, to:

- 1) Develop a phenotyping method that combines the two temperature exposures traditionally used with temporal data collection.
- 2) Validate this novel phenotyping approach compared to existing approaches in terms of agreement, effectiveness, and efficiency.

**Chapter 3:** To establish response patterns of thermal taste. Specifically, to:

- 1) Establish groups of thermal tasters based on their thermal taste response patterns.
- 2) Establish links between specific thermal tastes and temperatures and temperature ramps.

**Chapter 4:** To establish potential relationships between TT status groups and TT subgroups and other relevant measures. Specifically, to:

- 1) Profile a group of phenotyped participants on demographic measures, food behaviour, personality and orosensory sensitivity.
- 2) Establish the relative importance of different measures and their constituents in predicting TT Status and TT subgroups.

**Chapter 5:** To establish the relative relationship between thermal taste and food choice. Specifically, to:

- 1) Establish the relative importance of thermal taste in determining food choice when also considering demographic, food behavioural, personality and orosensory sensitivity measures.
- 2) Establish patterns of food consumption behaviour in relation to thermal taste.

### 1.7.1 Thesis structure

Each research chapter draws an English idiom to explain its central theme.

**Chapter 2**, “*Fast and furious*”, presents and validates a novel phenotyping method for thermal taste. The question of whether the method is “faster”, and more “furious”; or more effective is addressed. **Chapter 3**, “*Divide and conquer*” clusters (“dividing”) thermal taste data obtained using the novel method presented in Chapter 2 and then further extracts (“conquering”) time and taste relationships. In **Chapter 4**, “*Seeing the forest and the trees*”, the participants phenotyped in Chapter 3 were profiled for demographic, food behavioural, personality and orosensory sensitivity measures. A machine learning approach was used to untangle the relationships between thermal taste and these measures, focussing both on the bigger picture (“the forest”) and the effects of specific factors (“the trees”). **Chapter 5**, “*The path to enlightenment*” evaluates the impact of thermal taste in the context of food choice by using a path-modelling-related approach to demystify the relative relationships between thermal taste, food choice and other factors (demographic, food behavioural, personality and orosensory sensitivity).

## **Chapter 2 *Fast and Furious*: Developing and validating a rapid (fast) portable temporal phenotyping method for phenotyping thermal taste (RapCoTT) with improved efficiency (furious)**

Note: the research in this chapter was published as Botha et al. (2021) *Comparing a new rapid combined method (RapCoTT) with traditional approaches for phenotyping thermal taste*. *Physiology & Behavior*. 238:113482. Contextually, it has been edited to better align with the objectives of the thesis.

### **2.1 Introduction**

Orosensory differences between thermal tasters (TT) and thermal non-tasters (TnT), as phenotyped with existing methods, were detailed in Chapter 1. In short, traditionally phenotyped TT exhibit higher orosensory responsiveness than TnT (Bajec & Pickering, 2008; Eldeghaidy et al., 2021; Green et al., 2005; Green & George, 2004; Hort et al., 2016; Pickering et al., 2009; Small-Kelly & Pickering, 2019; Thibodeau et al., 2020; Thibodeau & Pickering, 2021; Thibodeau et al., 2019; Yang et al., 2014; Yang et al., 2020). However, differences reported in the literature are between “classifiable” TT and TnT, representing only a subset of the population. Traditional phenotyping methods leave approximately 50% of participants unclassified (Thibodeau et al., 2019), meaning that previously described orosensory differences may, at least partially, be an artefact of the differences between participants that can be reliably phenotyped. One aspect of this is that most phenotyping methods require TT to report thermal taste above a specific threshold, and the intensity of thermal taste reported may indicate orosensory responsiveness more generally. Thibodeau et al. (2020) found that the orosensory responsiveness of unclassified (UC) participants falls between that of TT and TnT, further strengthening this idea. Therefore, a phenotyping method that is more effective (with a greater

proportion of participants correctly classified) and does not rely on the intensity of the reported thermal taste is a prerequisite for clarifying the validity of previously reported differences.

Although there is evidence that TT differ from TnT in their attitudes to food (Pickering & Pickering, 2022; Thibodeau et al., 2023), studies comparing the hedonic responses of TT and TnT have generally been inconclusive (Pickering et al., 2010; Pickering et al., 2009; Q. Yang et al., 2018) or showed only minor differences (Bajec & Pickering, 2010; Pickering & Klodnicki, 2016; Pickering et al., 2016). The lack of conclusive findings is likely due to low sample sizes (n ranges from 24-76), resulting in low power. Typically, to achieve sufficient power to establish consumer-relevant differences, a sample size 75-150 is recommended (Lawless & Heymann, 2010) when considering expected variability in consumer responses. Using larger populations when studying thermal taste and food choice could improve understanding in this area.

The inefficiency of phenotyping methods is a crucial contributor to historically small sample sizes, as, in addition to their low classification rates, existing approaches are time-consuming. Current phenotyping involves applying a temperature stimulus to the tongue in two separate trials. During the first (warming) trial, the thermode tip, placed on the tongue, is cooled at a rate of  $1^{\circ}\text{C} / \text{s}$  ( $35^{\circ}\text{C}$  to  $15^{\circ}\text{C}$ ), then warmed up ( $15^{\circ}\text{C}$  to  $35^{\circ}\text{C}$ ), and participants report the intensity of any tastes perceived during the warming portion on the general labelled magnitude scale (gLMS). This is followed by a cooling trial, where the thermode tip is cooled down ( $35^{\circ}\text{C}$  to  $5^{\circ}\text{C}$ ) and held at  $5^{\circ}\text{C}$  for 5s. Based on the intensity of their taste responses, participants are then classified as either TT or TnT.

Typically, participants need to report the same taste above weak on two replicates

(Bajec & Pickering, 2008; Green & George, 2004) but some studies have removed the requirement for the tastes to be identical on both trials (Skinner et al., 2018; Q. Yang et al., 2018; Yang et al., 2020) and others have removed the requirement for taste ratings to be above weak (Nachtigal & Green, 2020; Skinner et al., 2018; Yang et al., 2020). Early thermal taste phenotyping studies only recognised sweet, sour, bitter and salty reported tastes (Bajec & Pickering, 2008; Green & George, 2004), but Savoury/Umami and metallic were subsequently included (Yang et al., 2014), and some studies recognise spicy and minty (Skinner et al., 2018).

The rules adopted result in many participants being classified as neither TnT nor TT. These participants report taste but do not meet the requirements for classification as TT and constitute a third unclassified (UC) group. Bajec and Pickering (2008) introduced the criteria to control for response bias but likely confounds classification as a side effect. For example, differences in scale use and the ability of participants to accurately describe taste influence current classification (Skinner et al., 2018). It is also possible that, on the first replicate, some participants perceive taste but are either unaware that what they perceive is taste or do not categorise it as such. Skinner et al. (2018) included a third phenotyping replicate for unclassified participants and found a minor increase in phenotyping efficiency. By re-examining historical data, Thibodeau et al. (2019) made a compelling case that it was highly likely that some unclassified participants were misclassified, as additional orosensory responsiveness data to tastes and temperature showed they mirrored the responses of either TT or TnT. The same authors also found that unclassified participants who reported a taste but did so inconsistently between replicates demonstrated similar orosensory responsiveness to TT (Thibodeau et al., 2020), further supporting the need for better classification criteria and phenotyping methodology.

The sensory experience induced by a changing temperature stimulus is temporal by nature, but earlier studies have failed to examine this. Collecting temporal taste data during phenotyping would enable the temperature stimulus linked to each taste response to be identified. In addition, temporal methods could improve the accuracy of taste data, as the practice of delayed data collection could affect reported perceived tastes through primacy and recency effects. When participants are presented with sequential taste stimuli with differing delays between reporting, they are better able to recognise the first taste (primacy) after a longer (60s) delay and the last (recency) after a short (15s) delay (Daniel & Katz, 2018) indicating that both effects can occur in the 30-45s taken for thermal taste phenotyping. Primacy effects of tastes perceived during the warming trial's initial (cooling) portion are of particular concern as individuals may still remember tastes from the cooling phase, which they are not supposed to be attending to. Temporal methods collect taste data in real-time, allowing the two trials to be combined. Preliminary work by Skinner et al. (2018) partially joined the trials and collected data throughout temperature stimulation in a study acquiring perceptual and functional Magnetic Resonance Imaging (fMRI) data. They collected time-intensity data during temperature exposure but proposed that temporal check-all-that-apply (TCATA) would be a suitable data collection method.

TCATA is typically used to characterise food products and requires participants to check and uncheck tastes as they are perceived (Castura et al., 2016). TCATA has been applied to characterise food and beverage matrices (Reyes et al., 2017 ; Tan et al., 2019) and is effective with consumers (Ares et al., 2017; Jaeger et al., 2018; Jaeger, Beresford, et al., 2017; Ramsey et al., 2018) including children (Velázquez et al., 2020a; Velázquez et al., 2020b). As the data obtained shows the start and end time of a sensation, the resolution obtained from TCATA data collection could significantly

aid in understanding thermal taste as a phenomenon.

### **2.1.1 Aims**

This chapter aimed to develop and validate a novel phenotyping method that combines the two temperature trials typically used (as first proposed in Bajec & Pickering, 2008) into a single temperature treatment with concurrent temporal data collection. Both improvements were inspired by the suggestions of Skinner (2017). It was expected that the new method would be more effective, and more efficient than existing methods, and would agree with existing phenotyping approaches.

From the objectives, the specific aims of this chapter were to:

- 1) develop a novel phenotyping method that incorporates TCATA data collection alongside a thermal trial.
- 2) validate the effectiveness of this new protocol compared to existing protocols.
- 3) validate the agreement of this new protocol with existing protocols and classification schemes.
- 4) establish any age or gender differences between phenotyped TT and TnT
- 5) establish potential improvements to the protocol for future implementation.

## **2.2 Materials & methods**

The work contained in this chapter consists of two parts: the development of the new thermal taste phenotyping method and the validation study testing this method against existing approaches. Method development included building the thermode device and its associated software, ensuring that the thermode could deliver the correct protocols and that the interface software could collect the appropriate data. During method validation, the newly developed protocol was tested against

conventional protocols for thermal taste phenotyping.

### **2.2.1 Design of thermode and associated protocols**

The author led the design of the thermode and the associated software. Although the following sections show the necessary details, Appendix C shows the final briefing document used as a reference during development.

#### *2.2.1.1 Thermode and data collection system*

A portable thermode device, akin to devices used by previous groups (Bajec & Pickering, 2008; Cruz & Green, 2000), was built in Massey's then School of Food & Advanced Technology's mechatronics lab with input from the author. The device consisted of a Peltier device (45mm<sup>2</sup>, model CM23-1.9, II-VI Marlow, USA), positioned on an aluminium rod enclosed by a plastic pipe (Figure 2-1 d). The aluminium rod served as a heat sink and handle, and the plastic enclosure thermally insulated the pipe and avoided heat transfer from a participant's hand. The device was connected to a box containing an Arduino microcontroller (Atmel ATmega328) and a Bluetooth connector (HC-05). The device's consistency in delivering temperatures (range and rate of change) for temperature trials was validated through calibration with an additional temperature probe.

The development of custom software for data collection and thermode control allowed for the setup, as shown in Figure 2-1. This software (Thermode Controller Software v.1.0, Palmerston North, New Zealand), was built by the School of Psychology, with input from the author using JavaScript on the Electron platform (<https://www.electronjs.org/>). The software simultaneously controlled and monitored the temperature of the thermode and collected participant response data.

When the system was fully set up, participants held the thermode against their tongues

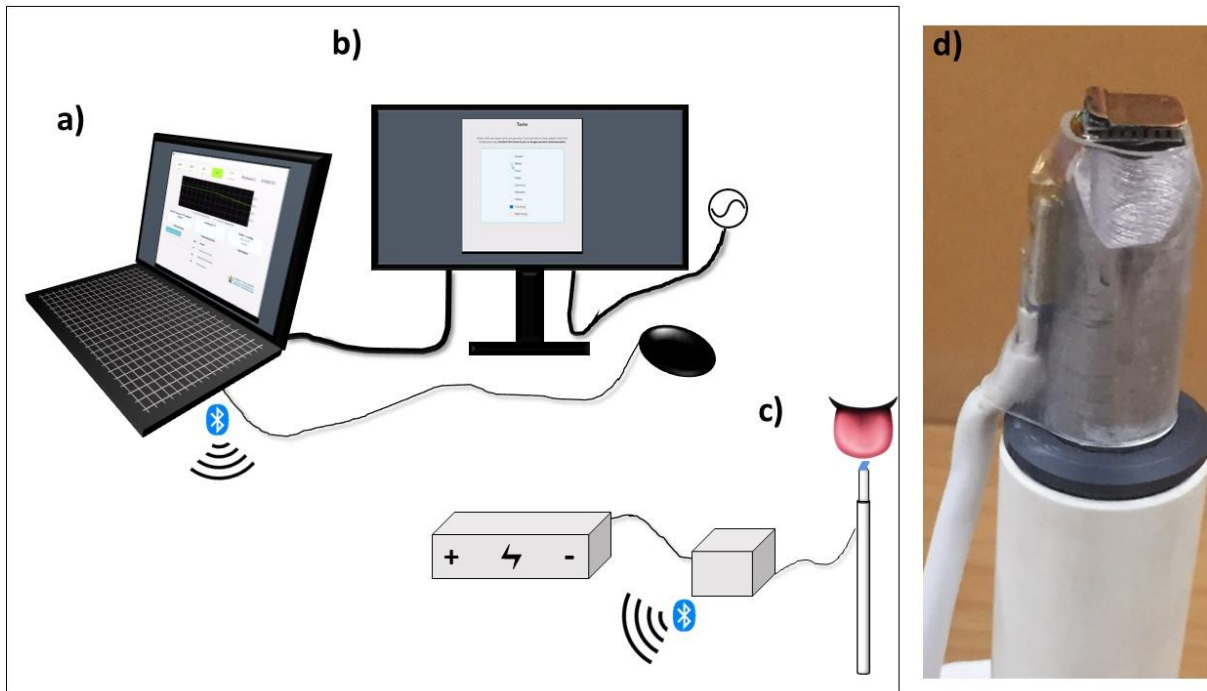


Figure 2-1 Overview of the system, connected via Bluetooth a) Laptop computer running joint thermode control/data collection software with primary display for researcher b) Secondary display and mouse (participant interface) c) Thermode consisting of a control box, the thermode itself and power supply d) detailed picture of the tip of the thermode showing the Peltier device sandwiched between a copper plate and the aluminium rod (covered with tasteless plastic film during phenotyping).

and interacted with the software through a mouse and computer monitor display. Concurrently, the researcher could monitor and control thermode functions and monitor participant responses on a laptop display. In addition, the researcher could see the placement of the thermode at all times, and monitor any small movements of the thermode, as movement would result in disturbances in temperature delivery. In the experimental setup, participants activated the protocol with a mouse click once the thermode was positioned correctly on their tongue (Figure 2-1).

### 2.2.1.2 Phenotyping protocols and classification schemes

The intent of the method validation study (section 2.2.2) was to compare the new phenotyping method with two existing methods. These methods and their corresponding data collection protocols were programmed into the thermode and its interface software. The new method had a unique, combined temperature protocol and classification approach, while the existing methods employed the same temperature

protocol (traditional) but differed in classification approach (Scheme A and Scheme B).

#### 2.2.1.2.1 Traditional Protocol

The traditional protocol was based on Bajec and Pickering (2008) and consisted of two trials; warming and cooling (Figure 2-2 a. and b.). During the warming trial, the surface temperature of the thermode decreased from 35°C to 15°C and then increased to 35°C at a rate of 1°C/s; participants were instructed only to report tastes perceived during the warming part. During the subsequent cooling trial, the temperature decreased from 35°C to 5°C, which was maintained for 5s; participants were instructed to report any tastes they perceived during the whole trial. As in earlier studies, the cooling trial followed the warming trial to eliminate potential numbing from prolonged cooling. After each trial, participants reported the maximum temperature intensity and any tastes perceived on the gLMS. Three replicates of both trials were performed, with a short break (30-60 s) in between.

Three replicates were included to facilitate the application of two classification schemes, as inconsistent classification criteria exist across studies for traditional phenotyping. Scheme A followed the current application of Bajec and Pickering (2008) as presented in Thibodeau et al. (2019) and Scheme B followed Skinner et al. (2018). Scheme A considered all participants who reported any taste (Sweet, Sour Bitter, Salty, Savoury, Metallic) consistently above weak in both initial replicates of the same trial as TT. Scheme B considered all participants who reported an expanded list of tastes (Sweet, Sour Bitter, Salty, Savoury, Metallic, Minty, Spicy), regardless of intensity, on any two replicates of the same trial as TT. In both schemes, participants who reported no tastes were considered TnT, and those who reported taste in any trials but did not meet the classification criteria (either through consistency or intensity) were considered UC.

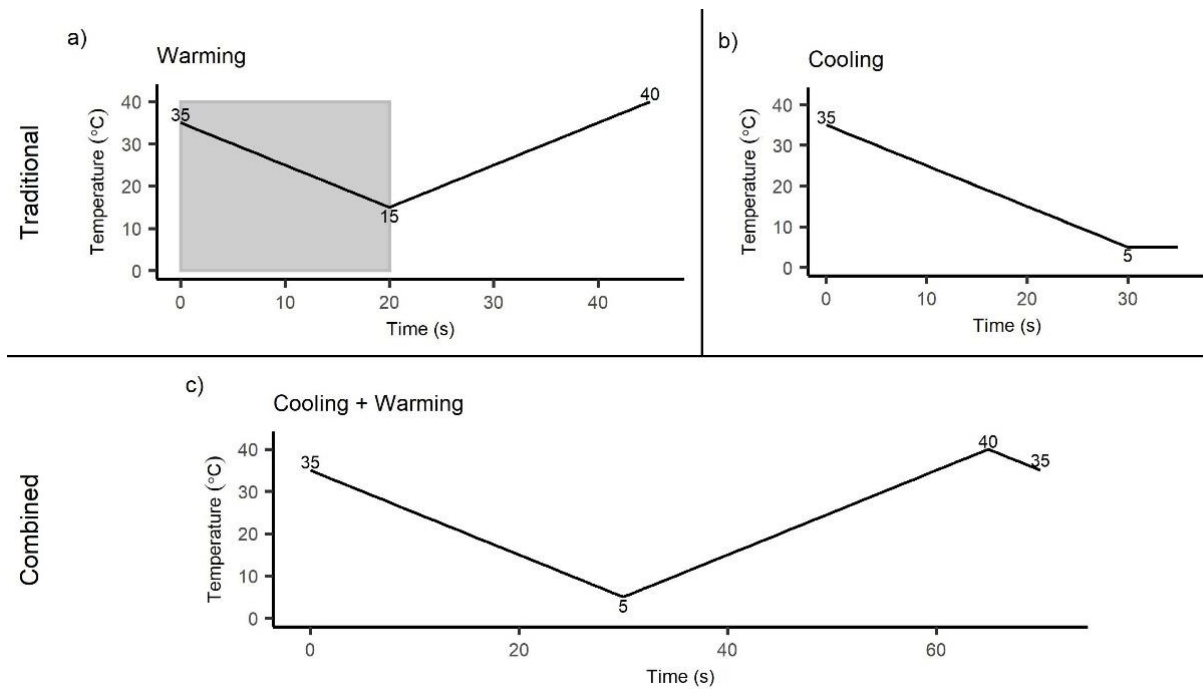


Figure 2-2 Summary of a) and b) the traditional protocol and c) the combined protocol. The traditional protocol consists of two trials, a) warming (35°C – 15°C – 40°C) and b) cooling (35°C – 5°C held for 5s). c) The combined protocol only consists of one temperature trial, which has a cooling portion (35°C – 5°C) a warming portion (5°C – 40°C) and return or end portion (40°C – 35°C).

#### 2.2.1.2.2 New Temperature Protocol

The new temperature protocol combined cooling and warming trials, as proposed by Skinner et al. (2018). The surface temperature of the probe decreased from 35°C down to 5°C, increased to 40°C, then decreased to 35°C at a rate of 1°C/s (Figure 2-2 c.). In the combined trial, the tongue is automatically returned to its baseline temperature, starting at the same temperature it finishes, reducing the need for a break between the two temperature trials. This approach was first used by Skinner et al. (2018) to overcome the movement constraints of an fMRI protocol.

Participants were instructed to use a temporal check-all-that-apply (TCATA) paradigm following Skinner et al. (2018) to indicate in real-time when they perceived a taste and/or temperature throughout the trial. Participants viewed a screen with tastes (Sweet, Sour, Bitter, Salty, Savoury, Metallic), Cooling, Warming and “Other” as checkboxes. They were instructed to select any taste or temperature sensation they

experienced and deselect it as soon as the experience stopped (attribute fading (Ares et al., 2017) was not used). Any participant who reported “Other” was asked to describe the sensation they experienced after the temperature exposure had ended. Participants were trained in the TCATA method by performing a test run, during which their responses to “Cooling” and “Warming” were monitored to assess whether they understood the instructions. After the test run, they were given feedback and were allowed to ask questions about the instructions. Two subsequent replicates were performed for classification. Scheme C considered participants to be TT if they reported any taste during both replicates – which is analogous to Scheme B. No intensity criteria were included, as this was not collected. With combined warming and cooling phases, the new method, concurrent data collection and classification Scheme C was named ‘Rapid Combined Phenotyping for Thermal Taste’, or RapCoTT for short.

### **2.2.2 Method validation study**

After developing the RapCoTT protocol, a method validation study validated the efficacy of the new protocol in an experimental setting.

#### *2.2.2.1 Participants*

Participants (n = 140) aged between 18 and 65 were recruited from the general New Zealand population using the Feast (Food Experience and Sensory Testing Lab) participant database, the university’s internal email system and posters on campus. Full ethics approval was obtained (Massey Human Ethics Committee Southern A application number 19/26, Appendix A). Participants were remunerated for their time with shopping vouchers (NZ\$40).

#### *2.2.2.2 Experimental procedure*

Each participant individually attended a 1-hour session. Participants were informed

that the experiment aimed to investigate responses to temperature stimulation of the tongue and that sometimes tastes may be perceived, but not always. Participants were trained to use the gLMS, rating the intensity of a range of real and imagined sensations on a gLMS scale (Bartoshuk et al., 2004). As a familiarisation, participants tasted and verbally identified Sweet, Sour, Bitter, Salty, Umami/Savoury and Metallic solutions (Table 2-1). Participants were prompted for associations if they struggled to identify tastes (Bartoshuk et al., 2004), often necessary for Umami/Savoury and Metallic. After tasting, participants rated the intensity of each solution, which served as a further familiarisation step with the gLMS.

*Table 2-1 Tastes and concentrations used for taste familiarisations. Adapted from Skinner (2017)*

<b>Taste</b>	<b>Reference</b>	<b>Concentration</b>	<b>Source</b>
Sweet	Glucose	116g/l	Sigma-Aldrich, NZ
Sour	Citric Acid	1.5g/l	Davis Food Ingredients, NZ
Bitter	Quinine Sulphate	0.017g/l	Sigma-Aldrich, NZ
Salty	Sodium Chloride (NaCl)	10g/l	Sigma-Aldrich, NZ
Umami/Savoury	MSG	20g/l	Davis Food Ingredients, NZ
Metallic	Copper sulphate pentahydrate	0.75g/l	Sigma-Aldrich, NZ

After familiarisation, the researcher gave a brief introduction to the thermode to acquaint participants with its purpose and functionality, as seeing unfamiliar equipment without explanation can be unsettling to participants. Before starting the first protocol, participants could use their fingers to acquaint themselves with the starting temperature of the thermode. They were instructed to position the thermode on the front of the anterior of the tongue in the middle, as this has previously been found to be the most responsive (Cruz & Green, 2000). The experimenter checked placement and adjusted as needed so that the positioning was the same for each run and across each participant. Participants were instructed not to move their tongues during the temperature runs, which was monitored by the researcher as described above. The first protocol that participants were to experience was outlined. The traditional protocol (Bajec & Pickering, 2008; Skinner et al., 2018) was conducted first

for half of the participants according to a balanced design, balancing potential learning effects. The remaining participants started with the new protocol.

At the end of the session, participant age and gender, according to Stats NZ (2019) criteria, were collected.

### *2.2.2.3 Evaluation of time cost of new protocol*

The time requirement for each step of each protocol was estimated. The average time taken for each participant for each protocol was compared.

### **2.2.3 Data analysis**

All data analyses were performed using R (R Core Team, 2018). An alpha level of 0.05 was adopted for all hypothesis tests. The packages `vcd` (Zeileis et al., 2017) and `exact2x2` (Fay, 2010) were used for hypothesis testing on contingency tables and visualising consistency. To correct for potential errors arising from small cell counts, tests for conditional independence were bootstrapped using Monte Carlo simulations (Friendly & Meyer, 2016), and the contribution of individual cells in Chi squared tests regarding deviation from independence were evaluated using the absolutes of Pearson residuals as detailed in Zeileis et al. (2007).

#### *2.2.3.1 Thermal Taster Status Classification*

Data were processed according to the classification schemes in section 2.2.1.2, resulting in three classifications per participant (Traditional – Scheme A, Traditional – Scheme B, Combined - Scheme C). To evaluate whether protocol order affected classification through inducing a learning effect, data were split according to which trial participants performed first, and a Pearson's chi-squared test for conditional independence was performed on each classification scheme. Classification schemes were evaluated separately to split the effect of the classification scheme from any order

effect.

The effects of age and gender on thermal taster status were evaluated using Pearson's chi-squared test. Age responses were converted to four categories of roughly equivalent size (19-26, 27-32, 33-39, 40+). Gender responses were converted into three categories: male, female and all other genders. Self-reported gender was used as a proxy for gender/sex, and any groupings smaller than 1% were removed only from analyses specifically related to gender.

#### *2.2.3.2 Method Efficacy comparison*

Efficacy was evaluated by comparing the number of participants classified (TT and TnT) and not classified (UC) for each method using a set of contingency tables. As the observations are paired (different classifications for the same participant), they are not independent, which makes tests for independence inappropriate (Agresti, 2002). The exact McNemar test (Fay, 2010) determined whether the odds ratio classification versus non-classification equalled 1. The exact version of this test was used to correct any inaccuracies that may arise from small sample counts in the relevant cells. The potential effect of protocol order on efficacy was further evaluated by splitting the data by protocol order.

#### *2.2.3.3 Consistency between methods*

RapCoTT was validated as an alternative to traditional approaches by comparing phenotyping outcomes. For example, if an individual was classified as TT using traditional methods, were they also classified as TT using RapCoTT? Were any classified individuals unclassifiable with RapCoTT? Three questions were of interest, namely whether participants were consistently classified into the same phenotype,

whether there were tendencies of disagreement and whether the nature of tastes reported was consistent across protocols.

The consistency of categorisation (overall scheme agreement) was evaluated using Cohen's  $\kappa$  (Cohen, 1960) following Thibodeau et al. (2019) and allowed results to be compared with previous evaluations of classification schemes. Agreement for each classification category (TnT, UC, TT) was similarly evaluated. It was expected that participants unclassified by Scheme A and Scheme B would be classified as either TnT or TT by RapCoTT; therefore, a weighted  $\kappa$  (Cohen, 1968), which accounts for movement to an adjacent category, was computed. The same diagnostics as in Thibodeau et al. (2019) were applied;  $\kappa < 0.2$  (poor),  $\kappa = 0.201-0.4$  (fair),  $\kappa = 0.401-0.6$  (moderate),  $\kappa = 0.601-0.8$  (good), or  $\kappa = 0.801-1$  (very good) (Kwiecien et al., 2011).

Observer agreement charts were also constructed. These charts (Bangdiwala, 1985) illustrate directional tendencies in classification, making it a complementary tool to traditional measures, like Cohen's  $\kappa$ , when categories are ordered (e.g. TnT, UC, TT) (Bangdiwala, 2017; Bangdiwala & Shankar, 2013). Agreement charts highlight trends in classification by showing the number of participants classified the same by two methods, as well as the proportion with changed categorisation. Weighting accounting for movement to an adjacent category (TnT/UC and TT/UC) was applied and visualised. Potential directional effects of protocol order were assessed by repeating these charts using the first protocol.

It was expected that participants would report similar tastes regardless of the phenotyping protocol. Pearson's chi-squared tests were applied to frequencies of tastes reported to evaluate the impact of the protocol on taste reported. Raw taste

counts were used from all three replicates of the traditional protocol and two replicates from the combined protocol. Responses from the combined protocol were split into three sets: cooling (35°C to 5°C), warming (5°C to 40°C), and end (40°C to 35°C). To correct for the multiple responses typically reported in TCATA, where a taste can be selected more than once in a run section, any one taste was only counted as selected once. As no analogous data exists in the traditional protocol for the end section, only data from the cooling and warming sections were considered. Frequencies were evaluated using two methods. First, for each protocol, the frequency of each taste reported was compared between cooling and warming. Second, for cooling and warming, the frequency of each taste was compared between protocols (traditional vs combined). Tests for conditional independence were performed for each set of variables (Cohen, 1980; Meyer et al., 2006).

#### *2.2.3.4 Method optimisation*

The combined protocol in this study followed Skinner et al. (2018), including the final five seconds (returning from 40°C to 35°C) per their protocol. This approach also allowed participants time to report tastes occurring at the warming element's end. The effect of ending the trial at 40°C without returning to 35°C was tested through comparison of the efficiency of the full and shortened trials using the McNemar test. To determine whether there were differences in how tastes were reported, frequencies of tastes during the warming portion (5°C to 40°C) were compared with the warming portion plus the final five seconds (5°C to 40°C to 35°C) using the chi-squared tests regarding tastes (as described in section 2.2.3.3). Potential taste reporting patterns during the final 5s were assessed by comparing these taste frequencies with the warming phase.

## 2.3 Results

### 2.3.1 Classification

Of the 140 participants, data were successfully collected from 132. Of those unable to be phenotyped, five (3.5%) were due to equipment failure, and three (2.1%) were due to failure to follow instructions. The phenotyped participants ranged in age from 19 to 65yrs (mean = 34.2 yrs, median = 31 yrs ;19-26 n =34, 27-32 n= 39, 33-39n=22, 40+ n = 37) and were 76.5% female (n=101), 22.7% male (n=30) and 0.8% non-specified (n=1). Classification frequencies resulting from the three classification methods are presented in Table 2-2.

*Table 2-2 Summary of percentages (numbers of participants) classified as either TT or TnT by two schemes (Scheme A + Scheme B) under the traditional protocol as well as the combined protocol (Scheme C)*

Category	Traditional Protocol Scheme A	Traditional Protocol Scheme B	Combined Protocol Scheme C
Classified	68% (89)	80% (106)	93% (122)
- Thermal non-tasters	23% (30)	15% (20)	29% (38)
- Thermal Tasters	45% (59)	65% (86)	64% (84)
Unclassified	33% (43)	19% (26)	7% (10)

The combined protocol and Scheme C (RapCoTT) reported the lowest number of unclassified participants (10 or 7%). Using the traditional protocol and Scheme A produced the highest number of unclassifieds (43 or 33%), followed by the traditional protocol and Scheme B (26 or 18%). RapCoTT and the traditional protocol with Scheme B classified a similar proportion as TT (84 or 64% vs 85 or 65%), with a lower proportion reported with the traditional protocol with Scheme A (59 or 45%). The traditional protocol with Scheme B classified the lowest number as TnT (20 or 15%), with the traditional protocol with Scheme A and RapCoTT delivering similar proportions of TnT (RapCoTT 38 or 29% vs Scheme A 30 23%).

Gender was not significantly associated with thermal taster status (TnT-UC-TT) for any classification approach (Scheme A:  $X^2 = 0.41$ ,  $df = 2$ ,  $p = .83$ ,  $V = 0.06$  95%CI

[0.00, 0.19]; Scheme B:  $X^2 = 0.15$ ,  $df = 2$ ,  $p = .95$ ,  $V = 0.03$  CI[0.00, 0.13]; RapCoTT:  $X^2 = 0.97$ ,  $df = 2$ ,  $p = .68$ ,  $V = 0.09$  CI[0.00, 0.23]). Age was also not significantly associated with thermal taster status for any approach (Scheme A:  $X^2 = 4.1$ ,  $df = 4$ ,  $p = .67$ ,  $V = 0.13$  CI[0.00,0.18]; Scheme B:  $X^2 = 5.58$ ,  $df = 4$ ,  $p = .45$   $V = 0.15$  CI[0.00,0.22]; RapCoTT:  $X^2 = 4.64$ ,  $df = 4$ ,  $p = .60$ ,  $V = 0.13$  CI[0.00,0.19]).

Protocol order was significantly associated with classification for Scheme A ( $X^2 = 8.22$ ,  $df = 2$ ,  $p = .016$ ,  $V = 0.25$  95%CI[0.04,0.41]) and approached significance for Scheme B ( $X^2 = 5.49$ ,  $df = 2$ ,  $p = .070$ ,  $V = 0.20$  CI[0.00,0.36]). Protocol order had a comparable effect size for RapCoTT, but this was insignificant ( $X^2 = 4.95$ ,  $df = 2$ ,  $p = .079$ ,  $V = 0.19$  CI [0,0.35]). The order effect had an impact on specific categorisations. Scheme A was more likely to classify participants as TT on their second protocol (Combined first: Pearson residual = 1.47,  $p = .041$ ; Traditional first: Pearson residual = -1.49,  $p = .032$ ). Scheme B was more likely to leave participants unclassified on their first protocol (Pearson residual = 1.45,  $p = .042$ ) when they completed the traditional protocol first. For Scheme C, no categorisation was more likely to occur based on protocol order, although there was moderate evidence (Pearson residual = -1.34,  $p = .072$ ) that participants were less likely to be TnT on their first protocol.

### **2.3.2 Efficacy of classification**

Classification frequency (Table 2-2) suggests that RapCoTT was the most effective phenotyping method, with only 7% of participants remaining unclassified. RapCoTT was more efficient than Scheme A (OR = 5.71,  $p < .001$ ) and Scheme B (OR = 3.28,  $p = .005$ ). Scheme B was also more effective than Scheme A (OR = 3.4,  $p = .004$ ) (Appendix D, Table D-1). When considering protocol order, RapCoTT remained more effective than Scheme A when completed first (OR = 3.2,  $p = .027$ ), but this advantage receded when compared to Scheme B (OR = 1.6,  $p = .58$ ). The difference in efficacy

between Scheme A and Scheme B approached significance when participants had completed the traditional protocol first (OR = 0.36  $p = .063$ ) (Appendix D, Table D-1).

### 2.3.3 Consistency across methods

Unweighted classification agreements between pairs of methods were ‘fair to moderate’ ( $\kappa = 0.31-0.58$ ), but when considering weighted classifications, agreement improved to ‘moderate to good’ ( $\kappa = 0.54-0.74$ ) (Table 2-3). Looking at thermal taster status categories in isolation, the agreement in classification between methods for TnT ranged from ‘moderate to good’ ( $\kappa = 0.57-0.76$ ) and for TT, it was ‘fair to good’ ( $\kappa = 0.34-0.6$ ). However, agreement was poor when comparing the UC category in RapCoTT with the two traditional methods ( $\kappa = -0.01 - 0.06$ ).

Table 2-3. Agreement of each pair of methods and classification category using Cohen’s  $\kappa$  (weighted and unweighted) as well as interpretation. Weighted Cohen’s  $\kappa$  allows for misclassification of one category (e.g. TnT to UC, UC to TT). Cohen’s  $K$ , diagnostic categories: poor ( $\kappa < 0.200$ ), fair ( $\kappa = 0.201-0.400$ ), moderate ( $\kappa = 0.401-0.600$ ), good ( $\kappa = 0.601-0.800$ ), or very good ( $\kappa = 0.801-1.0$ ).

Scheme	% match	Statistics					Agreement				
		All		Category			All		Category		
		U $\kappa$	W $\kappa$	TnT	UC	TT	U $\kappa$	W $\kappa$	TnT	UC	TT
RapCoTT vs Scheme A	57%	0.31	0.54	0.57	-0.01	0.34	Fair	Mod	Mod	Poor	Fair
RapCoTT vs Scheme B	69%	0.41	0.59	0.57	0.06	0.47	Mod	Mod	Mod	Poor	Mod
Scheme B vs Scheme A	74%	0.58	0.74	0.76	0.41	0.60	Mod	Good	Good	Mod	Good

Observer agreement charts, shown in Figure 2-3, visualise the maximum possible agreement between each pair of methods. Axes are scaled to the number of observations (participants) and divided proportionally to the number of observations in each category (TnT, UC, TT). Light grey rectangles represent the maximum possible agreement, while the black squares show the observed agreement. Dark grey shaded areas show partial agreement (TT to UC, TnT to UC, UC to TnT or UC to TT). The shape and position of light grey areas surrounding dark grey areas represent the numbers of participants classified in opposite categories (TT vs TnT). Light grey on

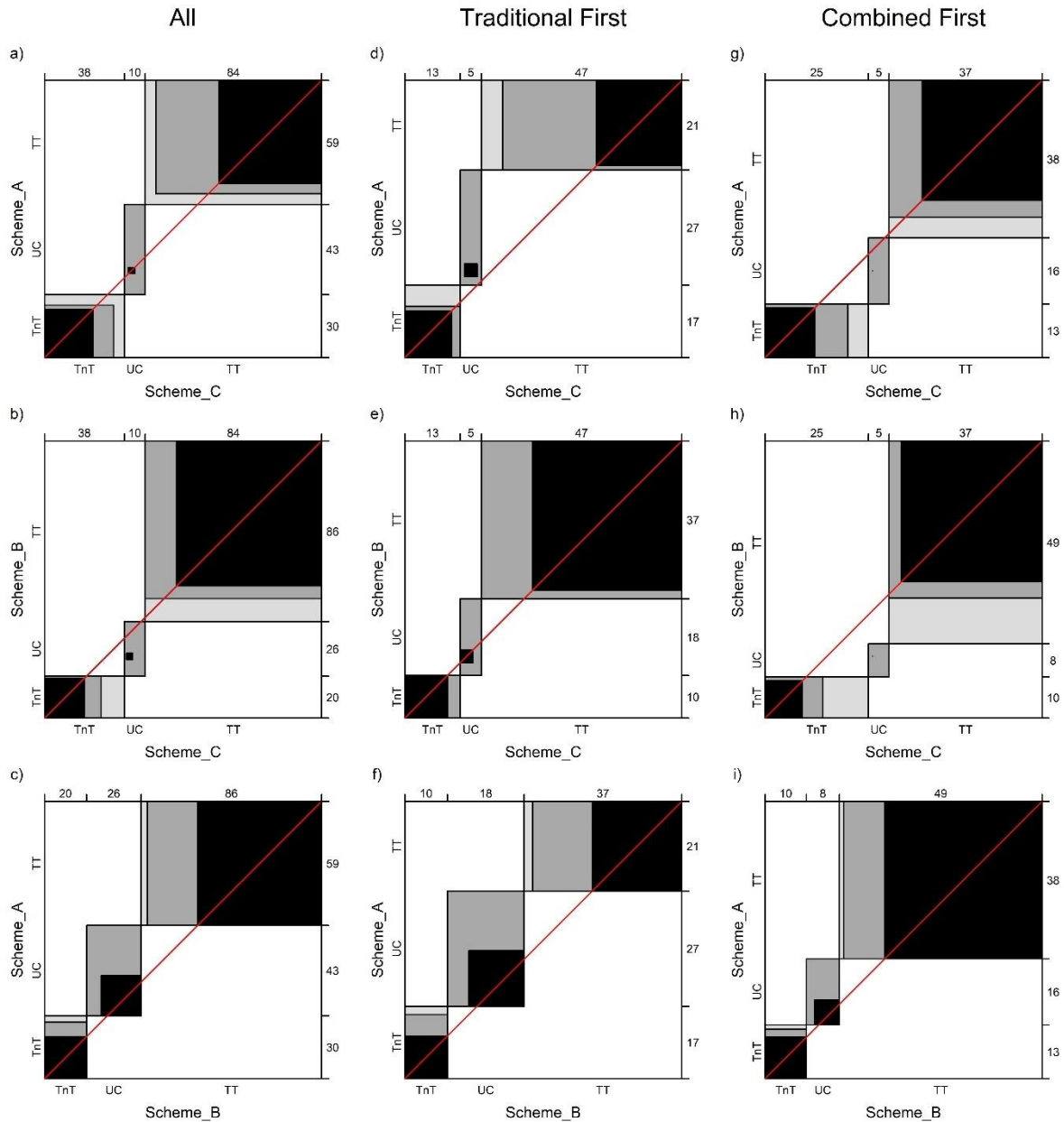


Figure 2-3 Agreement plots comparing each pair of methods using the full dataset (a-c) and subset (d-i) according to which protocol was performed first. In each plot, the axes are split into three categories (TnT, UC, TT) according to the number of participants classified in that category. The grey areas show the overlap between methods. Black squares are equal to the number of participants classified the same by both methods. Dark grey shading represents those classified in an adjacent category. The diagonal line is an interpretation guide for directional bias.

both sides of the dark grey indicated no tendency for one method to classify as TT and for the other to classify participants as TnT. The diagonal line indicates marginal homogeneity, again highlighting directional bias (the tendency of one method to classify in one direction). When the corners of all rectangles fall on the diagonal, this indicates the absence of bias (Bangdiwala, 2017).

The traditional protocol with Scheme A and RapCoTT agreed well when protocol order was not considered (Figure 2-3). The traditional protocol with Scheme B showed a slight tendency to classify participants as TT who were classified as TnT by RapCoTT; the proportion of participants classified the same was substantial ( $n = 91$ ) (Appendix D, Table D-2). Scheme A and B differed mainly in their tendency to classify participants in adjacent categories, an expected outcome given the differences in classification approach. When taking the order in which an individual performed the phenotyping into account, there was a tendency for some participants classified as TnT on the first phenotyping protocol and to be TT on the second (Traditional first: Scheme A vs Scheme C  $n = 5$ , Combined first Scheme A vs Scheme C  $n = 5$ , Combined first Scheme B vs Scheme C  $n = 11$ ) (Appendix D, Table D-2), with one exception. When the traditional protocol was first applied, RapCoTT and Scheme B classified all participants either the same or in adjacent categories, shown by the absence of light grey (Figure 2-3 e.). Patterns of agreement between Scheme A and B were unaffected by the sequential position of the traditional protocol. However, when participants completed the traditional protocol second, more participants were classified the same (Traditional first: Scheme A vs Scheme B  $n = 44$ , Combined first: Scheme A vs Scheme B  $n = 54$ ) (Appendix D, Table D-2).

### 2.3.3.1 Taste consistency

No significant difference in reported tastes existed between protocols in either the cooling ( $X^2 = 5.33$ ,  $df = 7$ ,  $p = .62$ ,  $V = 0.09$  95% CI[0.00, 0.12]) or warming trial ( $X^2 = 7.73$ ,  $df = 7$ ,  $p = .35$ ,  $V = 0.12$  CI[0.00, 0.16]). However, there were some relationships between tastes and temperature trials (Traditional:  $X^2 = 35.5$ ,  $df = 7$ ,  $p < .001$ ,  $V = 0.23$  CI[0.13, 0.29]; Combined:  $X^2 = 18.61$ ,  $df = 7$ ,  $p = .008$ ,  $V = 0.18$  CI[0.05, 0.24]). For both protocols, savoury was strongly associated with the warming trial (Combined:

Pearson residual = 2.19,  $p = .012$ , Traditional: Pearson residual = 3.19,  $p < .001$ ) and strongly negatively associated with the cooling trial (Combined: Pearson residual = -2.42,  $p = .002$ , Traditional: Pearson residual = -2.62,  $p < .001$ ). The traditional protocol showed some patterns that were not present in the combined protocol. Sour was negatively associated (Pearson residual = -1.93,  $p = .062$ ), and sweetness was positively associated (Pearson residual = 1.84  $p = .072$ ) with the warming trial of this protocol.

### **2.3.4 Temperature trial evaluation**

The classification of only one participant changed (TT to UC) when data from the final five seconds of RapCoTT were removed, which did not significantly reduce classification efficacy (OR =  $\infty$ ,  $p = 1.000$ ). Patterns of tastes reported were also unchanged between the warming phase and the inclusion of the final 5s ( $X^2 = 1.03$ ,  $df = 7$ ,  $p = 1.0$ ,  $V = 0.04$  95% CI[0.00, 0.00]). There were, however, differences when comparing taste frequencies directly between these two periods ( $X^2 = 24.78$ ,  $df = 7$ ,  $p < .001$ ,  $V = 0.23$  CI[0.10, 0.30]), savoury occurred more often during the final 5s (Pearson residual = 2.10  $p = .045$ ), and sour less often (Pearson residual = -2.21  $p = .040$ ).

The approximate time taken for the traditional method, RapCoTT, as applied in the current study, and time projections for the suggested optimised version of RapCoTT are tabulated in Table 2-4. The traditional protocol took, on average, 19 minutes per person, while RapCoTT as applied in this study took, on average, 10 minutes per person.

Table 2-4 Summary of approximated times for the traditional protocol, RapCoTT as applied during the present study, as well as proposed changes to RapCoTT

	Traditional	RapCoTT	RapCoTT shortened	RapCoTT proposed
<b>Training</b>				
gLMS training	5 minutes			
Taste training	5 minutes	5 minutes	5 minutes	5 minutes
Explanation	2 minutes			
Practice run TCATA				2 minutes
Practice temperature exposure		70 seconds	65 seconds	65 seconds
Break		60 seconds	60 seconds	60 seconds
<b>Rep 1</b>				
Combined temperature exposure and data collection		70 seconds	65 seconds	65 seconds
Warming exposure	45 seconds			
Warming Data collection	45 seconds			
Break	30 seconds			
Cooling exposure	35 seconds			
Cooling Data collection	45 seconds			
Break	30 seconds	30 seconds	30 seconds	30 seconds
<b>Rep 2</b>				
Combined temperature exposure and data collection		70 seconds	65 seconds	65 seconds
Warming exposure	45 seconds			
Warming Data collection	45 seconds			
Break	30 seconds			
Cooling exposure	35 seconds			
Cooling Data collection	45 seconds			
Break	30 seconds	30 seconds	30 seconds	30 seconds
<b>Total</b>	19 minutes, 10 seconds	10 minutes	9 minutes, 45 seconds	11 minutes, 45 seconds

## 2.4 Discussion

The inefficiency of existing phenotyping is a key limiting factor for thermal taste research, highlighting the need for a more effective method. This inefficiency is particularly relevant when considering the necessary participant sample sizes for consumer research. Thermal taste presents as a range of experiences, and factors beyond thermal taste itself (e.g. scale use and ability to identify taste) likely affect current classifications (Skinner et al., 2018). As the dynamic thermal stimulation traditionally used for phenotyping thermal taster status is by nature temporal, combining temperature protocols with temporal data collection presents as an obvious opportunity.

As expected from previous studies (Bajec & Pickering, 2008; Thibodeau et al., 2019; Yang et al., 2014; Yang et al., 2020), the present study showed no impact of sex/gender

on thermal taster status. A previous study has shown an effect of age on thermal taster status (Thibodeau et al., 2019), but the absence of an age effect in the present study is not strictly in contrast to this. The effects previously demonstrated were with the < 19-year-old age group, not represented in this study, and may be a factor of the large sample size ( $n = 394$ ), suggesting a weak association (calculated Cramer's  $V = 0.15$ , 95% CI [0.00, 0.23]). Age is just one example of a factor of which the effect the improved RapCoTT protocol will enable further study. The interaction between age and TT Status is of interest as age is known to impact taste (Mojet et al., 2001; Mojet et al., 2003; Puputti, Aisala, et al., 2019) and trigeminal perception (Piochi et al., 2020).

In the present study, the traditional approach produced classification outcomes similar to those of earlier studies, with some exceptions. Scheme A reported proportions of UC and TnT which fall within the ranges typically reported in the literature (UC:24–50% and TnT: 24–40%, summarised in Thibodeau et al. (2019)), but the proportion of TT was higher than typically reported (20–38% (Thibodeau et al., 2019)). This discrepancy is a likely consequence of the more recent inclusion of metallic as an acceptable taste for TT classification (Mitchell et al., 2019; Small-Kelly & Pickering, 2019). Scheme B reported proportions which differ from those reported previously (TnT 51%, UC 21%, TT 28%) (Skinner et al., 2018), which may have resulted from methodological differences between Skinner et al. (2018) and the present study. Here, participants were familiarised with basic tastes using tastant solutions (including metallic), while Skinner et al. (2018) only verbally familiarised participants with the tastes. Training with tastes increases taste recognition and identification ability (Hettinger et al., 1999), which may have led participants in the present study to be more tuned to tastes and more likely to perceive and report them, as has been shown

in the context of thermal taste phenotyping (Thibodeau et al., 2020). Skinner et al. (2018) also only performed a third replicate with participants unclassified after two replicates, whereas all participants in this study completed three replicates. As a result, Skinner et al. (2018) would have classified participants who only reported taste on the third replicate in the present study (UC) as TnT.

Classification frequency showed order effects, likely due to a learning effect and counteracted by the classification approaches of Scheme B and RapCoTT. Participants became more familiar with the process as they moved through trials, which likely increased confidence in their perception. Situational appropriateness may, however, also contribute to this phenomenon. Incongruity of experience with expectation can impact the intensity of experience (Carreiro et al., 2011), and participants do not expect to perceive taste from temperature stimulation. After the first experience of thermal taste, it becomes situationally appropriate, increasing the likelihood that participants will experience and report taste in subsequent trials (Nachtigal & Green, 2020). The lack of order effect on RapCoTT can be partially attributed to the absence of minimum intensity in classification criteria, as this controls for incongruence effects. The absence of an efficiency advantage over Scheme B when RapCoTT precedes the traditional protocol highlights the common benefit of dropping the intensity requirement. However, if the classification approach were RapCoTT's sole advantage, Scheme B would also consistently perform better than Scheme A. There is only moderate evidence of such an advantage on the first protocol performed (traditional protocol first), pointing to the advantages of TCATA.

The immediacy of TCATA data collection during temperature exposure, as opposed to data collection after temperature exposure, is a key factor in the overall increase in efficiency. In the traditional approach, participants must hold their response (taste

identity and intensity) in their working memory until the end of the trial. During the trial, participants may experience additional tastes alongside temperature sensation, resulting in the original taste information either decaying in terms of remembered intensity or being affected by interference (change in identity) (Daniel & Katz, 2018). This working memory load could also decrease the intensity of any further tastes experienced during the same trial, as cognitive task load reduces taste sensitivity (Liang et al., 2018) and taste responsiveness (van der Wal & van Dillen, 2013). The advantage of TCATA here is that it allows participants to report taste sensations immediately, avoiding memory effects and allowing participants to fully attend to any new taste experiences instead of trying to maintain their working memory.

Overall, RapCoTT agreed moderately to well with the two traditional methods, but the UC category agreed poorly, which is expected from the low proportion of unclassified individuals. This pattern mirrors previous findings where a classification scheme not allowing unclassified individuals (Green & George, 2004) agreed only moderately with Scheme A (Thibodeau et al., 2019). As RapCoTT's efficiency affects traditional agreement measures, the insights from agreement plots (Figure 2-3) are more relevant. They show no overall tendency for participants to be classified into opposing categories (TnT and TT). However, when taking protocol into account, agreement plots showed a tendency of some participants to be classified as TnT on the first protocol and TT on the second (Figure 2-3), perhaps reflecting learning or increased situational appropriateness as discussed above. This effect was most pronounced when RapCoTT was completed first and compared with Scheme B, where 11 initially TnT (RapCoTT) participants were classified as TT by Scheme B, and no participants changed in categorisation from TT to TnT. Interestingly, when the traditional protocol preceded RapCoTT, these schemes classified no participants in opposite categories. As

the classification approach is consistent between these schemes (no minimum intensity, identical list of approved tastes), the approach can be ruled out for this effect. However, the observed effect may be an interaction between the challenging nature of TCATA for some individuals missing a true taste response and/or response bias.

The fact that only one tongue location was tested for this study may be a limitation as the original protocol proposed by Bajec and Pickering (2008) tested three tongue locations, namely the anterior tip and a location to the left and the right of this point – an approach which has been adopted by some but not all researchers (see Thibodeau et al., 2019). Nevertheless, using only the anterior tip has been favoured by some researchers as it is the most responsive to thermal taste (Cruz & Green, 2000) and is the most innervated (Shahbake et al., 2005). During the current study, the researcher observed large variations in tongue morphology and extensibility, which may also introduce methodological variability regarding where the thermode can be applied. There is evidence that using only one tongue location may impact classification as classification schemes only using one tongue site delivered different classification proportions to those using all three sites (Thibodeau et al., 2019). However, it is essential to note that in the referenced study, the primary purpose was to compare classification schemes, and no direct comparison was drawn between tongue locations. To definitively answer the question of the number of locations, further work should be carried out evaluating RapCoTT across additional locations. It is also important to consider that the order effects observed in the present study would likely have been magnified by increasing the number of runs, which further draws into question whether differences seen through the three locations are true differences or whether they are order effects which have come about from greater response bias.

A minimum reported taste intensity has been previously used to control response bias

(saying a taste is perceived when it is not), but as taste intensity is not reported in RapCoTT, this could be a limitation. However, this minimum taste criterium has fallen out of favour recently with some researchers for traditional phenotyping (Nachtigal & Green, 2020; Skinner et al., 2018; Yang et al., 2020), as there is no direct evidence that “weak” on the gLMS is an appropriate cut-off point for categorising a participant as a TT (Nachtigal & Green, 2020). This cut-off point likely excludes participants who use the scale conservatively from thermal taster classification (Skinner et al., 2018). Response bias can be considered a form of (benign) deception. Reduced working memory capacity decreases the fluency at which participants can deceive an experimenter (Lisofsky et al., 2014; Sporer, 2016) especially the cognitive load accompanying a dual-task paradigm (Hu et al., 2015). Although the cognitive aspects of TCATA have not been studied to date, RapCoTT can be considered a multi-task paradigm as participants have to focus on holding the probe and tongue in position, focus on temperature and taste sensations and respond to them. Therefore, the complexity of TCATA is expected to inherently counteract response bias. In addition, including cooling, warming and “other” as options in the current protocol is intended to capture any potential response dumping. However, the protocol might be improved by including options for several trigeminal orosensations, to avoid further response dumping (Clark & Lawless, 1994) and by including a task which assesses participants for cheating behaviour, such as the difference spotting task (Liu et al., 2021). All protocol improvements present avenues for further research.

There was consistency in the nature of tastes evoked by both protocols, but differences in within-protocol comparisons point to previous evidence (Skinner et al., 2018; Thibodeau et al., 2020) that thermal taste is a complex phenomenon which can be better described through the full application of RapCoTT. The finding that sweetness

is more common during the warming trial of the traditional protocol agrees with others (Skinner et al., 2018; Yang et al., 2020). However, there were differences in patterns of tastes reported between warming and cooling trials/phases between the two protocols. These differences are likely a consequence of a combination of three factors. Firstly, sweetness may occur towards the end of the warming trial, and participants are unable to respond soon enough in RapCoTT for their responses to be considered at the cut-off point. However, during the traditional protocol, these participants recall the experience and report them after temperature exposure. In this case, due to its temporal proximity and familiarity (Hettinger et al., 1999), participants may remember perceiving sweetness better and are more likely to report it than other tastes (Daniel & Katz, 2018). Being able to report tastes immediately in RapCoTT is an advantage. Secondly, sourness may persist from cooling into warming (Skinner et al., 2018) in RapCoTT, making it equally likely to occur in either phase. Third, multiple mechanisms may evoke sweetness, causing it to occur at two time/temperature points (Nachtigal & Green, 2018). All these aspects point to the additional potential of RapCoTT to deliver deeper insights into thermal taste responses beyond its phenotyping application.

It is imaginable that minty and spicy could be proxies for cooling and warming sensations, especially as the sensations tend to occur more often in the “appropriate” slopes (Skinner et al., 2018). However, in this study, no significant associations were found between minty or spicy and any of the temperature trials, which could be a consequence of the low frequencies at which participants reported these sensations. Thibodeau et al. (2020) has demonstrated that the training participants receive and the options with which they are presented influence the sensations reported. However, in Thibodeau et al. (2020), participants were not trained on recognising orosensations

other than the prototypical orosensations, nor were they presented as an option to select minty or spicy (only under “other”). With appropriate assessor training, the high-resolution data obtained from RapCoTT will allow this issue to be explored further, which points to the potential of RapCoTT as an approach to investigate thermal taste beyond just its phenotyping.

Truncating the temperature trial can improve the RapCoTT method employed in the current study. Including the final five seconds of the combined protocol did not improve efficiency. However, additional detail gained regarding savoury suggests that information from the final five seconds may be useful in studies aiming to characterise the phenomenon of thermal taste.

#### 2.4.1 A new, more efficient protocol for phenotyping Thermal Taster Status

This study presents considerable evidence to adopt a more rapid and more efficient approach for thermal taste phenotyping, as proposed in Figure 2-4.

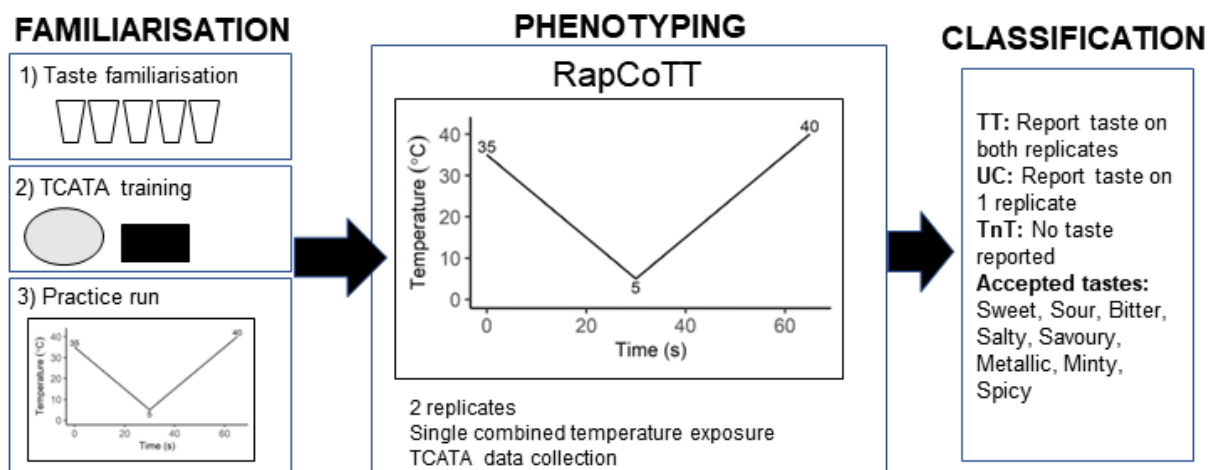


Figure 2-4 The RapCoTT method, showing three training steps namely taste training, TCATA training and a practice run with the protocol. This is followed by two replicates of a temperature exposure (35°C, 5°C, 40°C 1°C/s) with concurrent TCATA data collection.

Firstly, TCATA, when combined with the unfamiliarity of a thermode on the tongue, can be a cognitively demanding task, decreasing the experienced intensity of taste (Liang et al., 2018). A training step can increase the discriminability of TCATA (Jaeger,

Beresford, et al., 2017), and it is therefore recommended that this precedes phenotyping. Training, similar to the approach taken by Velázquez et al. (2020b) using coloured circles that appear and disappear on a screen, is recommended to familiarise participants with how to respond to stimuli in TCATA. Secondly, the unfamiliarity of experiencing taste from temperature and the fact that it is not ‘appropriate to the situation’ may decrease the intensity of the taste signal and prevent it from being reported (Carreiro et al., 2011; Hasbroucq & Guiard, 1992). If participants expect taste, they may experience (and report) taste at a higher intensity (Wilton et al., 2019), explaining the learning effect observed in this study, and by Nachtigal and Green (2020). To control for this learning effect, a training trial should precede the phenotyping, as was performed in this study. Finally, as the task of identifying tastes is not commonly executed outside a sensory lab, a training step with reference solutions is recommended to increase the ability of participants to recognise and verbally identify tastes (Hettinger et al., 1999). The new RapCoTT method, therefore, includes all three steps (Figure 2-4).

Although accurately classifying TT and TnT was a primary aspect of improving phenotyping approaches, the time taken was also of concern. When testing one tongue site, the traditional approach was estimated to take approximately twice as long as RapCoTT. Introducing a new TCATA training step would add approximately 2 minutes, slightly increasing the time for the final RapCoTT method. Nevertheless, this is still considerably shorter than the existing approach.

After the three training steps, participants should complete two replicates of the temperature protocol with TCATA data collection using a system allowing concurrent temperature exposure and data collection. Individuals who report taste on both replicates are considered TT, whereas those who report no taste in either trial are TnT.

Any participant reporting taste on only one trial would be considered UC. Valid orosensations are sweet, sour, bitter, salty, umami, metallic, minty, and spicy, but researchers should review these as the understanding of taste and the mechanisms behind thermal taste continue to evolve.

## **2.5 Limitations**

Although the benefits of RapCoTT have been thoroughly explored in this chapter, the method also has two key limitations. The first is the cognitive demand of TCATA as a data collection method. The second is the inability to control for response bias or demand bias using a minimum threshold as used in traditional (Bajec & Pickering, 2008) thermal taste phenotyping approaches.

In general, temporal sensory data collection methods are criticized for being more cognitively demanding than static data collection methods (Castura, 2018). However, it is important to note that, although valid, due to the differences between thermal taste phenotyping and product characterisation, these criticisms may not necessarily hold true for RapCoTT. When assessors evaluate food or beverage products using traditional sensory methods, there is no restriction on reporting on their experience immediately, as there is with thermal taste phenotyping. In addition, when participants evaluate food or beverage products, the sensory experience is constant, and is multi-modal, comprising of taste, texture, temperature and aroma. In contrast, the temperature ramps involved in thermal taste phenotyping have set times, preventing participants from immediately reporting sensations. In addition, with thermal taste phenotyping, there are large portions of the phenotyping trial where participants experience sensation from maximum one modality (temperature), with the production of taste being a sudden occurrence. Interestingly, experiments

comparing different temporal methods (Visalli et al., 2023) have found TCATA to be most sensitive to changes in sensory experience, making TCATA data collection the most suitable for thermal taste phenotyping, despite the limitations of temporal methods.

A second limitation of RapCoTT is the demand bias (or response bias) that has been demonstrated with temporal methods (Visalli et al., 2023). One way to counteract this would be to introduce an intensity measure, either by collecting time-intensity data or by allowing participants to report this intensity after phenotyping. However, setting a minimum intensity of thermal taste reported as a threshold to counteract response bias is arbitrary, as there is no empirical evidence to support this practice. A better solution would be to use a signal detection theory-based approach (Lawless, 2013), potentially employing novel modelling techniques such as permutation entropy (Huang et al., 2022). Note, however, that some aspects of the RapCoTT were designed to intentionally subvert the researcher-subject power dynamic, to reduce the desire of participants to please the experimenter, reducing experimental “demand”. Specifically, participants held the thermode themselves, and had control over the start of their temperature treatments, restoring their agency in the researcher-participant dynamic.

## **2.6 Conclusion**

This chapter aimed to develop a novel phenotyping method for thermal taste and to validate its efficiency and effectiveness compared to existing approaches. A method with a single temperature exposure and concurrent TCATA data collection was expected to be more effective at classifying TT Status, regardless of scheme. In addition, this approach was expected to be more efficient, phenotyping participants in

less time than existing approaches.

When comparing the new method, RapCoTT, with existing approaches, it agreed with traditional approaches while being more effective and efficient. In addition, no age or gender differences were found between TT and TnT. However, several areas of improvement were also established, including shortening the temperature treatment and including additional training steps.

The work presented in this chapter formed the foundation for the rest of the work in this thesis. RapCoTT not only facilitated the phenotyping of participants, but the detailed TCATA data was used to describe the thermal response in more detail, as presented in Chapter 3. In addition, participants could be grouped according to their thermal taste responses, also presented in Chapter 3. These groupings were used to achieve the objectives of Chapters 4 and 5.

## **Chapter 3 *Divide and conquer*: Exploring underlying mechanisms of thermal taste by clustering and describing thermal taste response collected via RapCoTT using Mixture Hidden Markov Models and functional data analysis**

### **3.1 Introduction**

This chapter aimed to establish clusters of TT based on their thermal taste response and to characterise the response of each of these clusters.

As established in Chapter 1, due to differences in orosensory responsiveness between TT and TnT, differences in food choice are expected between these two groups. However, these differences have yet to be established. One potential cause of the absence of differences is the small sample sizes used in previous research due to inefficient phenotyping methods. As TT have a heterogenous thermal taste response (Skinner et al., 2018), the second potential cause is TT subgroups that differ in their food choice behaviours, obscuring differences between TT and TnT. The development of a more efficient method, RapCoTT, in Chapter 2, addressed the challenge of obtaining larger sample sizes. To address the second challenge, the thermal taste response needed to be described, and the existence of subgroups needed to be established.

#### **3.1.1 Background on thermal taste clusters and mechanisms**

Skinner et al. (2018) established the existence of thermal taster response subgroups by collecting taste intensity data during ten temperature trials, finding four clusters per trial differing in taste reporting pattern. However, their study failed to establish relationships between clusters for the two trials and did not compare cluster memberships. Although Skinner et al. (2018) collected taste identity data by having participants point at tastes perceived, this information was not included in clustering. The taste qualities were also established by having the participants report this after

the intensity data collection, meaning that it lacked the precision of the intensity data. Thibodeau et al. (2020) grouped participants based on the trial (Cooling or Warming) in which they perceived taste, tongue locations eliciting those tastes, and the tastes reported. However, this grouping was researcher-driven, and not algorithmic response-based clustering driven. Thibodeau et al. (2020) found relationships between tastes and temperature trials: Sweet and Warm, Sour and Cool. Additionally, participants reporting Sour were less likely to report Metallic or Sweet, suggesting the involvement of different mechanisms behind thermal taste perception, further underlining Skinner et al. (2018) 's findings. To fully understand groupings of thermal taste response, it is necessary to cluster participants in a manner that considers both time (and, by extension, temperature) and the identity of their responses.

RapCoTT also provided the potential to elucidate the mechanisms behind thermal taste by linking specific taste-temperature relationships with activation temperature ranges of temperature-sensitive taste channels. Talavera et al. (2005) postulated that the TRPM5 channel, activated between 15°C and 35°C, is responsible for Sweet thermal taste and is often referenced as a potential mechanism. Although a relationship exists between Warming and Sweet thermal taste, Green and Nachtigal (2015) suggested two distinct mechanisms for Sweet thermal taste active at different temperatures. The TRPM5 channel is also involved in the perception of Bitter, Umami/Savoury (Zhang et al., 2003), and Metallic (Riera et al., 2009), indicating it is not solely the channel's activity that is triggering thermal taste (Table 3-1).

Besides TRPM5, there are additional mechanisms and channels of interest for thermal taste (Table 3-1). TRPV1 is activated at a higher temperature and responds to capsaicin and salty stimuli (Dias et al., 2013), and is also involved in Metallic sensation (Cao et al., 2014; Luebbert et al., 2010; Riera et al., 2009). The ENaC channel is also involved

Table 3-1 Potential mechanisms and related temperatures for thermal taste and related sensations

Channel	Activation Temperature	Taste	Reference	Details
TRPV1	43° Heating	Salty Metallic	Dias et al. (2013), Cao et al. (2014); Luebbert et al. (2010); Riera et al. (2009).	
TRPV4	24-27°C	Sour	Matsumoto et al. (2019)	
TRPM5	15-35°C Heating	Sweet Bitter Umami/Savoury Metallic	Talavera et al. (2005) Zhang et al. (2003) Zhang et al. (2003) Riera et al. (2009)	
TRPM4	15-35°C Heating	Bitter Sweet Umami/Savoury Sour	Banik et al. (2018); Dutta Banik and Medler (2023)	Component of broadly responsive taste cells
TRPA1	17°C Cooling	Carbonation Sour  Sweet Umami/Savoury	Wang et al. (2010) de la Roche et al. (2013); Garrity (2011); Wang et al. (2011) Matsumoto et al. (2023) Matsumoto et al. (2023)	Through response to weak acids
ENaC	10°C cooling	Salty	Talavera et al. (2008)	

in Salty perception and is activated at 10°C (Askwith et al., 2001), and suggested as a mechanism for Salty thermal taste (Talavera et al., 2008). The TRPA1 channel, activated through noxious cold (Latorre et al., 2009), is part of the nociceptive response of carbonation (Wang et al., 2010) and responds to weak acids (de la Roche et al., 2013; Garrity, 2011; Wang et al., 2011), suggesting that it might contribute to Sour thermal taste. In addition, Matsumoto et al. (2023) recently linked TRPA1 with Umami/Savoury and Sweet taste (Table 3-1).

### 3.1.2 Modelling thermal response data: key concepts

Modelling is the process of extracting a signal from noise (Yokoo et al., 2001). In the context of RapCoTT, the signals of interest are the taste and temperature relationships, and noise would be all factors that impact the accuracy of those taste and temperature relationships, namely distinct latent responses, taste misreporting, delays in taste reporting and inconsistent reporting. In a RapCoTT dataset, unidentified subgroups with different responses could lead to spurious null conclusions (as demonstrated

regarding temperature responses to ethyl caprylate by Powers (1977)) or spurious conclusions which contradict the true underlying pattern (known as Simpson's paradox, see Kievit et al. (2013)).

Untrained individuals often confuse or misreport taste and are more likely to misreport taste when the tastes are fleeting (Kelling & Halpern, 1987), which means that relying on taste reporting (identity) alone could obscure groupings, as different participants who experience the same sensation through the similar mechanisms may report different tastes and be classified differently. For example, if two participants both perceive Savoury/Umami with a common underlying cause, one might falsely report Salty. Participants who are inconsistent in their taste response present a similar problem. Participants are also expected to vary in the speed of their taste (and subsequent verbal) processing, impacting response speeds and leading to temporal noise. A delay in reporting between 400 ms (Kelling & Halpern, 1983; Kelling & Halpern, 1987), and 900 ms (Kelling & Halpern, 1987), is expected, which is impacted by stimulus concentration (and by extension taste intensity), taste duration and taste identity (Kelling & Halpern, 1987).

The appropriate modelling approach could address the above-mentioned sources of noise, but as RapCoTT is a novel technique, there are no existing modelling approaches to achieve signal extraction. Consequently, this required a review and understanding of modelling approaches applicable to RapCoTT data.

#### *3.1.2.1 Typical modelling of TCATA data*

TCATA data is typically collected in the context of product evaluation and used to compare products. In RapCoTT, no products are involved, making product comparison paradigms inappropriate. Conventional TCATA data and RapCoTT data,

however, have a similar structure in the simplified form, where both have start and stop times for each instance of taste reported, with details of the relevant participant, replicate etc. RapCoTT, however, also contains data on the start and stop temperatures. From this form, the data can be expanded into a matrix of ones and zeroes, with ones indicating a taste being reported and zeroes that a taste is not reported (Castura et al., 2016). The expanded RapCoTT matrix differs from the traditional TCATA matrix in that RapCoTT data is sparse in comparison, with many areas containing no reported sensations. In contrast, when participants taste products, the resulting taste sensation is continuous despite the changing nature of taste.

To identify areas of interest in TCATA, the approach initially proposed was to compare the citation proportions of products at each time slice using pairwise comparisons (Castura et al., 2016), an evaluation approach used by the `tempR` package (Castura, 2022) by many built-in functions. However, adjacent time points are likely collinear, calling this approach into question (Meyners & Castura, 2018) and is discouraged (Castura, 2018). Collinearity concerns are particularly relevant for RapCoTT as the treatment is not a single sip or bite but a continuous, changing process. Meyners and Castura (2018) proposed a sensible approach of using bootstrapping and a randomised panel. In this approach, many virtual panels are constructed randomly from the collected data, and the obtained confidence intervals used for product comparison. A similar approach could be used in RapCoTT analysis to determine during which time/temperature portions groups differ in their citation frequency of specific tastes.

### *3.1.2.2 Modelling processes*

It is possible to use citation proportions obtained from TCATA data as a stand-in for time-intensity data (Jaeger et al., 2020; Vidal et al., 2021), and therefore, time-intensity data analysis approaches could be applied to RapCoTT data. Time-intensity

data analysis typically involves extracting a set of parameters, such as the time to the peak, the height of the peak (intensity), the duration of the peak etc. (Cliff & Heymann, 1993), based on theoretical knowledge of the taste response. However, the underlying properties of thermal taste response are relatively unknown, making it disingenuous to superimpose these constructs and the accompanying constraints onto a RapCoTT dataset. Alternatively, curves can be fitted to time-intensity responses, allowing for computational determination of critical parameters (Wendin et al., 2003) although a single function or curve is likely insufficient to describe responses in RapCoTT, as several underlying mechanisms are expected.

Functional data analysis (fda) (Ramsay & Dalzell, 1991) is a computational approach which dynamically fits curves to data based on knots or intervals, enabling the extraction of more complex patterns than possible with linear algebra (Ramsay et al., 2009). Bi and Kuesten (2013) proposed applying fda to time-intensity data, as it can extract curve features, such as the exact time (or times) of a peak, when the sensation is increasing, etc. In the context of RapCoTT, fda would allow extraction of the exact temperatures at which thermal taste phenomena were reported, allowing for more direct inference and the potential to reveal thermal taste mechanisms.

As different underlying thermal taste response patterns are expected, RapCoTT requires an approach incorporating modelling and clustering of responses. As raw RapCoTT data consists of binary (1/0 indicating present/absent) responses, it is categorical, so it is necessary to think beyond the scope of the clustering of functional data. Although not strictly relevant to RapCoTT, the work of Jacques and Preda (2014a) and Leroy et al. (2018) on clustering of functional data informed the approach eventually taken for data analysis. In their review of functional data clustering, Jacques and Preda (2014a) highlight three overarching types of modelling approaches

based on whether the modelling step or the clustering step takes place first or simultaneously. Both these reviews of functional data clustering (Jacques & Preda, 2014a; Leroy et al., 2018) recommend using mixture modelling.

Mixture modelling assumes that a dataset contains a mixture of models and, using an algorithm tests different models and iterates between combinations of model membership of observations. In sensory science, this approach has been used to classify the sensory responses of autistics (Dwyer et al., 2022; Dwyer et al., 2020), to cluster participants based on their response style in sensory tests (Cleaver & Wedel, 2001) and to segment consumer liking based on non-liking factors (Kornelis et al., 2010). Of interest for this study is the use of Gaussian mixture modelling to find different subgroups based on responses to capsaicin and pain (Kringel et al., 2018), and noxious cold (Weyer-Menkhoff et al., 2018), which relate to TRPV1 and TRPA1 channel expression respectively, both of which are potentially involved in thermal taste. However, the most helpful example of mixture modelling is the application of mixture semi-Markov chains to temporal dominance of sensation (TDS) data (Cardot et al., 2019; Lecuelle et al., 2018).

Cardot et al. (2019) demonstrates using a mixture semi-Markov model to find groups with different sequences of dominant attributes. Markov models model stochastic (random) processes by modelling the transition from one state to the next while only considering that present state (Lecuelle et al., 2018). In the context of TDS data, a state would be a dominant sensation, and the Markov model would model how participants move from one (sensation/taste) state to the next. It is tempting to apply the approach of Cardot et al. (2019) to RapCoTT data but, this is not directly possible. TDS data concerns a dominant attribute, making it a univariate categorical time series, with the category response being the dominant sensation. In contrast, RapCoTT/TCATA data

can have many different values at each time point, as more than one taste can be selected concurrently, necessitating an approach suitable for multivariate categorical time series data.

One such approach is multichannel state sequence analysis, typically used in the social sciences to map life trajectories (Gauthier et al., 2010). State sequence analysis considers the categorical temporal sequence as a whole and accounts for all periods and events of interest (Studer & Ritschard, 2016). A RapCoTT trial could be considered a trajectory through distinct stages of a treatment (changing temperature), meaning that state sequence approaches could be applied to RapCoTT data. The initial step is the creation of a difference matrix, which can then be subjected to further analysis, such as finding representative sequences, ANOVA-like analyses, or hierarchical clustering (Gabadinho et al., 2011). For RapCoTT, this represents an initial step leading to further analyses.

Mixture Hidden Markov Models (MHMM) of state sequences (Helske & Helske, 2019) combine the strengths of several of the approaches discussed above with an additional strength particularly relevant to RapCoTT. Hidden Markov Models (HMM) differ from traditional Markov models in that the states are hidden or latent and cannot be directly observed (Langrock et al., 2012; McClintock et al., 2020). In the context of a TDS data example (Lecuelle et al., 2018), the states are the dominant sensations in question, between which participants are moving. Hidden states can be thought of as analogous to principal components in PCA as these states describe a component which the variables themselves cannot describe. For example, in social sciences, HMM can be used to detect aspects of life and employment trajectories that contain measurement error (Pankowska et al., 2018), and in ecology, HMM help account for the periodic absence of the subject of interest (e. g. an animal).

With RapCoTT, using an HMM would enable misreporting of taste (Hettinger et al., 1999; Puputti, Aisala, et al., 2019) to be accounted for. In addition, this approach would account for the fact that thermal taste is not directly observable – instead, secondary taste responses are observed. When combining HMM with mixture modelling in the context of state sequences, the approach considers the total RapCoTT treatment (state sequences), can look at changes in response at specific times/temperatures (Markov models), and compensates for likely misreporting of tastes both in approach to clustering (mixture modelling) and modelling (hidden model). Consequently, a mixture HMM of state sequences was deemed a suitable approach for analysing RapCoTT data.

### *3.1.2.3 Technical details of Mixture Hidden Markov Models*

Table 3-2 shows an overview of the relevant parameters of Mixture Hidden Markov Models, showing those parameters that cannot be estimated and need to be supplied to the model (hyperparameters), and others that can either be supplied or estimated.

#### *3.1.2.3.1 Hyperparameters*

Two parameters, the number of clusters and the number of states per cluster, cannot be estimated by the MHMM and therefore need to be supplied by the researcher. Helske and Helske (2019) recommends that the researcher uses their domain expertise to make sensible parameter choices. However, the examples in Helske and Helske (2019) relate to life trajectories, making it comparatively easy to observe and form sensible hypotheses. The present study, however, is the first of its type, necessitating an alternative approach. A potential approach to estimating the number of clusters is to do exploratory hierarchical clustering of the state sequence data, aiming to establish the number of clusters prior to the Markov modelling process.

Table 3-2 Parameters of Mixture Hidden Markov Models. Abbreviations: Spec = specified. Res = Results

Parameter	Source	Definition	Notes/comments	Example															
Period	Specified	Categorical time unit	Specified during data cleaning	t = 5s															
Nr of clusters	Specified	Number behavioural groups, linked to distinct models.	Exploratory data analysis is recommended to determine k.	k = 3															
Nr of states per cluster	Specified	The expected number of distinct underlying behaviours per cluster	Nr of combinations and models increases with complexity of model (nr of clusters).	A sequence of numbers of the length of the number of clusters, e. g., 3,2,3															
Initial probabilities	Specified/ Estimated	Probabilities of participants being in a specific state at the start of the observed period	If any probabilities are 0, these are taken to be structural zeros and thus, a model constraint	For each cluster, a set of probabilities (one per state), e. g.:															
				<table border="1"> <thead> <tr> <th>Constrained model</th> <th>Unconstrained model (Spec)</th> <th>Unconstrained model (Res)</th> </tr> </thead> <tbody> <tr> <td>Cluster 1: 1,0,0</td> <td>Cluster 1: 0.8, 0.1, 0.1</td> <td>Cluster 1: 0.62, 0.31, 0.07</td> </tr> <tr> <td>Cluster 2: 0,1</td> <td>Cluster 2: 0.1, 0.9</td> <td>Cluster 2: 0.22, 0.78</td> </tr> <tr> <td>Cluster 3: 1,0,0</td> <td>Cluster 3: 0.7, 0.15, 0.15</td> <td>Cluster 3: 0.8, 0.11, 0.09</td> </tr> </tbody> </table>	Constrained model	Unconstrained model (Spec)	Unconstrained model (Res)	Cluster 1: 1,0,0	Cluster 1: 0.8, 0.1, 0.1	Cluster 1: 0.62, 0.31, 0.07	Cluster 2: 0,1	Cluster 2: 0.1, 0.9	Cluster 2: 0.22, 0.78	Cluster 3: 1,0,0	Cluster 3: 0.7, 0.15, 0.15	Cluster 3: 0.8, 0.11, 0.09			
Constrained model	Unconstrained model (Spec)	Unconstrained model (Res)																	
Cluster 1: 1,0,0	Cluster 1: 0.8, 0.1, 0.1	Cluster 1: 0.62, 0.31, 0.07																	
Cluster 2: 0,1	Cluster 2: 0.1, 0.9	Cluster 2: 0.22, 0.78																	
Cluster 3: 1,0,0	Cluster 3: 0.7, 0.15, 0.15	Cluster 3: 0.8, 0.11, 0.09																	
Transition probabilities	Specified/ Estimated	The probabilities of movement between states at the end of each period.	If any probabilities are 0, these are taken to be structural zeros and thus a model constraint.	For each pair of states in a specific cluster for transition both to and from: e. g. Cluster 1															
				<table border="1"> <thead> <tr> <th>Constrained model</th> <th>Unconstrained model (Spec)</th> <th>Unconstrained model (Res)</th> </tr> </thead> <tbody> <tr> <td>To S1 S2 S3</td> <td>To S1 S2 S3</td> <td>To S1 S2 S3</td> </tr> <tr> <td>From S1 0.8 0.1 0.1</td> <td>From S1 0.4 0.3 0.3</td> <td>From S1 0.9 0.1 0.1</td> </tr> <tr> <td>S2 0 0.7 0.3</td> <td>S2 0.3 0.4 0.3</td> <td>S2 0.1 0.7 0.1</td> </tr> <tr> <td>S3 0 0.3 0.7</td> <td>S3 0.4 0.3 0.4</td> <td>S3 0.1 0.2 0.8</td> </tr> </tbody> </table>	Constrained model	Unconstrained model (Spec)	Unconstrained model (Res)	To S1 S2 S3	To S1 S2 S3	To S1 S2 S3	From S1 0.8 0.1 0.1	From S1 0.4 0.3 0.3	From S1 0.9 0.1 0.1	S2 0 0.7 0.3	S2 0.3 0.4 0.3	S2 0.1 0.7 0.1	S3 0 0.3 0.7	S3 0.4 0.3 0.4	S3 0.1 0.2 0.8
				Constrained model	Unconstrained model (Spec)	Unconstrained model (Res)													
To S1 S2 S3	To S1 S2 S3	To S1 S2 S3																	
From S1 0.8 0.1 0.1	From S1 0.4 0.3 0.3	From S1 0.9 0.1 0.1																	
S2 0 0.7 0.3	S2 0.3 0.4 0.3	S2 0.1 0.7 0.1																	
S3 0 0.3 0.7	S3 0.4 0.3 0.4	S3 0.1 0.2 0.8																	
Emission probabilities	Specified/ Estimated	The relationships between each hidden state and taste	Nr of levels in categorical variables determine number of probabilities (e.g. TT, TnT, NC = 3 probabilities/ state)	For each state/taste combination, a pair of probabilities for presence/absence of taste.															
				<table border="1"> <thead> <tr> <th>Constrained model</th> <th>Unconstrained model</th> <th>Unconstrained model (Res)</th> </tr> </thead> <tbody> <tr> <td>Sweet No Sweet</td> <td>Sweet No Sweet</td> <td>Sweet No Sweet</td> </tr> <tr> <td>S1 0 1</td> <td>S1 0.1 0.9</td> <td>S1 0 1</td> </tr> <tr> <td>S2 0.5 0.5</td> <td>S2 0.5 0.5</td> <td>S2 0.81 0.19</td> </tr> <tr> <td>S3 0.5 0.5</td> <td>S3 0.5 0.5</td> <td>S3 0.23 0.77</td> </tr> </tbody> </table>	Constrained model	Unconstrained model	Unconstrained model (Res)	Sweet No Sweet	Sweet No Sweet	Sweet No Sweet	S1 0 1	S1 0.1 0.9	S1 0 1	S2 0.5 0.5	S2 0.5 0.5	S2 0.81 0.19	S3 0.5 0.5	S3 0.5 0.5	S3 0.23 0.77
				Constrained model	Unconstrained model	Unconstrained model (Res)													
Sweet No Sweet	Sweet No Sweet	Sweet No Sweet																	
S1 0 1	S1 0.1 0.9	S1 0 1																	
S2 0.5 0.5	S2 0.5 0.5	S2 0.81 0.19																	
S3 0.5 0.5	S3 0.5 0.5	S3 0.23 0.77																	

Another challenging aspect is establishing reasonable values for the number of states. Traditional model selection approaches would fit several different models and choose the one that has the best fit using a goodness-of-fit metric such as the log-likelihood (Bilder & Loughin, 2014). However, when evaluating models with different numbers of parameters, it is inappropriate, as a model with a higher number of parameters (states) will, by definition, fit better (Vrieze, 2012), unless compensating for the numbers of parameters. The drawback is that the model will be overfitted, modelling noise rather than the underlying phenomena and will not be informative. Alternative model metrics such as the BIC (Bayesian Information Criterion) (Schwarz, 1978) and the AIC (Akaike Information Criterion) (Akaike, 1974) assess model fit whilst penalising models for the number of parameters.

Although using an appropriate metric can improve model selection, as the number of potential models increases, the time it takes to fit these can become prohibitive. Additionally, when there are many different models in question and several are likely to be suitable, minor differences in model assessment values are not meaningful (Schwarz, 1978), especially as goodness-of-fit values are arbitrary, not relating to an underlying physical phenomenon. A more practical approach is to systematically estimate the model parameters using machine learning, a full review of which is out of scope. However, the general principle is to obtain the properties of a good model, rather than the ideal model (Scrucca, 2013), solving the problem of choosing a model based on arbitrary parameters, making this a suitable alternative approach for looking for the correct model.

#### 3.1.2.3.2 Model parameters

The MHMM process estimates three sets of model parameters: initial probabilities, transition probabilities, and emission probabilities. The model can estimate these

fully, or the researcher can supply them as starting values or constraints. In an MHMM, each cluster has a set of parameters (initial probabilities, transition probabilities, emission probabilities), and estimation, constraint, and full searching can be combined in many ways depending on the researcher's insight and other assumptions (Helske, 2018).

In broad terms, on the one extreme, constraining a model allows for simpler, more interpretable models. Constrained models are also computationally less onerous, expediting the model fitting process. However, poor choices in constraints can lead to an underfitted model not fully describing underlying phenomena. The other extreme is estimating all parameters, a lengthy process often resulting in a difficult-to-interpret and unhelpful model, modelling noise rather than signal. A compromise between the two is to give parameters starting values that are unconstrained. However, some constraints can prevent a model from considering impossible scenarios (Helske, 2018).

#### *3.1.2.3.2.1 Emission probabilities*

Emission probabilities are values describing the properties of hidden states and how they relate to the measured variables. For each hidden state, there is a probability (emission probability) for each possible value of each observed variable, which is relatively simple in the case of TCATA/RapCoTT data, with each variable only having two potential values in the form of present/absent (e. g., Sweet/No\_Sweet). However, a categorical variable with three or more levels would have more potential values. The sum of emission probabilities (both starting/supplied and estimated) for each state and variable is 1 – however, when a variable only has two levels (present/absent), it is unnecessary to report both values (Helske, 2018).

In practice, knowledge about the process informs supplied starting values. For example, if participants are unlikely to report taste during the initial period, it is sensible to assign the initial state with a low probability for the specific tastes. Conversely, constraining these values would mean that it is known that a particular state and observation must be linked. For example, to force a particular state not to have any tastes associated with it, the emission probabilities for all tastes can be set to zero.

#### *3.1.2.3.2.2 Initial probabilities*

Initial probabilities relate to the likelihood that the participants are in a specific state at the start of the modelled process. For example, if, at the start of the temperature treatment most participants are not perceiving thermal tastes and therefore start in State 1, this will have a high initial probability. However, if the researcher desires participants to start in a particular state, it can be achieved through constraints, by setting the initial probability for that state to one and for all other states to zero.

#### *3.1.2.3.2.3 Transition probabilities*

Transition probabilities are the probabilities that participants move between states (or remain in a state) at the end of each defined period. For example, with two states, the relevant probabilities will be as follows: State 1 to State 1, State 1 to State 2, State 2 to State 2, and State 2 to State 1. Emission probabilities are most helpful in interpreting relationships between states and the likely process. For example, if participants are unlikely to move from one state to another, these states tend not to follow each other.

Constraining transition probabilities can be helpful if there are relationships between the known states or if there is an apparent initial or terminal state. For example, when observing the lives of animals in an ecosystem, there will be a state where they die, and

it is impossible to move from this state to another state. In the life trajectory example, once participants get married, they will never move back to being single, as they will either become divorced, separated, or widowed. Similarly, it does not make sense for participants to move from single to either of these states. In the context of thermal taste, if it were known that participants only perceived taste once during the temperature treatment, it would be useful to constrain the model to contain three sequential states (“start”, “taste” and “finish”).

In the present study, a method sensitive to temporal shifts would have less tolerance for participants reporting taste earlier or later, which is not of concern as the period length (5s) is much longer than the expected temporal error (Kelling & Halpern, 1987). The trimming step in data pre-treatment (section 3.2.2.1) also addressed the expected error. Inconsistent taste reporting is expected (Hettinger et al., 1999; Puputti, Aisala, et al., 2019), which requires a distance metric that is more sensitive to difference than similarity.

### **3.1.3 Aims**

This chapter had two main objectives. The first objective was to establish clusters of thermal taste responses, to group participants based on their responses. Based on previous findings (Skinner et al., 2018), different taste responses clusters were expected, likely three or four groups. In addition, some participants were expected to be inconsistent in their thermal taste responses, as previously demonstrated (Skinner et al., 2018).

The second objective was to describe thermal taste responses by establishing links between temperature ramps (cooling/warming) and tastes and specific temperatures and tastes. Although previous research (Skinner et al., 2018; Thibodeau et al., 2020)

has linked temperature ramps with specific tastes, it is expected that analysis of RapCoTT data will produce more nuanced patterns of taste responses than previously found. In addition, it is expected that it is possible to link specific thermal tastes with their underlying mechanisms. Skinner et al. (2018) previously presented evidence of a connection between TRPM5 and Sweet thermal taste, but, likely, some of the mechanisms presented in section 3.1.1 could also be linked to specific tastes.

Therefore, from these objectives, this chapter aimed to cluster (divide) and fully describe (conquer) thermal taste responses. Specifically, it aimed to:

- 1) evaluate overall relationships between reported tastes and temperature ramps and co-reported tastes.
- 2) cluster the thermal taste responses of participants using a model-based clustering approach.
- 3) describe the differences in thermal taste response between the different clusters.
- 4) determine relationships between temperature and taste response in thermal taste and establish potential mechanisms.
- 5) establish clusters of TT for use in follow-on research (Chapter 4 and Chapter 5)

### **3.2 Materials and Methods**

To understand underlying groupings in thermal taste response, taste responses from three studies were combined and clustered using a mixture modelling approach.

Data were combined from studies with ethics approval numbers SOA 19/26, SOA 19/50, and SOA 20/40 (see Appendix A and B). All data were collected using the RapCoTT protocol using TCATA although there were subtle differences between the

studies, as outlined in section 3.2.1.

### 3.2.1 Participants and protocols

**Study 1:** The first cohort ( $n = 143$ ) were those from the method validation study (Chapter 2). Participants had therefore completed both protocols (RapCoTT and the traditional).

**Study 2:** A second researcher, trained in the use of RapCoTT by the author, recruited a second cohort ( $n = 36$ ) as part of a different project. Only minor differences existed in data collection compared to cohort 1: participants only completed RapCoTT and had been familiarised with taste solutions to recognise the different tastes.

**Study 3:** The author phenotyped a third cohort ( $n = 65$ ) using RapCoTT for a more extensive study on thermal taste reported in Chapters 4 and 5. This cohort was also familiarised with different tastes using taste solutions. The cohort also completed an additional activity with the thermode on the tip of their finger whilst only responding to cooling and warming sensations as a practice run for the TCATA aspect of RapCoTT, following comments in Chapter 2 that an additional practice run would aid participants with the TCATA task and would increase phenotyping efficiency.

As is typical with TCATA data, the resulting dataset ( $n = 244$ ) was in the form of start and stop times for each instance of sensation for each participant and replicate. Each start and stop time had a corresponding temperature. When participants indicated “Other” to signify that they experienced a sensation not listed, this was noted and incorporated into their data.

### 3.2.2 Data analysis

Data were analysed using R (version 4.1) (R. Core Team, 2021), using the `tidyverse`

data manipulation framework (Wickham, 2017). The `TramineR` package (Gabadinho et al., 2011) was used to perform sequence analysis and the associated clustering using the package `cluster` (Maechler et al., 2021). Model parameters were optimised using the genetic algorithm infrastructure in the `GA` package (Scrucca, 2013), in conjunction with mixture hidden Markov model functionality included in the `seqHMM` package (Helske & Helske, 2019). The tools included in `tempR` (Castura, 2022) facilitated TCATA data analysis investigating differences in citation proportions. The `fda` package (Ramsay et al., 2021) was used for functional data analysis.

### *3.2.2.1 Data pre-treatment*

Data consisted of start and stop times (1/100s or cs) and temperatures (°C) for each TCATA response for different temperature sensations (Cooling/Warming) and tastes for each participant x replicate combination. The RapCoTT system dynamically delivered temperature (the parameter of interest) according to the preset time parameters for temperature modulation during the trial. However, as most temperature values occurred twice during the trial (in Cooling and Warming), it was impractical to use these to model the entire trial, and time values were used as a compromise. However, temperature was the parameter being controlled, meaning there were small shifts in the temperature delivery based on external factors. One key cause of such shifts was that participants had the agency to trigger the temperature trials themselves, resulting in large differences in trial start times. In addition, external factors such as participant physiology sometimes caused momentary shifts in temperature delivery, causing the thermode to self-correct.

A dummy time dataset was created by splitting the raw data into three sets according to the taste or sensation phase: Cooling, Warming, or end, resulting in multiple

observations for tastes that started in one phase and ended in another. For example, if a taste started in the cooling phase and ended in the warming phase, the taste was split across two responses. The first response had the first phase's observed start and end temperatures (5°C). The second response had the start temperature of the second phase as its start temperature and the observed end temperature as its end temperature. Similarly, if a participant clicked a taste in the first cooling phase and unclicked in the second cooling phase, it was split into three responses: one with the start and the end temperature of the first phase, the second with the start and end temperature of the second phase and the third with the start temperature of the final phase and the observed unclick temperature as its end. Responses continuing beyond the end of the trial had the end temperature of the trial (35°C).

An additional transformation was required to make the data compatible with state sequence analysis and the associated modelling (Mixture Hidden Markov Models of state sequences). Data were binned into a series of 13 periods, each corresponding to 5°C (5s), using a threshold of 5s. Response segments smaller than 1s were trimmed to compensate for the expected error of up to 1s in TCATA data (Kelling & Halpern, 1987); for example, if a participant clicked a taste at 4.97s and unclicked it at 7s, the taste response was included in the second 5s bin but not the first. As a final pre-treatment step, each taste variable was converted into a data channel by recoding the presence/absence values into descriptive categorical values (e. g., Bitter and No\_Bitter), as necessitated by the sequence analysis algorithm.

### *3.2.2.2 Determining relationships between tastes and temperature ramps*

An approach similar to Thibodeau et al. (2020) was used to find associations between tastes and temperature ramps, to facilitate comparison with their findings. As the present study did not include consistency in the definition of thermal taste

(participants did not need to report the same tastes or report taste in the same trial portion), there were several ways of approaching this comparison. The initial comparison (approach 1) applied the traditional definition of TT (reporting the same taste consistently during the same trial portion)- enabling direct comparison with the results in Thibodeau et al. (2020). Next (approach 2), considered each participant as an experimental unit; all tastes reported by a participant were pooled into a singular observation – highlighting which tastes tended to be reported by the same participants. Finally (approach 3) considered each participant x replicate (response) as an experimental unit, and raw responses were used - highlighting which tastes were typically reported together.

A 2x2 contingency table was created for each taste x taste and taste x temperature combination, and the exact Fishers 2x2 test was applied (Fay, 2010). The resulting  $p$ -values were adjusted for the false discovery rate (fdr adjustment) according to the Benjamini Hochberg procedure (Benjamini & Hochberg, 1995), enabling multiple comparisons. Significant associations ( $p < 0.05$  after fdr adjustment) in the form of odds ratios (OR) were visualised.

### *3.2.2.3 Specifying Mixture Hidden Markov Models*

Data were simultaneously clustered and modelled using a Mixture Hidden Markov Model (MHMM) of state sequences (Helske & Helske, 2019). This approach required two parameters: the number of clusters and states per cluster. Without pre-existing knowledge, these parameters were determined through exploratory modelling.

#### *3.2.2.3.1 Determining the number of clusters (exploratory hierarchical clustering)*

The number of clusters was established through exploratory state-sequence analysis and clustering by creating a distance matrix, which was subsequently clustered using

an agglomerative hierarchical clustering approach on the matrix (Figure 3-1).

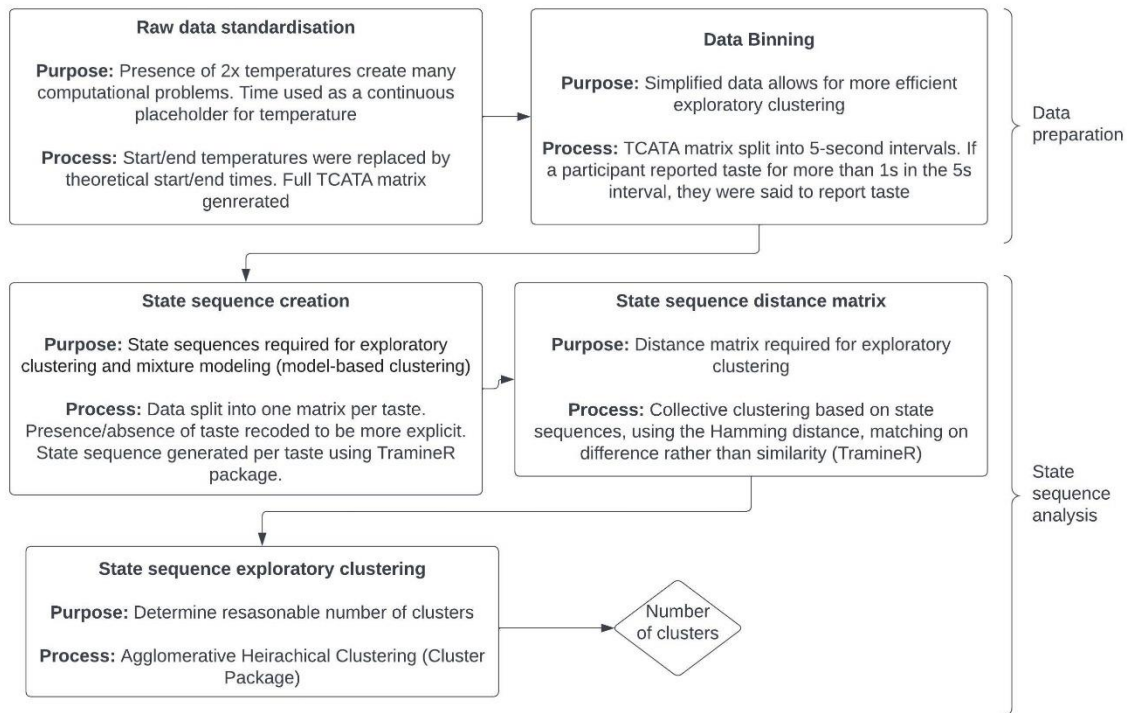


Figure 3-1 Overview of the steps taken to determine the number of clusters in thermal taste responses, showing the data processing steps (standardisation, binning and creation of state sequences) and the clustering steps (creation of a distance matrix and hierarchal clustering)

During distance matrix creation, the Hamming distance (Hamming, 1950) was used as a distance metric, due to its suitability for state sequences that are of the same length, its relative sensitivity to temporal shifts (Studer & Ritschard, 2016), and its relative weighting of difference versus similarity (difference weighted more heavily than similarity) (Studer & Ritschard, 2016). The Hopkins statistic was used to evaluate clustering tendency (Kassambara, 2017), ranging from 0 to 1, with a value close to 1 indicating the presence of a clustering structure.

Hierarchical clustering was performed on the distance matrix using Ward's method through the functions available in the `cluster` package (Maechler et al., 2021). The optimal number of clusters was determined using the silhouette (Rousseeuw, 1987) and Dunn methods (Dunn, 1974) through the functionality in the `NbClust` package

(Charrad et al., 2014). The silhouette method measures how well each observation fits into its cluster and recommends a value for k, which maximises the average “fit” value (Kassambara, 2017). The Dunn method maximises the ratio of minimum separation (distance between objects between clusters) and maximum diameter (distance between observations in the same cluster), favouring a value of k where clusters are well separated.

### 3.2.2.3.2 MHMM constraints

When fitting the MHMM, starting values for emission, initial and transition probabilities were provided to the model (Table 3-3), enabling it to reach a solution more quickly. Note that these constraints were used when determining the optimal number of states per cluster (section 3.2.2.3.3) and finally fitting the model (section 3.2.2.4.1).

*Table 3-3 Parameters used to input into the MHMM, showing an example for a four-state cluster. Note that the initial probabilities indicate a constrained model, which “forces” the process to start in State 1. The emission probabilities are unconstrained, but the starting values for State 1 are set to indicate that it is expected that State 1 would be dominated by “no sensation”. Transition probabilities were set to 1/n, which means that these had to be modelled from scratch. Each state refers to a distinct experience, linked to a likely taste mechanism.*

<b>Initial probabilities</b>				
	State 1	State 2	State 3	State 4
	1	0	0	0

<b>Emission probabilities (example)</b>									
	Sweet	No Sweet	Bitter	No Bitter	Sour	No Sour	...	Metallic	No Metallic
State 1	0.1	0.9	0.1	0.9	0.1	0.9		0.1	0.9
State 2	0.5	0.5	0.5	0.5	0.5	0.5		0.5	0.5
State 3	0.5	0.5	0.5	0.5	0.5	0.5		0.5	0.5
State 4	0.5	0.5	0.5	0.5	0.5	0.5		0.5	0.5

<b>Transition probabilities</b>				
	State 1	State 2	State 3	State 4
State 1	0.25	0.25	0.25	0.25
State 2	0.25	0.25	0.25	0.25
State 3	0.25	0.25	0.25	0.25
State 4	0.25	0.25	0.25	0.25

The model was constrained in that all participants started the trial in State 1, in which taste was unlikely to be reported. To force all participants to start the trial in State 1, the initial probability of State 1 was set to 1, and the remainder of the states to 0 (Table 3-3). To define State 1 as a state in which taste was unlikely to be reported, the starting

emission probability for reporting each taste” (e. g. Sweet) was set to 0.1, and the complementary emission probabilities (e. g. No\_Sweet) were set to 0.9. Note that these emission probabilities are simply starting values to guide the model in a specific direction and are not considered a constraint.

The remainder of the values (starting emission and transition probabilities) were unconstrained and were therefore set to chance values – emission probabilities were all set to 0.5, and transition probabilities were set to  $1/n$ , where  $n$  equals the number of states for that given cluster. For example, in a four-state model, the transition probabilities equal  $1/4$  or 0.25.

#### 3.2.2.3.3 Determining the number of states per cluster

A second hyperparameter required by the MHMM was the number of hidden states per cluster. Hidden states (or simply states) can be considered different experiences participants have during RapCoTT, linked to underlying mechanisms. As participants were expected to cluster based on differences in underlying thermal taste mechanisms, the assumption that all clusters would experience (or move through) the same number of states could not be made.

This problem was solved using a hyperparameter tuning approach (Figure 3-2), which tested model performance with different potential sets of parameters. A genetic algorithm (Scrucca, 2013) found the set of numbers of states per cluster, which allowed for the fitting of an MHMM that had the best fit but was also the least complex (lowest number of total states) by maximising the negative Bayesian Information Criterion (BIC) (i.e. minimising the BIC) (Figure 3-2). The genetic algorithm worked iteratively by partially fitting 20 random candidate models, identifying the five best models (lowest BIC), and including those in the subsequent iteration. Once the model did not

improve (the BIC did not decrease) after 20 consecutive iterations, the genetic algorithm stopped.

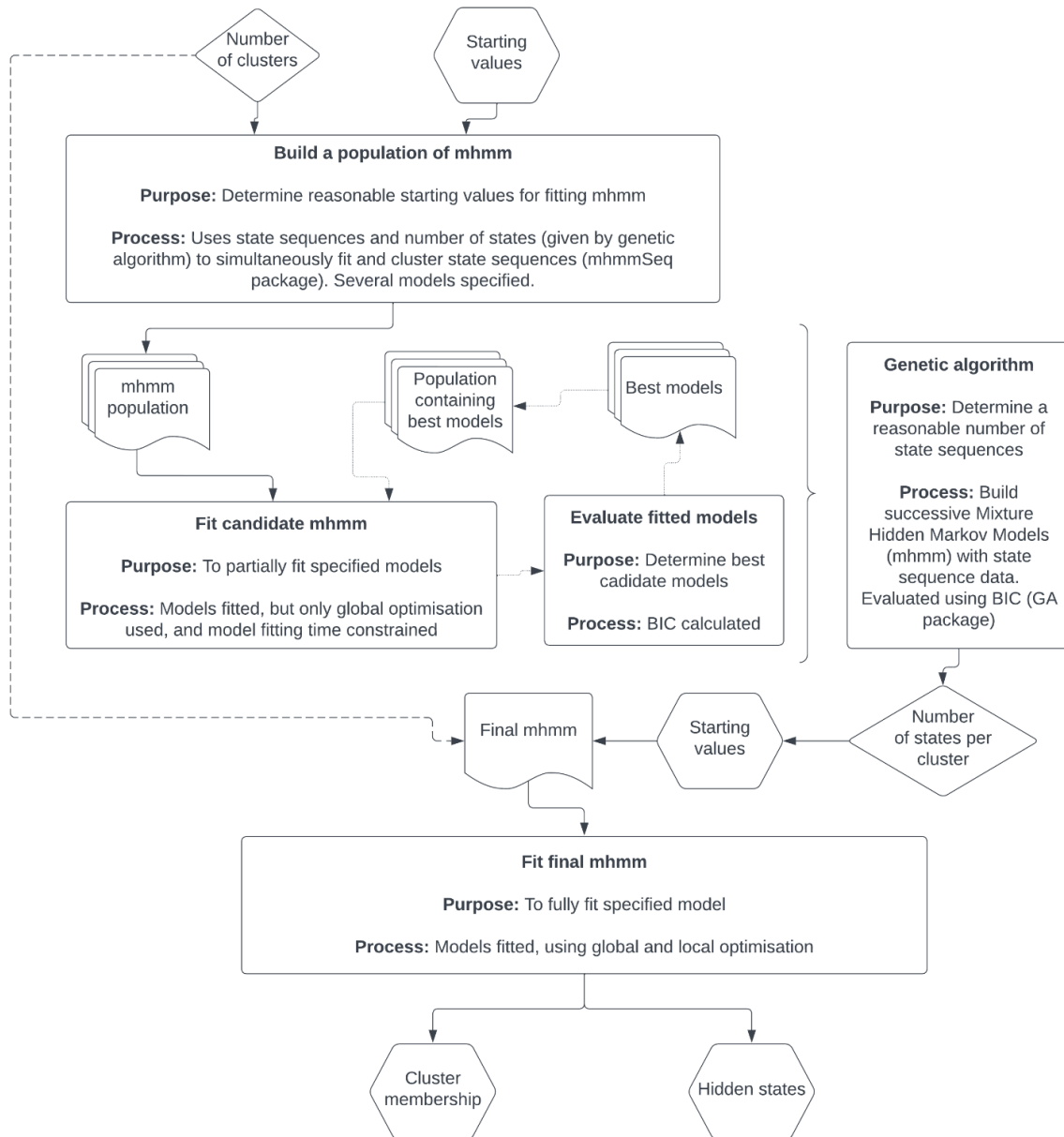


Figure 3-2 Modelling process, which shows how the final MHMM was obtained. The left shows the iterative process using the genetic algorithm, which determines the number of states per cluster. On the right, the process of specifying (building) and fitting the model is shown, resulting in cluster membership and hidden states

The genetic algorithm tested hidden states per cluster ranging from three to eight. Three was chosen as the bottom of the range as the thermal taste process should have a beginning, middle and end. Eight was used as a top value as there are six potential

tastes, allowing for a “no taste” state and a mixed state in the most complex possible model.

#### *3.2.2.4 Fitting and evaluating MHMM*

##### *3.2.2.4.1 Final model fitting*

The final MHMM was fitted using the number of clusters obtained from hierarchical clustering (section 3.2.2.3.1), with the constraints given in section 3.2.2.3.2 and the number of states per cluster obtained from the genetic algorithm (section 3.2.2.3.3). In this final step, infinite time was allowed for global optimisation, and global optimisation was followed by local optimisation, fully fitting the model. A successful model was one where the BIC obtained was smaller than the smallest BIC obtained by the genetic algorithm, indicating a better fit as the complexity (number of states per cluster) remained constant. In addition, the model was considered successful if the mean cluster probabilities of the most probable cluster were close to 1 for all clusters.

The final model obtained was described by visualising the final emission probabilities of each state in each cluster, providing insight into which cluster or clusters tended to contain which taste and whether those tastes were reported in isolation or in conjunction with another taste.

##### *3.2.2.4.2 Determining individual cluster membership*

Each participant participated in two replicates of RapCoTT, the replicates they completed each represented two different responses or observations, which were clustered independently. Cohen’s Kappa (Cohen, 1968) was used to evaluate whether participants fell into the same cluster, regardless of replicate.

Two approaches were assessed to determine the best cluster assignment for

participants for whom the cluster assignment was unclear (different cluster assignments across replicates). Initially, the mean cluster probability per participant was found across both replicates and participants were assigned to the cluster for which they had the highest affinity (mean probability). Next, participants were assigned to the cluster with the highest overall affinity across both replicates. The effectiveness of the two approaches was evaluated by visualising the distribution of mean cluster probabilities.

#### *3.2.2.5 Comparing cluster citation proportions*

Significant differences in citation proportions between clusters were established using a bootstrapping process (Castura et al., 2022; Meyners & Castura, 2018), which produced an upper and lower confidence limit for the cluster and taste in question, and for the remaining combined clusters. Citation proportions were considered significantly different at any point where, for a specific taste, the lower confidence limit exceeded the upper limit of the remaining (combined) clusters. Differences in citation proportions from the other clusters were plotted, highlighting points where each cluster differed significantly. All underlying calculations were performed using the `tempR` package (Castura, 2022), using `ggplot2` (Wickham, 2016), for manual visualisation.

#### *3.2.2.6 Functional data analysis*

Functional data analysis was used to extract patterns of times/temperature and reported taste, focussing on cluster and taste combinations that were significant in the TCATA analysis. The approach in Bi and Kuesten (2013) was applied, with changes to two hyperparameters, namely roughness penalty ( $\lambda$ ) and breaks (knots).

A hyperparameter tuning approach was used to determine the ideal value for  $\lambda$ .

For each dataset (taste x cluster), lambda values between  $10^{-6}$  and  $10^2$  were tested to determine the value lambda, which minimises the generalised cross-validation (GCV), thereby optimising the balance between smoothness and fit (Bi, 2015; Ramsay et al., 2009). The range to test was chosen based on the examples in Bi (2015) and Ramsay et al. (2009), and the results of all tests were inspected visually to confirm that the values included a minimum.

The breaks (or knots) – which are intervals in model fitting – were set at 5s intervals, in keeping with the approach employed for the MHMM. The approach of Bi and Kuesten (2013) of using time slices (1s) as breaks was initially tested. However, this resulted in models which were overfitted and difficult to interpret.

Fda resulted in functional data objects used to visualise relevant taste responses and obtain relevant curve parameters. The curve properties, for example, maximum points and points of sharpest increase/decrease, were extracted to provide insight into potential mechanisms for thermal taste. These curves represent citation frequency, but are also used as a proxy for intensity (Jaeger et al., 2020; Vidal et al., 2021).

### **3.3 Results**

When considering the two responses from each participant independently, most responses (41%) contained only one reported taste. When considering both responses from each participant together, a lower number of participants reported only one taste (34%) (Table 3-4). However, when considering all responses, a sizeable proportion (66%) of participants reported more than one taste. Five participants (3%) reported all six tastes, and there were three instances (responses) (1%) where the participant reported all six tastes in the same trial.

Table 3-4 Summary of the number of tastes reported per participant and contained per response. All the tastes reported in both replicates are considered when participants are counted. Each replicate  $x$  participant is treated as an experimental unit when observations are counted.

Number of tastes	Participants (replicates combined)		Individual responses (replicates treated separately)	
	Count	Percentage	Count	Percentage
1	59	34%	133	41%
2	42	24%	90	28%
3	33	19%	58	18%
4	23	13%	32	10%
5	13	7%	5	2%
6	5	3%	3	1%

### 3.3.1 Relationships between tastes and phases and intra-taste relationships

When only considering duplicate responses (i.e. the traditional TT), Cool TT were more likely to be Bitter TT (OR = 5.14,  $p = .006$ ) (Figure 3-3). Metallic was negatively associated with Sweet (OR = 0.26,  $p = .008$ ) and Sour (OR = 0.29,  $p = .006$ ); Metallic TT were three times (1/0.3) less likely to also be Sweet TT or Sour TT.

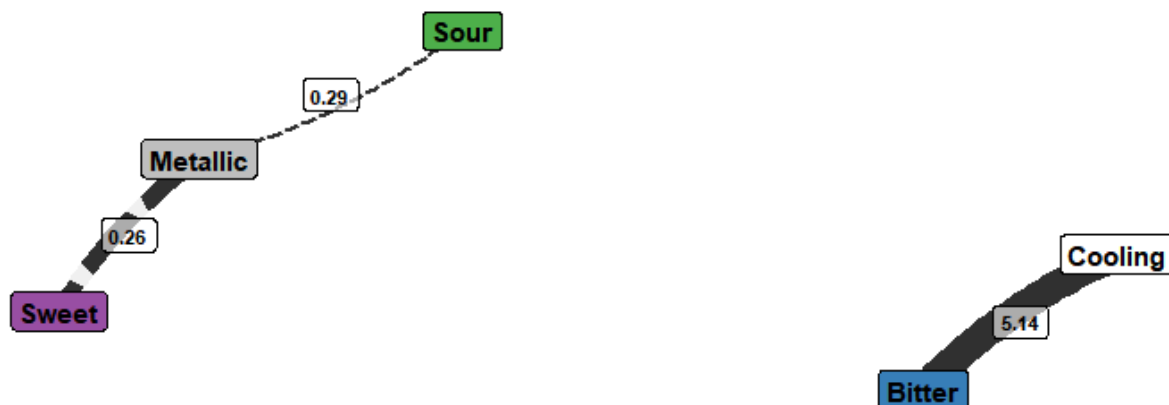


Figure 3-3 Relationships between tastes and trial portions when only duplicate tastes are considered for Thermal Taster classification, labelled with the odds ratio (OR). Line thickness is proportional to the strength of association in the form of odds ratio or inverse of odds ratio in the case of negative association ( $OR < 1$ ). Negative associations ( $OR < 1$ ) are indicated with a dashed line, and positive associations ( $OR > 1$ ) with a solid line. Light grey lines are less significant. ( $0.01 < p < 0.05$ ) and black lines are more significant ( $p < 0.01$ ). Figure after Thibodeau et al. (2020).

When considering pooled responses across participants, participants who reported taste during the cooling trial portion were four times more likely to report Metallic thermal taste ( $p = .009$ , OR = 3.95) (Figure 3-4 a). Note that this does not necessarily indicate that Metallic was reported more frequently during the cooling portion. Participants who reported Savoury thermal taste were almost five times more likely to

also report Salty thermal taste ( $p < 0.001$ , OR = 4.61) and three times more likely to report Sour ( $p = .005$ , OR = 2.96), although there was no significant relationship between Salty and Sour. Participants who reported Sour were significantly more likely to also report Bitter ( $p = .002$ , OR = 2.67), and those reporting Bitter were significantly more likely to also report Sweet ( $p = .001$ , OR = 3.23).

a)



b)

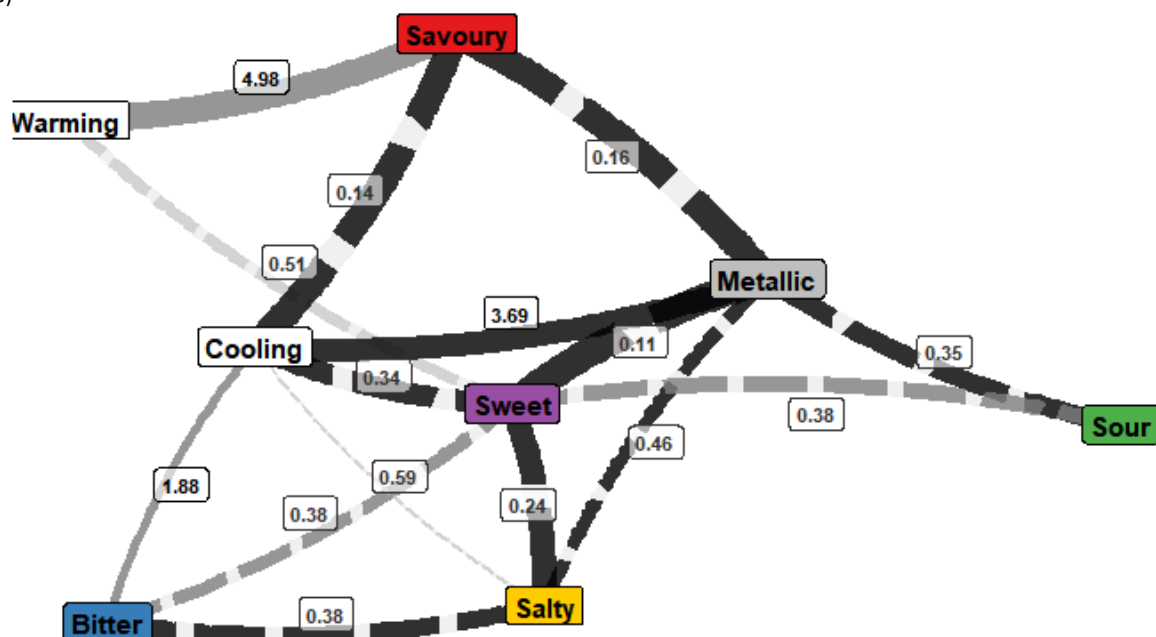


Figure 3-4 Relationships between tastes and trial portions when a) responses are pooled per participant b) responses are considered individually. Thickness of lines are proportional to the strength of association in the form of odds ratio or inverse of odds ratio in the case of negative association ( $OR < 1$ ). Negative associations are indicated with a dashed line, and positive associations with a solid line. Light grey lines are less significant ( $0.01 < p < 0.05$ ) and black lines are more significant ( $p < 0.01$ ). Figure after Thibodeau et al. (2020).

When considering each response (participant x replicate) separately, Metallic was reported approximately four times more often during the cooling trial portion than

would be expected by chance ( $p < 0.001$ , OR = 3.69) (Figure 3-4 b). Savoury was reported during the warming trial portion, as it was strongly positively associated with the warming trial ( $p = .013$ , OR = 4.98) and negatively associated with the cooling trial portion ( $p < 0.001$ , OR = 0.14). In addition, Sweet was negatively associated with both the Cooling ( $p < 0.001$ , OR = 0.34) and Warming ( $p = .043$ , OR = 0.51) trial portions. Salty was negatively associated with the cooling trial portion ( $p = .027$ , OR = 0.59), and Bitter was positively associated with the cooling trial portion ( $p = .010$ , OR = 1.88).

As participants tended to not report Savoury and Metallic in the same trial portion, they also tended not to be reported together ( $p < .001$  OR = 0.14). Metallic also tended not to be reported in the same trial portion as Sweet ( $p < 0.001$ , OR = 0.11) or Salty ( $p = .002$ , OR = 0.46). Bitter was also negatively associated with Sweet ( $p = .013$ , OR = 0.38) and Salty ( $p = .003$ , OR = 0.38). Despite the negative associations of both tastes with Metallic and Bitter, Sweet and Salty tended not to be reported together ( $p = .002$ , OR = 0.24) (Figure 3-4 b).

### **3.3.2 Clustering of thermal taste response**

#### *3.3.2.1 Exploratory clustering, model specification and effectiveness of clustering*

The Hopkins statistic of the dataset was 0.85, indicating the presence of a cluster structure. The silhouette and Dunn methods indicated three as the optimal number of clusters. The clusters contained 121, 51 and 140 responses, respectively (Figure 3-5).

The genetic algorithm terminated after 41 iterations, as the BIC (10964.15) did not decrease in the 20 preceding generations. The suggested model was one where two clusters contained four hidden states, and a third cluster contained five hidden states. These parameters were included in the final MHMM.

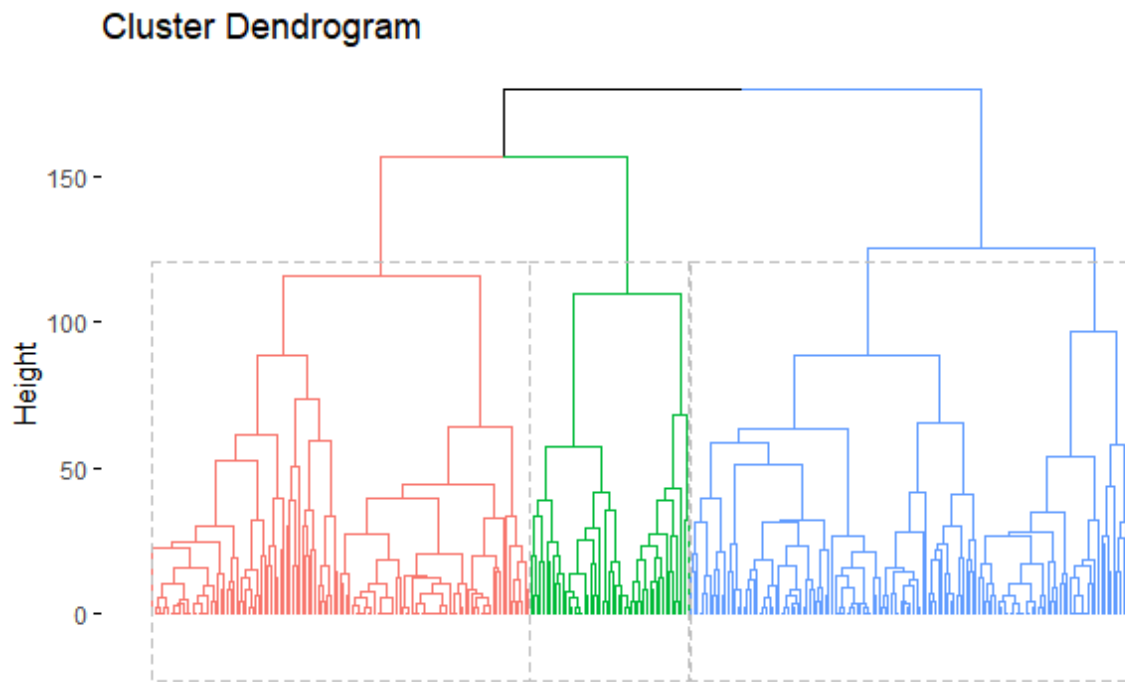


Figure 3-5 Dendrogram obtained from the exploratory hierarchical clustering of the state sequence data, using the dissimilarity matrix obtained using Hamming distances, highlighting the three clusters.

The MHMM's final BIC was 10755.68, indicating an improvement in fit from the genetic algorithm. Cluster 1 (4 hidden states) contained 133 responses, Cluster 2 (4 hidden states) contained 77 responses and Cluster 3 (5 hidden states) contained 102. The model produced a probability for each cluster and response. When calculating and comparing the mean cluster probabilities with the most probable cluster (cluster with the highest probability for a specific response), all values fell close to 1 (Cluster 1 = 0.92, Cluster 2 = 0.97, Cluster 3 = 0.97), indicating a successful model.

#### 3.3.2.1.1 Individual cluster membership

Overall, both observations (responses) of most participants (91 or 64%) were included in the same cluster, but 51 participants were not classified consistently (Table 3-5). Cohen's K comparing the classification of participants between two replicates was 0.45, indicating moderate agreement (Kwiecien et al., 2011).

The initial approach to the final cluster assignment (individual mean cluster probabilities) yielded a split distribution (Figure 3-6), whereas assigning participants

Table 3-5 Contingency table of classification of participants in the same cluster across the two different replicates

Replicate 1	Replicate 2			Total	% correct class
	Cluster 1	Cluster 2	Cluster 3		
Cluster 1	41	9	7	57	72%
Cluster 2	11	18	9	38	47%
Cluster 3	8	7	32	47	68%
<b>Total</b>	60	34	48		
<b>% correct class</b>	68%	53%	67%		

to the cluster for which they had the highest affinity yielded a distribution where most participants had a cluster probability over 0.8, indicating a higher affinity to one specific cluster. Therefore, the second approach to cluster assignment was adopted.

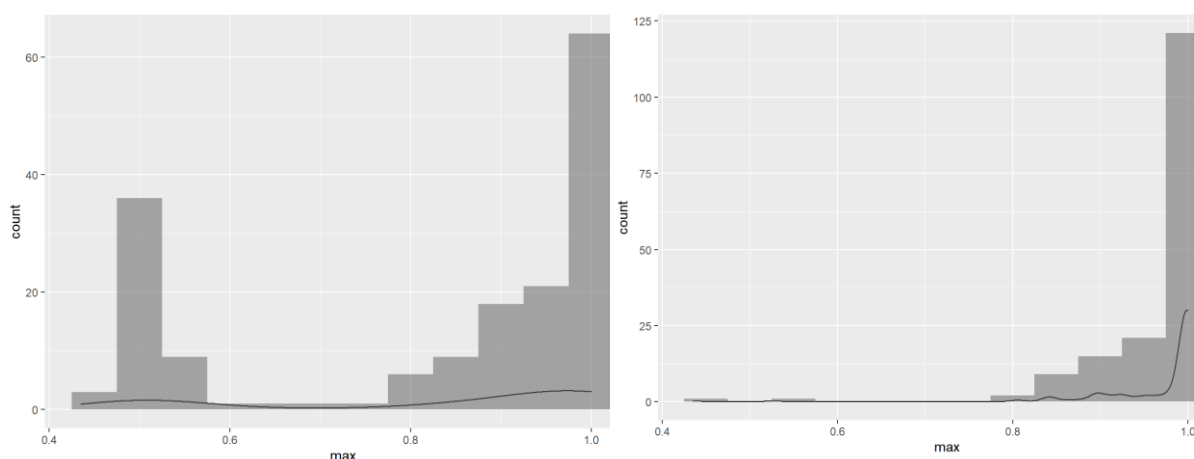


Figure 3-6 Distributions of cluster probabilities of individuals considering two approaches to final clustering; a) Assignment to cluster with highest mean across two replicates b) Assignment to highest affinity across two replicates

### 3.3.2.2 Cluster, State (mechanism) and taste relationships

Emission probabilities (Figure 3-7) described relationships between reported tastes and hidden states and, by extension, potential mechanisms. Emission probabilities were interpreted as the probability that the relevant taste was included in the responses of that state. An emission probability of 1 means that the relevant taste was included in all responses in the hidden state in question. As expected from the modelling parameters, in State 1, emission probabilities were close to 0, regardless of cluster or taste (Figure 3-7).

Cluster 1 mainly included responses where participants reported Metallic and Salty.

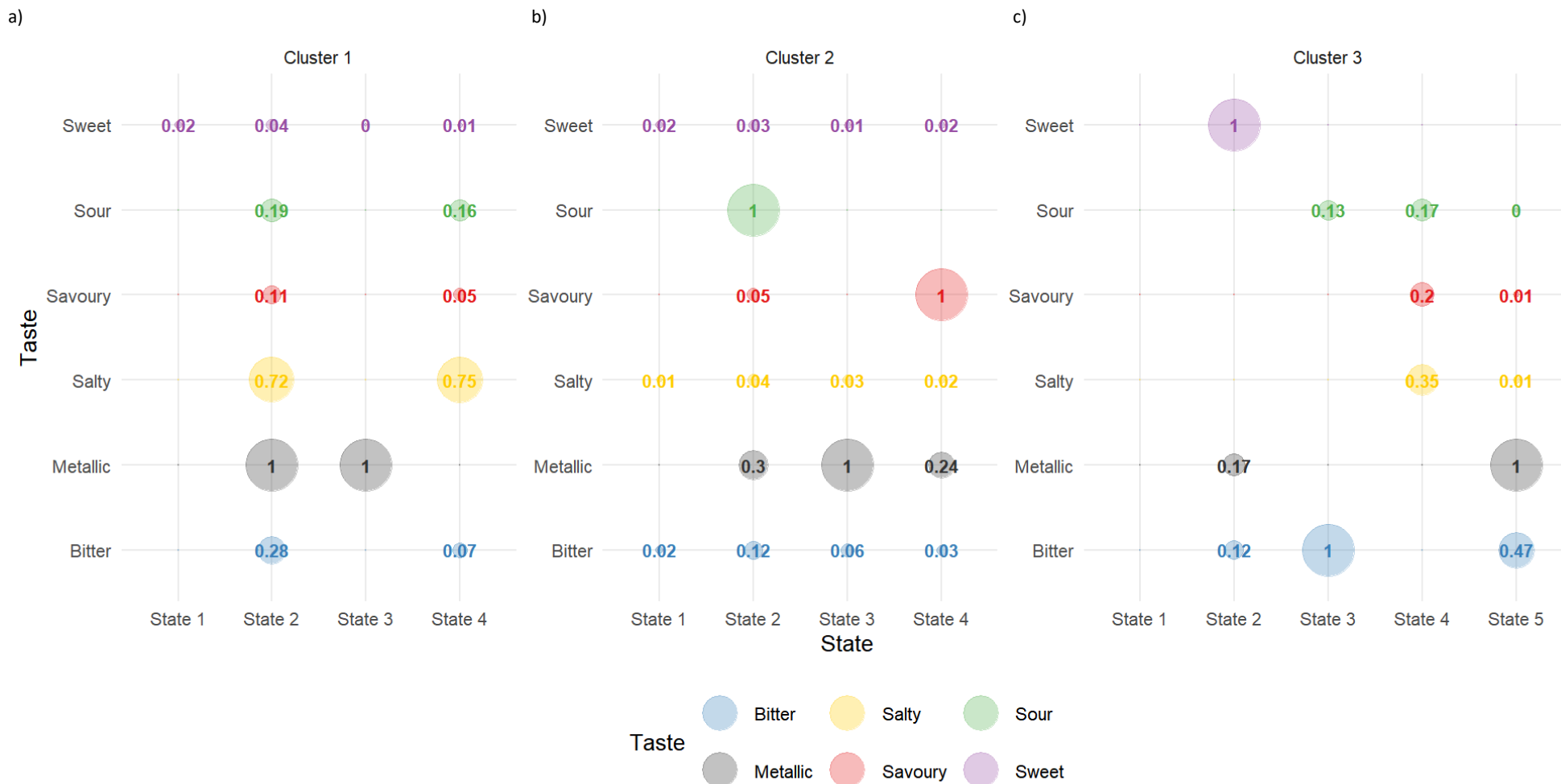


Figure 3-7 Emission probabilities of each of the hidden states for a) Cluster 1, b) Cluster 2 and c) Cluster 3. The size of the bubbles is proportional to the emission probabilities and are also coloured accordingly. A taste was considered to characterise a state if the emission probability exceeded 0.7. Note that the numbers do not add up to 1, as there is an accompanying (complementary) emission probability for a taste not being reported.

State 3, Cluster 1 (S3C1) was the only hidden state across all responses where participants reported Metallic in isolation (Metallic  $p = 1.0$ , and  $p = 0$  for the remaining tastes). During State 2 (S2C1), both Metallic ( $p = 1.0$ ) and Salty ( $p = .72$ ) characterised the responses in Cluster 1 – the only state across clusters with two dominant states. As Metallic was always reported during S2C1, when Salty was reported during S2C1, it was reported alongside Metallic. Bitter ( $p = .28$ ), Sour ( $p = .19$ ) and Savoury ( $p = .11$ ) were reported at lower frequencies during S2C1. Salty was reported most frequently ( $p = .75$ ) during State 4 (S4C1), and Metallic was not reported ( $p = 0$ ). During this state, Sour was reported infrequently ( $p = .16$ ), with low residual frequencies for the remaining tastes.

Cluster 2 was described by a state where Metallic was always reported (State 3 i.e. S3C2,  $p = 1.0$ ). In addition, the responses in Cluster 2 were described by states in which Sour (State 2 i.e. S2C2,  $p = 1.0$ ) and Savoury (State 4 i.e. S4C2,  $p = 1.0$ ) were respectively reported. However, participants sometimes reported Metallic alongside Sour in S2C2 ( $p = .30$ ) and Savoury in S4C2 ( $p = .24$ ). Also, note that Salty and Bitter were reported at low frequencies during all four hidden states.

Five hidden states described the responses in Cluster 3. In three of these states, all responses contained Sweet (State 2 S2C3,  $p = 1.0$ ), Bitter (State 3 S3C3,  $p = 1.0$ ) or Metallic (State 5 S5C3,  $p = 1.0$ ), respectively. However, in State 5, Bitter ( $p = .48$ ) was reported alongside Metallic in almost 50% of cases. No tastes dominated State 4 (S4C3), although Salty ( $p = .35$ ), Savoury ( $p = .20$ ) and Sour ( $p = .17$ ) were sometimes reported.

In summary, responses in Cluster 1 tended to contain reports of Salty, in combination with any taste except Metallic (State 4) or with Metallic (State 2). Responses in Cluster

2 tended to contain reports of Sour (State 2), Savoury (State 4) and Metallic. Responses in Cluster 3 tended to contain reports of Bitter, either in conjunction with dominant Sweet (State 2), in conjunction with dominant Metallic (State 5), or in conjunction with Sour (State 3).

### **3.3.3 Taste and time relationships**

Significant differences in citation proportions are highlighted in the TCATA difference plots in Figure 3-8 (Cluster 1), Figure 3-10 (Cluster 2) and Figure 3-12 (Cluster 3). Results of fda for significant taste and cluster combinations (as identified through differences in TCATA citation proportions) are shown in Figure 3-9 (Cluster 1), Figure 3-11 (Cluster 2) and Figure 3-13 (Cluster 3). Note that the points indicated on the fda figures were obtained empirically from the first and second derivatives of the curves. As these curves represent both citation frequency and intensity (Jaeger et al., 2020; Vidal et al., 2021), any increases or decreases described refer to both these phenomena. An expanded version of the functional analysis, showing the derivative curves, is included in Appendix E.

#### **3.3.3.1 Cluster 1**

When considering TCATA citation proportions, the responses in Cluster 1 more frequently contained reports of Metallic from the end of the cooling trial portion (8.7°C) to approximately halfway through the warming trial portion (19.7°C) (Figure 3-8). In addition, Cluster 1 more frequently contained reports of Salty for a sizeable portion of the trial – from 24.9°C in cooling to 17.4°C in warming (Figure 3-8).

Based on the fda, responses in Cluster 1 had the highest proportion of Metallic responses at 9°C during the cooling trial portion (Figure 3-9). However, there are three critical points during cooling, namely 25°C, 18°C and 15°C. At these critical points, the

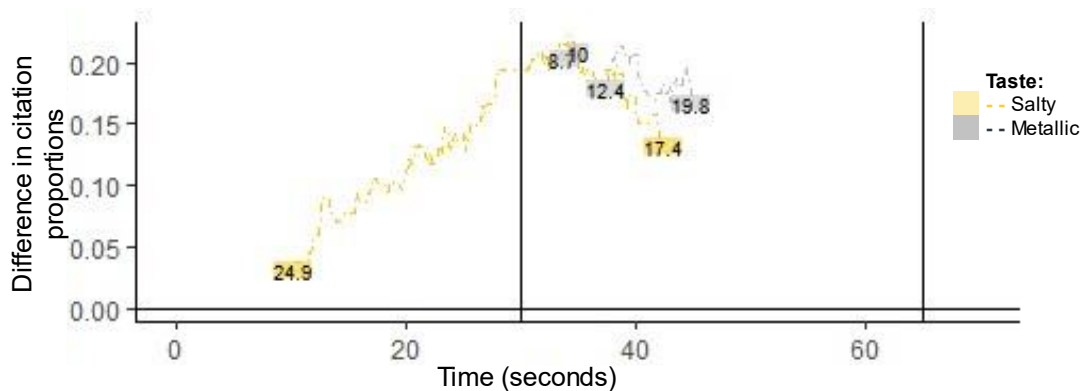


Figure 3-8 TCATA difference plot showing the citation proportion difference between Cluster 1 and the entire sample. Significance levels were obtained using bootstrapped confidence intervals. Only areas of the plot where a taste was reported more frequently than the overall sample are shown. Shaded numbers highlight start and end temperatures of significant portions.

second derivative is zero, indicating a change in predominant clicking behaviour (either clicking or unclicking) – pointing to temperatures where the taste is either first induced or stops, likely linked to potential mechanisms. After the maximum at 9°C, the Metallic curve reaches a minimum at 36°C. However, there are several critical trigger points along the way, namely 7°C during Cooling, 7°C during Warming, 16°C during Warming and 29°C and 32°C during Warming. The curve has a slight plateau, ending at 16°C. The number of critical points (changes in curve trajectory) suggests the involvement of several mechanisms for inducing the Metallic taste responses contained in this cluster.

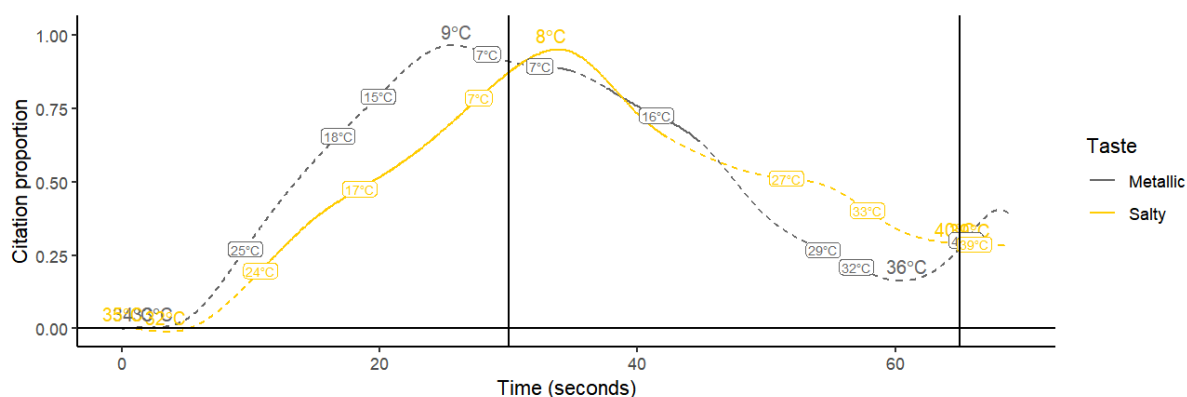


Figure 3-9 Functional curves obtained from Metallic and Salty reported frequencies for responses contained in Cluster 1. Numbers in bold indicate temperatures where local maxima or minima are reached (first derivative zero), and boxed numbers indicate “trigger points” where the first derivative reaches a local minimum/indicating (second derivative zero) indicating that the taste is being frequently clicked or unclicked. Solid lines indicate temperatures where the taste in question was more frequently reported by Cluster 1 (Figure 3-8)

Although there are differences in some critical points, the functional curve for the Salty responses in Cluster 1 follows a similar trajectory to Metallic. This curve reaches its maximum at 8°C during Warming and has a minimum at 40°C (Warming)(Figure 3-9). In the Cooling portion of the trial, the critical points are at 24°C, 17°C, and 7°C. During Warming, there is a plateau ending at 27°C, with the curve decreasing sharply again at 33°C. Again, this suggests that the Salty responses in this cluster have several underlying mechanisms.

### 3.3.3.2 Cluster 2

When considering TCATA citation proportions, the responses in Cluster 2 contained more reports of Salty than the other two clusters at 15.2°C during Cooling (Figure 3-10). Cluster 2 also contained more frequent reports of Sour for most of the trial – starting from 24.2°C during Cooling and ending at 29.8°C during Warming (Figure 3-10) and again from 39.2°C during Warming until the end of the trial. In addition, Cluster 2 contained more frequent reports of Savoury for most of the warming trial portion – starting at 5°C and ending at the end of the trial (Figure 3-10).

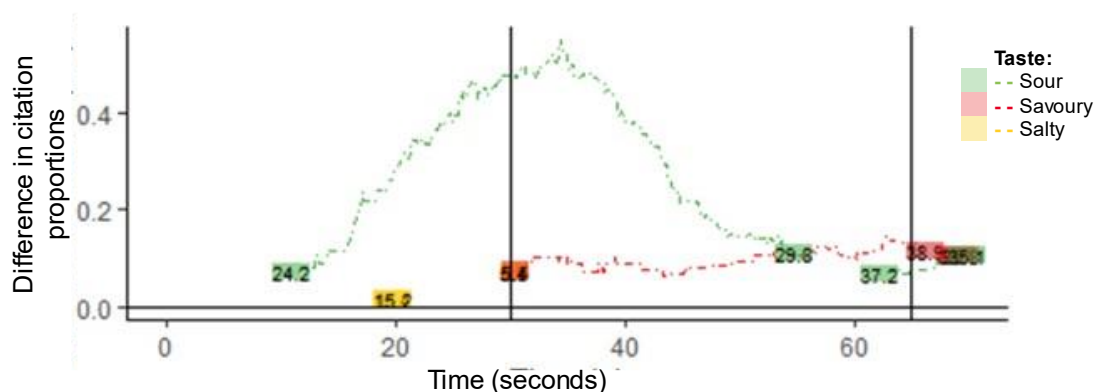


Figure 3-10 TCATA difference plot showing the citation proportion difference between Cluster 2 and the entire sample. Significance levels were obtained using bootstrapped confidence intervals. Only areas of the plot where a taste was reported more frequently than the overall sample are shown. Shaded numbers highlight start and end temperatures of significant portions.

The responses in Cluster 2 mostly contained Sour responses at 8°C, with an initial critical point at 16°C during Cooling (Figure 3-11), and a subsequent critical point at 17°C during warming, suggesting that Sour was induced through a mechanism linked to 16-17°C and that it was activated and deactivated at similar temperatures during Cooling and Warming. A local minimum for Sour occurs at 36°C during Warming, after two critical points at 25°C and 28°C, respectively. This suggests that Sour is induced, although at a lower frequency, through a second warming-linked mechanism.

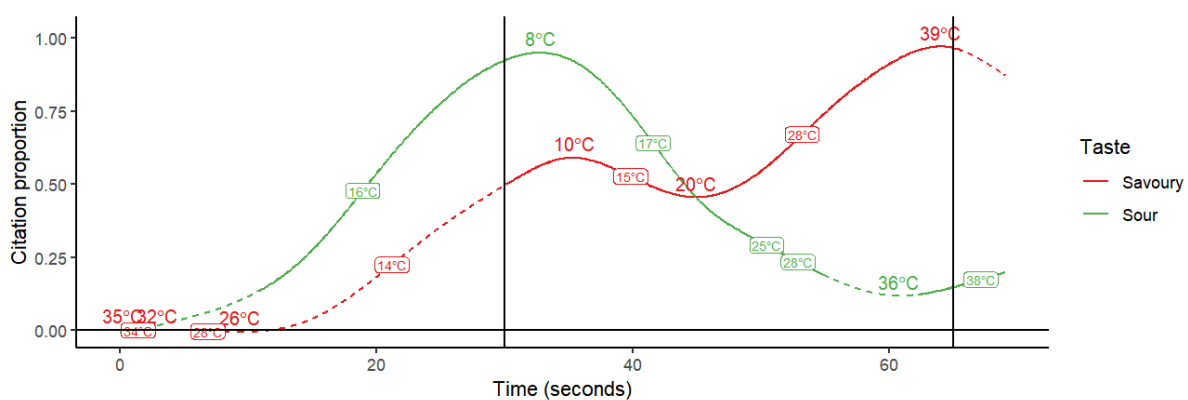


Figure 3-11 Functional curves obtained from Savoury and Sour reported frequencies for responses contained in Cluster 2. Numbers in bold indicate temperatures where local maxima or minima are reached, and boxed numbers indicate “trigger points” where the first derivative reaches a local minimum/indicating points where the taste is either being clicked or unclicked. Solid lines indicate temperatures where the taste in question was more frequently reported by Cluster 2 (Figure 3-10)

The functional curve for Savoury (Figure 3-11) initially followed a similar trajectory to Sour, although here, the local maximum occurred at 10°C during Warming, with the mirrored critical points at 14°C in Cooling and 15°C in Warming, respectively. However, Savoury reaches its maximum at 39°C (Warming) after a critical point at 28°C, indicating that Savoury is induced through a secondary mechanism in Warming, which is most active at 28°C.

### 3.3.3.3 Cluster 3

When considering TCATA citation proportions, the responses in Cluster 3 more frequently contained Bitter responses for a sizable portion of the trial portion, starting

in the cooling portion (18.7°C) and ending in the warming portion (27°C) (Figure 3-12). In contrast, Cluster 3 more frequently contained Sweet responses during four distinct portions of the trial – twice during Cooling (28°C - 17°C and 12°C – 8°C) and twice during Warming (6°C to 10°C and 30 - 37°C) (Figure 3-12). Cluster 3 also more frequently contained Salty responses towards the end of the warming trial portion (between 28.2°C and 35°C) (Figure 3-12).

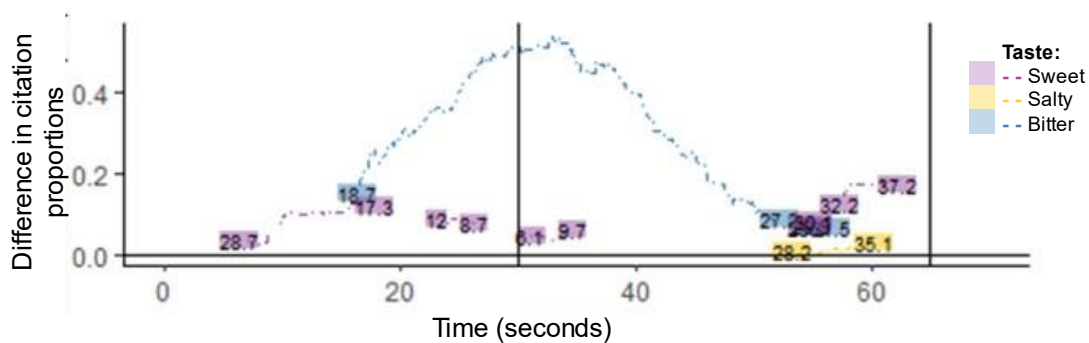


Figure 3-12 TCATA difference plot showing the citation proportion difference between Cluster 3 and the entire sample. Significance levels were obtained using bootstrapped confidence intervals. Only areas of the plot where a taste was reported more frequently than the overall sample are shown. Shaded numbers highlight start and end temperatures of significant portions.

In the curve extracted from the responses in Cluster 3, Sweet reaches its first maximum at 21°C during Cooling, with the critical point at 27°C in Cooling (Figure 3-13). It reaches a local minimum at 6°C (at the start of Warming) after two critical points at 18°C (Cooling) and 9°C (Cooling). Sweet reaches a second local minimum at 12°C (Warming), and after a critical point at 8°C (Warming), the curve decreases to 23°C. The next significant trigger point is at 31°C, leading Sweet to reach a maximum proportion at 37°C. Overall, the shape of the curve suggests at least three underlying mechanisms for Sweet – one activated during the Cooling trial portion and two during the Warming trial portion.

In responses contained in Cluster 3, Bitter citation frequency increases from Cooling, with its critical point at 18°C, reaching its maximum at 7°C (Warming)

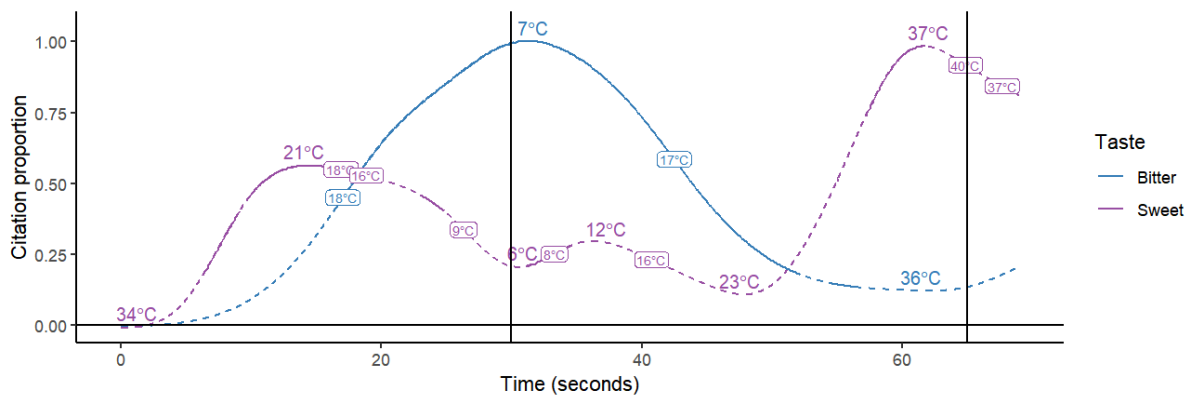


Figure 3-13 Functional curves obtained from Bitter and Sweet reported frequencies for responses contained in Cluster 3. Numbers in bold indicate temperatures where local maxima or minima are reached, and boxed numbers indicate “trigger points” where the first derivative reaches a local minimum, indicating points where the taste is either being clicked or unclicked. Solid lines indicate temperatures where the taste in question was more frequently reported by Cluster 3 (Figure 3-12)

(Figure 3-13). The proportion of Bitter decreases, with a trigger point at 17°C (Warming), finally reaching its minimum at 36°C (Warming). This suggests a singular mechanism for bitter thermal taste, with a symmetrical induction/stopping point at 17-18°C.

### 3.4 Discussion

This chapter evaluated overall thermal taste and cluster response patterns using a model-based approach. Additionally, differences between clusters in thermal taste response were evaluated, facilitating the discovery of relationships between taste responses and temperatures. These relationships inform potential mechanisms for thermal taste.

This discussion first explores the patterns in which tastes are reported together and which tastes are associated with which temperature ramp. Next, detailed response patterns are focused on to shed light on the potential underlying mechanisms of thermal taste. Finally, clustering outcomes are discussed.

### 3.4.1 General Patterns of taste response

#### 3.4.1.1 Co-reporting of tastes

Less than half of the thermal taste responses (41%) contained a single taste, considerably less than the 69% of responses in Skinner et al. (2018). The higher proportion of multiple tastes is likely a result of the lower working memory load requirement of RapCoTT, as allowing participants to report thermal taste in real-time avoids the effects of recency and memory corruption (see Chapter 2's discussion for more details). In previous thermal taste studies, more participants may have perceived multiple tastes but only reported the most dominant taste perceived, potentially invalidating studies comparing taste subgroups defined by a singular taste (for example, Thibodeau et al., 2020).

##### 3.4.1.1.1 Applying thermal taster classification

When considering TT as participants who reported a specific taste in duplicate on the same temperature ramp, Metallic TT were almost four times less likely ( $1/OR = 3.8$ ) to be Sweet TT and three and a half times less likely ( $1/OR = 3.4$ ) to be Sour TT, the latter finding agreeing with Thibodeau et al. (2020). Thibodeau et al. (2020) also found a direct negative relationship between Sour TT and Sweet TT, likely an indirect temperature trial relationship, as they found strong positive relationships between Sweet and Warming and Sour and Cooling. The present study failed to replicate these temperature trial relationships – discussed further in section 3.4.1.2.

##### 3.4.1.1.2 Pooled per participant

When considering all responses pooled per participant and temperature trial portion, there were no negative associations, with each positive association occurring in a chain-like pattern: Salty<-> Savoury<-> Sour <-> Bitter <-> Sweet (<-> indicating an

association). These relationships represent known frequently confused tastes (Salty/Savoury and Sour/Bitter), tastes grouped in the same response cluster (Savoury/Sour and Bitter/Sweet) and tastes with similar functional response patterns (Sour/Bitter).

Relationships between Salty and Savoury and Sour and Bitter could result from common confusion of these tastes (Doty et al., 2017; Puputti, Aisala, et al., 2019). Participants may be reporting different tastes to represent the same underlying experience (e. g. Salty on Rep1 and Savoury on Rep 2) or reporting both tastes to express their uncertainty about the experience (e. g. Sour and Bitter reported concurrently due to Sour/Bitter uncertainty). Patterns in emission probabilities support the hypothesis of multiple taste reporting when Bitter/Sour confusion occurs – in Sour-dominant states, there are residual probabilities of Bitter responses and vice versa. In contrast, there are only low residual probabilities of Salty responses during Savoury-dominant states and vice versa. Emission probability patterns support the idea that participants will respond with either Salty or Savoury but will report Bitter and Sour concurrently. The similar underlying functional response pattern between Bitter (Cluster 4) and Sour (Cluster 3) may also explain this relationship.

The remainder of pooled relationships reflect cluster structure: Savoury and Sour are associated, and both tend to be contained in responses in Cluster 2. Similarly, Sweet and Bitter both tend to be contained in responses in Cluster 3.

#### 3.4.1.1.3 Taste reported during the same trial portion

When considering responses individually split by trial portion, there were only negative relationships between tastes, suggesting that participants often tended to report only one taste per trial portion, but that there were specific patterns to these

responses. In addition, the negative associations suggest that participants who reported taste during both trial portions, tended to report different tastes during those trial portions.

#### *3.4.1.2 Relationships between temperature trials and tastes reported*

##### *3.4.1.2.1 Sweet*

The present study found no relationship between Sweet thermal taste and the Warming trial portion, contradicting previous findings (Skinner et al., 2018; Thibodeau et al., 2020; Yang et al., 2020). The advantages of RapCoTT likely underlie this discrepancy. In RapCoTT, participants report tastes immediately, so tastes are not stored in the working memory (as discussed in Chapter 2). As a result, participants are more likely to report what is perceived accurately, as interference effects of working memory (Daniel & Katz, 2018; Liang et al., 2018) are no longer at play. The fact that Sweet tends to be reported more often in other studies might have two underlying causes. The first is that Sweet might be easier to recognise, as naïve participants group stimuli into “Sweet” and “Non-Sweet” (Hettinger et al., 1999), likely causing perceptual bias towards Sweet when working memory is constrained. Secondly, Sweet tended to occur towards the beginning of the cooling trial portion of RapCoTT. Because of recency (Daniel & Katz, 2018), more recently perceived tastes might replace taste in the working memory, resulting in a lower overall proportion of Sweet responses. Similarly, with the warming trial portion, Sweet tends to be reported towards the end, meaning that the combination of perceptual bias towards sweet (Hettinger et al., 1999), and recency (Daniel & Katz, 2018), contributes to a tendency of participants to overreport Sweet thermal taste and underreport other thermal tastes that may have occurred during the warming trial portion – as this is the taste that they can best recognise and recall.

#### 3.4.1.2.2 Savoury

Savoury showed the most robust taste/temperature ramp relationships; individual responses were positively associated with the warming trial portion (5 times more likely to be reported) and negatively associated with the cooling trial portion (7 times less likely). Apart from the results reported in Chapter 2, this study was the first to find a relationship between Savoury/Umami reporting and the warming trial portion, but the lack of association in previous studies (Skinner et al., 2018 ; Thibodeau et al., 2020) is likely the result of the relatively low proportions of Savoury/Umami TT, stemming from methodological approaches which hinder Savoury/Umami reporting. Thibodeau et al. (2020) found no relationship between Savoury/Umami and temperature treatment, and relatively few Savoury/Umami TT overall, attributed to the fact that not all participants completed Umami stimulus training or received a ballot for Savoury/Umami reporting. Similarly Skinner et al. (2018) provided participants with ballot a scale for Umami response but no training with a Savoury/Umami stimulus, resulting in low proportions of Savoury/Umami thermal taste responses.

#### 3.4.1.2.3 Bitter and Sour

As Sour and Bitter show similar response patterns during fda and are prone to perceptual confusion (Doty et al., 2017; Puputti, Aisala, et al., 2019), their temperature-taste relationships are discussed jointly.

Bitter TT (participants who reported Bitter consistently on the same trial portion) were more likely to be Cooling TT (reporting thermal taste consistently in the Cooling trial portion), and individual Bitter responses were more likely to occur during the cooling trial portion. These findings agree with previous findings associating Bitter thermal taste with the cooling phenotyping treatment. (Cruz & Green, 2000; Skinner et al.,

2018; Yang et al., 2014; Yang et al., 2020). However, some previous studies found Bitter to be reported during both the cooling and warming treatments (Bajec et al., 2012; Hort et al., 2016; Pickering & Kvas, 2016; Thibodeau et al., 2020). The present study also failed to show any Sour-temperature trial (trial portion) associations, in contrast to previous findings associating Sour and the cooling treatment (Bajec et al., 2012; Pickering & Kvas, 2016; Skinner et al., 2018; Thibodeau et al., 2020; Yang et al., 2014).

Some of these differences may again be because of the increased performance of the RapCoTT protocol. The fda results show that both Bitter and Sour are induced towards the end of the cooling trial portion but extended into the warming trial portion. It is possible that, in previous studies, Sour experienced during warming was not reported - due to working memory interference (Daniel & Katz, 2018; Liang et al., 2018) previously discussed.

Differences in temperature treatments between previous and present studies may be a second underlying cause of the differences in Sour reporting patterns. The cooling treatment in the traditional phenotyping protocol ends with the tongue held at 5°C for 5s, which may have either resulted in more frequent induction or reporting of Sour thermal taste. Nachtigal and Green (2020) suggest that some thermal tastes could be a perceptual interaction of the physiological effects of tongue extension (drying out) and perceptual inference (astringency inferred). In the case of Sour, an opposite effect is likely occurring: cooling the tongue causes salivation (Dawes et al., 2000; Lee et al., 2006), likely through TRPA1 activation (Liu et al., 2018). As sour foods tend to increase salivation (Keesman et al., 2016), salivation may be perceptually confused with Sour taste (resulting in misreported Sour thermal taste) or increase the probability of participants who perceive Sour thermal taste (correctly) reporting it as

such. Note that the arguments around recency are also relevant here – in traditional thermal taste protocols, (salivation-enhanced) Sour thermal taste would occur at the end of the cooling temperature trial, potentially resulting in overreporting of Sour thermal taste and underreporting of other thermal tastes experienced during the (traditional) cooling trial.

#### 3.4.1.2.4 Metallic

Participants who reported Metallic thermal taste were more likely to report a thermal taste during the cooling trial portion, and Metallic thermal taste specifically was more likely to be reported during the cooling trial portion. Although the small cohort in Hort et al. (2016) reported Metallic more frequently during the cooling trial, in other studies (Skinner et al., 2018; Yang et al., 2020), participants reported Metallic at similar proportions during cooling and warming trials. In the present study, the highest frequency of Metallic responses in Cluster 1 occurred in the cooling trial portion, and Cluster 1 Metallic responses were only significantly more frequent at the start of the warming trial portion, which means that responses from all clusters contained Metallic responses during the cooling trial portion. This increased frequency of Metallic responses compared to previous studies could be related to the fact that RapCoTT facilitated reporting multiple tastes, allowing participants to report Metallic thermal taste alongside other tastes. Divalent metallic salts can produce several taste qualities, including bitter, sweet and salty (Skinner et al., 2017), which makes it unsurprising that Metallic thermal taste tends to occur alongside other thermal tastes.

#### 3.4.1.2.5 Salty

Salty thermal taste was weakly negatively associated with the cooling trial portion – thermal taste responses during the cooling trial portion were approximately 1.7 times less likely to contain a Salty response, contradicting earlier studies associating Salty

thermal taste with a cooling treatment (Cruz & Green, 2000; Pickering & Kvas, 2016). Other studies associated Salty thermal taste with neither trials/treatments (Skinner et al., 2018; Yang et al., 2014; Yang et al., 2020). A notable feature of studies associating Salty thermal taste with a Cooling treatment is that those studies allowed for reporting of Metallic thermal taste, suggesting that some of the association of Salty thermal taste with a cooling treatment is due to Metallic thermal taste being misreported as Salty. In other words, in studies where Metallic was not given as a reporting option, participants likely dumped Metallic sensation experienced under Salty, obscuring Salty-Cooling associations. In addition, some of the Salty responses were likely misreported Savoury responses. As Savoury was strongly associated with the warming trial portion, the weak negative association between Salty and cooling could be a spurious result.

### **3.4.2 Potential mechanistic links**

Examining the functional curves of specific thermal tastes led to insights regarding underlying mechanisms for thermal taste. To facilitate discussion, tastes with similar observed function patterns have been grouped.

#### **3.4.2.1 Sweet and Umami/Savoury**

The functional curves for Sweet and Umami/Savoury followed a similar trajectory during the warming trial portion, matching the pattern of principal component (PC) 1 of the 40°C trial in Skinner et al. (2018). Although Skinner et al. (2018) does not directly link this PC to Sweet thermal taste, it is plausible to assume that there is a link because a large proportion of the responses in Skinner et al. (2018) was Sweet thermal taste, and the first PC explained the most variation (although some responses may also have been Savoury/Umami thermal taste). Two peaks in the warming trial portion indicate two underlying mechanisms for Sweet (and Savoury/Umami) thermal taste,

in agreement with previous evidence (Green & Nachtigal, 2015) of two temperature-dependent pathways for sweet thermal taste.

Both Sweet thermal taste and Savoury/Umami thermal taste increase in the temperature range between 23-36°C (Sweet 23-36°C, Savoury 20-39°C with a trigger point at 28°C), corresponding to the temperature range at which TRPM5 is activated (15 - 35°C: Talavera et al., 2005). This temperature link between warming and Sweet thermal taste has been experimentally validated, with previous authors linking warming temperature ranges of 22 - 38°C (Skinner et al., 2018), 20 - 35°C (Nachtigal & Green, 2020) and 25 - 35°C (Ullah et al., 2022) to Sweet thermal taste. One study has linked a Salty-Sweet sensation to warming between 25°C and 35°C (Ullah et al., 2022), possibly corresponding to Savoury/Umami thermal taste. The direct link between these temperatures and Sweet thermal taste and Savoury/Umami thermal taste further supports the involvement of TRPM5 in both Sweet and Savoury/Umami thermal taste.

Sweet and Savoury/Umami thermal taste also had a secondary, smaller peak at the start of the warming trial portion. Savoury/Umami thermal taste shows clear activation at 14°C during cooling and deactivation at 15°C during warming. The deactivation point is similar for Sweet (16°C during warming), although the activation point is less prominent, as Sweet thermal taste was also produced during the cooling trial portion, and the curves overlap. The similarity in deactivation points indicates a secondary common underlying mechanism for Sweet and Savoury/Umami. TRPA1 is activated by cooling between 24°C and 12°C (Story et al., 2003) and has recently been shown to be necessary for Sweet and Umami taste perception (Matsumoto et al., 2023). Sweet and Savoury/Umami tastes' activation/deactivation temperatures in the present study closely match the mean activation temperature of 17.5°C  $\pm$  3.5°C of

TRPA1 (Story et al., 2003). In other words, the temperature at which TRPA1 was most likely to be activated/deactivated corresponds with the temperature at which participants were most likely to click or unclick the relevant thermal tastes.

The link between Sweet taste and TRPA1 is very recent and has not been further validated beyond Matsumoto et al. (2023). However, there is some other experimental evidence of the temperature range of TRPA1 activation being involved in Sweet taste. Green and Nachtigal (2015) found temperatures of 10°C and 5°C to affect sweetness perception and linked this to a second mechanism for the influence of temperature on sweet taste. In rodents, a temperature of 10°C induced a flavour preference (Kay et al., 2020), and in humans, exposure to water at 0°C induced a similar hypothalamus response (linked to satiation) to glucose at 22°C (van Opstal et al., 2018). TT specifically showed a higher cortical response than TnT to “cold sweet” samples (Eldeghaidy et al., 2021), although all participants showed a difference in response across temperature. Taken together, this underscores the validity of a cold-activated (below 20°C) mechanism for sweet taste- and a likely mechanism for cold, Sweet thermal taste.

Cooling the tongue produced Sweet thermal taste in a similar temperature range (34-21°C) to the Sweet and Savoury thermal tastes produced during warming. This pattern corresponds with PC4 for both trials in Skinner et al. (2018), and was also reported by Ullah et al. (2022). The fact that Sweet thermal taste, but not Savoury thermal taste, is produced through cooling at this temperature suggests a separate secondary TRPM5-dependent mechanism for Sweet thermal taste. A secondary pathway, independent of the TAS1R2 + TAS1R3 TRPM5 cascade, was recently proposed (von Molitor et al., 2020), likely involving the glucose transporters (GLUT) or sodium-glucose co-transporter (SGLT) (Breslin et al., 2021). There is evidence that this pathway requires

TRPM5 (von Molitor et al., 2020), as TRPM5 was required for polysaccharide taste response in mice (Sclafani et al., 2007). A second possibility is that Sweet thermal taste during cooling is induced via TRPM4, which shows a similar activation temperature to TRPM5 (Talavera et al., 2005), and is required for normal taste function (Banik et al., 2018). However, TRPM4 plays a role in transducing several tastes through broadly responsive Type III taste cells (Dutta Banik & Medler, 2023), which does not explain why only Sweet thermal taste is produced through cooling in this temperature range.

#### *3.4.2.2 Bitter and Sour*

Bitter and Sour responses follow a similar trajectory for most of the treatment, with a peak at the start of the warming trial portion and trigger points at 16°C/17°C (Sour) and 18°C/17°C (Bitter). This pattern resembles the trajectory of the curves of PC1 on the 5°C and PC2 in the 40°C trials in Skinner et al. (2018). Similar trigger points suggest a common underlying mechanism for Bitter and Sour thermal taste, with TRPA1 (most active at 17.5°C) again being a viable candidate. Although recent findings in mice (Matsumoto et al., 2023) only implicated TRPA1 in Sweet and Savoury/Umami sensation, Rhyu et al. (2021) makes a case for the involvement of TRPA1 in Bitter and Sour tastes. Also note that TRPA1 may still be involved in other taste sensations in humans, as differences between rodent and human TRPA1 are well-documented, differing for, for example, polyphenol (Takahashi et al., 2021), caffeine (Nagatomo & Kubol, 2008), and temperature (Laursen et al., 2015) responses.

Despite no direct evidence of the involvement of TRPA1 in Sour and Bitter taste, evidence exists linking TRPA1 to bitter-like and sour-like orosensations. TRPA1 is an acid-sensor (de la Roche et al., 2013; Wang et al., 2011), activated during the consumption of acidic (Sour) foods and beverages. Even if TRPA1 plays no direct role in Sour taste perception, the sensation of its activation may be associated with and

confused for Sour taste. Similarly, TRPA1 is part of the nociceptive response to CO<sub>2</sub> (Wang et al., 2010), and as many carbonated beverages are Sour, and CO<sub>2</sub> is immediately converted to carbonic acid in the oral cavity, the sensation of carbonation may be perceptually confused with Sour taste.

As with Sour taste, some Bitter-related responses involve TRPA1. TRPA1 responds to polyphenols (Takahashi et al., 2021), including astringent compounds in green tea (Kurogi et al., 2015). It is, therefore, possible that TRPA1 activation induces an astringent sensation in some participants, which is reported as bitter. In addition, caffeine deactivates (agonises) human TRPA1 (Nagatomo & Kubol, 2008). Caffeine also produces a Bitter taste through several different mechanisms (Poole & Tordoff, 2017), and, as with sourness, the strong association with this specific sensation might cause a bitter-related sensation to be reported as Bitter thermal taste in the absence of true Bitter thermal taste.

An additional potential underlying mechanism for Sour and Bitter thermal taste likely stems from the potential involvement of the sour taste receptor, OTOP1 (Turner & Liman, 2022), which could also be involved in the temperature-dependency of Bitter and Sour thermal taste. The structure of OTOP1 was only recently discovered, and there is evidence that it affects thermoregulation in mice (Tu et al., 2023), a function which also requires TRP channels (Señarís et al., 2018). In addition, there is evidence that OTOP1 is involved in the taste sensation of ammonium chloride (Liang et al., 2023) and can sense alkali (Tian et al., 2023). Both sensations could be potentially categorised/confused with bitter taste. Therefore, more detailed studies of the “sour sensor”, thermal response and its role in taste function more generally could uncover a potential mechanism relating to both sour and bitter taste.

The Sour response curve deviates from Bitter towards the end of the warming trial portion, with trigger points at 25°C and 28°C, suggesting that although the dominant behaviour is unclicking (curve sloping downwards), some participants are clicking and unclicking Sour at these temperatures. These temperatures correspond to the activation temperature of TRPV4 (24 - 27°C: Ferrandiz-Huertas et al., 2014), involved in Sour taste perception (Matsumoto & Kato, 2019; Matsumoto et al., 2019), and is expressed in the fungiform papillae (Moayedi et al., 2022). Interestingly, earlier studies identified TRPV4 in the elastic membrane of the tongue rather than the taste buds (Suzuki et al., 2003), and this expression of TRPV4 outside of the taste buds could explain Thibodeau et al. (2020)'s finding associating Sour thermal taste with thermal taste reporting from temperature stimulus on the side of the tongue.

#### *3.4.2.3 Salty and Metallic*

The functional curves for both Metallic and Salty show several “kinks”, indicating multiple underlying and potentially overlapping mechanisms. Also note that, due to the perceptual confusion between Salty and Savoury, some responses contained in Salty could be Savoury responses.

At least three mechanisms appear to be involved in Metallic thermal taste. Three temperature-relevant mechanisms have been linked to Metallic taste namely TRPM5 (Riera et al., 2009), TRPA1 (Ecarma & Nolden, 2021) and TRPV1 (Riera et al., 2008, 2009). Metallic thermal taste was initially induced between 25°C and 18°C, corresponding to the activation and deactivation temperatures of TRPM5. Similarly, the activation and deactivation temperatures of 15°C (Cooling) and 16°C (Warming) correspond to the activation temperature of TRPA1. In addition, the increase of Metallic citation at the end of the warming trial portion corresponds to the activation temperature of TRPV1. Therefore, all three temperature-sensitive taste channels

previously linked with Metallic perception appear to be involved in Metallic thermal taste.

The Metallic taste response curve shows some critical points not currently linked to taste sensation, such as 7°C and the 29 - 32°C (Warming) range. However, many TRP channels are either activated by metal ions, deactivated by metal ions, or act as channels for metal ions (Bouron et al., 2015), meaning many potential mechanisms for Metallic thermal taste could still be discovered.

Like Metallic thermal taste, Salty thermal taste was induced in the inverse activation range of TRPM5 (24°C – 17°C in cooling), involved in salt taste perception (Ren et al., 2013). In addition, Salty thermal taste was induced at the end of the Cooling trial portion (approximately 7°C), aligning with early findings of Salty thermal taste induction between 10°C and 5°C during cooling (Cruz & Green, 2000), aligning with the activation temperatures of the ENaC salt taste channel (Askwith et al., 2001), previously linked (Talavera et al., 2007) to Salty thermal taste during cooling.

Salty thermal taste again increased at the end of the Warming trial portion, with two areas of interest. The end of the Warming trial portion corresponds to the activation temperatures of TRPV1 (Talavera et al., 2007) and TMC4 (Kasahara, Narukawa, Kanda, et al., 2021), both linked to salt taste perception (Kasahara, Narukawa, Ishimaru, et al., 2021; Kasahara, Narukawa, Kanda, et al., 2021; Kasahara et al., 2022; Kim et al., 2014; Lyall et al., 2004). However, there is also evidence of activation between 27°C and 33°C, possibly an artefact of misreported Savoury thermal taste (linked to TRPM5).

### **3.4.3 Clusters of thermal taste responses**

This study found three clusters of thermal taste responses.

#### *3.4.3.1 Clustering efficacy and consistency*

The two responses (replicates) obtained from each participant tended to be clustered in the same MHMM cluster, indicating that participants were generally consistent in their thermal taste response. However, around 35% of participants were not consistently classified, with their two responses falling into different clusters. However, of the two responses obtained from these participants, one response had a high affinity for a specific cluster, and the other response had a low affinity for the cluster in which it fell (see Figure 3-6). This phenomenon justifies the use of these assigned clusters in future chapters.

Skinner et al. (2018) found between 19% (5°C trial) and 31% (40°C trial) of participants to have inconsistent responses. However, when the inconsistent participants in Skinner et al. (2018) were pooled (any participants who were inconsistent on any replicate), this group made up approximately 35% of their sample. The similarity in proportion of participants reporting thermal taste inconsistently between the present study and Skinner et al. (2018) suggests a stable subpopulation of TT who experience thermal taste consistently, but with an inconsistent experiential or reporting pattern.

This subgroup of TT represents a spectrum of thermal taste experience, ranging from intense and consistent to weaker and inconsistent. In previous studies, these participants may have been unclassified or unmatched. Thibodeau et al. (2019) explored the characteristics of unclassified TT, and divided unclassified participants based on the intensity of taste reported and the reproducibility (consistency across trials) of their taste response. Participants classified as “below weak only/reproducible” by Thibodeau et al. (2019) would have been classified as TT in the present study. It is not possible to compare the remainder of their unclassified participants with those who were inconsistently clustered in this study as the present

study had less stringent criteria for participants to be considered TT (participants could report any taste at any intensity at any point during the temperature trial, as long as participants reported taste during both replicates) and had more stringent criteria for consistency (participants were required to report thermal taste in a consistent pattern). However, this does raise the question of whether participants who were inconsistent in their taste response would have been unclassified using more traditional thermal taste phenotyping methods.

Clusters exhibited specific misclassification patterns. Participants who were classified in Cluster 3 in one replicate had the highest probability of being classified as Cluster 3 in both replicates, indicating highest cluster stability. In addition, if one response from a participant was included in Cluster 3, it was unlikely that the other response would be included in Cluster 1, indicating a more apparent distinction between the responses in Clusters 1 and 3 and a less clear distinction between Cluster 1 and 2 and Cluster 2 and 3. In other words, inconsistently classified participants were more likely to have one of their responses included in Cluster 2. Note, however, that when responses were inconsistently classified, there was a stronger affinity for one of the clusters, indicating a true cluster assignment. This assignment was chosen for use in the rest of this thesis.

Taste misidentification, confusion or perceptual similarity likely underlie misclassification patterns. Cluster 2 included Savoury and Sour responses, which are often confused with Salty (Cluster 1) and Bitter (Cluster 3), respectively (Puputti, Aisala, et al., 2019). In addition, the functional curves of Sour (Cluster 2) and Bitter (Cluster 3) showed many similarities, as did the patterns for Savoury (Cluster 2) and Sweet (Cluster 3) during the warming trial portion. Therefore, some participants may have experienced thermal taste during the same temperature portions through a common underlying mechanism, manifesting as different taste experiences across

different trials. In addition, more responses from Rep1 were classified as Cluster 1, whereas the responses from Rep2 from those participants were included in Cluster 2. The participants may have been true Savoury TT, who misreported Savoury as Salty but could recognise Savoury thermal taster in a second replicate - after reflection.

#### *3.4.3.2 Cluster characteristics*

When clustering RapCoTT responses using an MHMM approach, this study found three clusters of thermal taste responses. Previously, Skinner et al. (2018) clustered TT according to the temperatures at which they perceived taste most intensely and Bajec et al. (2012) and Thibodeau et al. (2020) grouped TT according to induced tastes (Bajec et al., 2012; Thibodeau et al., 2020), temperature ramp of induced tastes (Bajec et al., 2012; Thibodeau et al., 2020) and tongue location which induced tastes (Thibodeau et al., 2020). The present study considered temperature, temperature ramp, and taste identity and found three clusters, suggesting at least three distinct groups of thermal taste response.

The three clusters found in the present study differ from the four clusters found by Skinner et al. (2018). Skinner et al. (2018) clustered participants primarily on the temperatures at which their thermal taste response was the most intense, whereas current clustering considered both temperature and taste response. In the present study, only Cluster 1 contained responses of two different tastes (Metallic and Salty) exhibiting similar temperature-frequency relationships.

Although Metallic and Salty response curves for Cluster 1 were not identical, both increased during the cooling trial portion and decreased during the warming trial portion. Sour (Cluster 2) and Bitter (Cluster 3) response curves differed towards the end of the trial, but during the cooling trial and the warming trial portion (up to 24°C)

were strikingly similar. In addition, the response curves for Savoury (Cluster 2) and Sweet (Cluster 3) during the warming trial portion are nearly identical. These similar patterns suggest similar underlying mechanisms for each taste pair (Sour/Bitter and Savoury/Sweet). The complex relationship between the Metallic and Salty curves suggests that some Metallic/Salty responses have the same underlying mechanism, whereas other responses are produced through unique mechanisms that produce only Metallic or Salty thermal taste - but not both.

It is plausible that similar underlying mechanisms causing different thermal tastes (in the same temperature ranges) result from diversity in the expression of temperature-sensitive taste channels. There is some evidence of genetic differences in temperature-sensitive taste channels resulting in differences in taste experiences. Genetic differences in the expression of TRPV1 are related to differences in salt taste sensitivity (Dias et al., 2013; Pilic et al., 2020; Tapanee et al., 2021), and the temporal sensory quality of ethanol (Allen et al., 2014), and polymorphism differences in TRPA1 (Knaapila et al., 2012) are linked to the overall liking of cilantro. In addition, differences in the expression of TRPV1 and TRPA1 result in differences in sensitivity to capsaicin (Kringel et al., 2018) and differences in sensitivity to pain (Lötsch et al., 2020). However, only one study has examined the relationships between SNP's and thermal taste (Bering, 2013), finding no relationship between the SNP's in TRPM5 and TT Status. However, their study considered TT as a homogenous group and not as subgroups who experience and report different tastes, which means that a genetic basis for different tastes coming from a similar mechanism is still a plausible explanation.

### 3.5 Limitations

The first limitation of the research in this chapter relates to the data collection intent and the resulting processes. All data was collected with the primary intent to phenotype (classify) participants, and not with the intent of describing thermal taste as a phenomenon. As a result, the data collection protocols differed between cohorts, and learning effects demonstrated in Chapter 2 may have affected the data obtained from the portion of the cohort who completed the method validation study and completed the traditional phenotyping protocol preceding RapCoTT.

The limited training of participants prior to phenotyping is also a limitation of this study. Naïve consumers frequently misidentify taste sensations (Puputti et al., 2019), likely contributing to the error in the dataset, and potentially inconsistent classification. More extensive training would have improved the accuracy of tastes and sensations reported, as training and exposure enhances the ability of participants to identify stimuli, including aroma (Chambers et al., 2004; Clapperton & Piggott, 1979; Labbe et al., 2004), tastes (Labbe et al., 2004), textures (Chambers et al., 2004) and trigeminal (Byrnes et al., 2015; Han et al., 2020) sensations.

Although the most suitable approach for the available data, the constraints of the MHMM approach also present a limitation of this study. Two pieces of relevant information was not accounted for by the model, namely the relationship between the two responses obtained from the same participant, and the relationships between “reported” responses of tastes. (The model did not understand, for example, that “Sweet” and “Savoury” both represented taste reporting, and handled these as completely independent.) The first concern could be partially addressed by including a co-classification measure (e.g. Cohen’s Kappa) as a success measure in the genetic

algorithm. The second concern could be addressed by including an additional variable that encodes whether taste was reported

A final limitation is that participants could not indicate the intensity at which they experience tastes as previous research (Skinner et al., 2018) has shown differences in the intensity of TT response. In theory, replacing the TCATA paradigm with a multi-attribute time-intensity paradigm would achieve this, but this task might be too demanding for participants, limiting their ability to perform. One potential solution is to include a step at the end of the trial which allows participants to report the maximum intensity of taste perceived. Another alternative is to collect several replicates per participant, as this can be converted to citation proportions and treated as time-intensity. In this case, it can then be treated as functional data, which opens up many new possibilities for data analysis.

### **3.6 Conclusion**

Overall, this chapter found that thermal taste responses recorded through RapCoTT did not exhibit the same patterns as those reported in literature, likely due to different cognitive and physiological processes independent of thermal taste at play during the RapCoTT procedure.

Model-based clustering successfully found three distinct taste response patterns, which included Salty/Metallic (Cluster 1), Savoury/Salty (Cluster 2) and Sweet/Bitter (Cluster 3). Although not all responses from the same participants were consistently included in the same clusters, participants showed a higher affinity for a specific taste cluster. As a result, participants could be grouped into three TT Clusters for use in Chapters 4 and 5.

Functional data analysis confirmed existing hypotheses regarding thermal taste mechanisms and linked new mechanisms to thermal taste. Most notably, multiple mechanisms for Sweet thermal taste were identified. In addition, TRPA1 activation was linked to several thermal tastes, although some of the reported tastes may be other taste-related orosensation misreported as thermal taste. Finally, clear mechanistic links were established to Metallic thermal taste, disconfirming previous work speculating that Metallic thermal taste is a spurious sensation resulting from the mental association of Metallic taste with a metal thermode on one's tongue.

This chapter established TT Clusters used in the remainder of the thesis to explore differences between TT Clusters (Chapter 4) and the effect of thermal taste response on food choice behaviour (Chapter 5).

## **Chapter 4 *Seeing the forest and the trees*: Demystifying complex relationships between thermal taste response and demographic, personality, behavioural and orosensory responsiveness variables**

### **4.1 Introduction**

In Chapter 3, three distinct subgroups of TT were identified. Each subgroup (TT Cluster) reported *different thermal tastes* and tended to report thermal taste *at different temperatures* during the phenotyping run, supporting the findings of Skinner et al. (2018). The presence of TT subgroups suggests that subgroup differences skewed previous studies investigating differences between TT and TnT, perhaps confounding findings of no differences between TT and TnT (Bajec & Pickering, 2008; Thibodeau et al., 2019; Yang et al., 2014; Yang et al., 2020) or resulting in contradictory findings (for example food neophobia (Bajec & Pickering, 2010; Yang, 2015) and ethnicity (Thibodeau et al., 2020; Yang et al., 2020)). It is therefore necessary to re-evaluate previous findings regarding TT and contextualise their importance. In addition, some aspects of interest regarding thermal taste (as detailed in Chapter 1) have not yet been studied, warranting further investigation. These factors can be broadly categorised as demographic (gender, age and ethnicity), food behaviour (food involvement, food neophobia, oral processing behaviour, food modification and spicy food liking), personality (Private Body Consciousness and the Big Five Inventory) and orosensory sensitivity (Sweet Liker Status, PROP status, and fatty acid sensitivity).

Beyond the possibility of inter-TT-group characteristic differences potentially confounding findings, relationships between these different factors also exist. For example, Food Neophobia (Nezlek & Forestell, 2019) and oral processing behaviour (Laureati et al., 2020) are both related to personality as measured by using the Big

Five Inventory (BFI). There is some evidence of Food Neophobia differences between TT and TnT (Yang, 2015), and of BFI personality differences (Yang, 2015), but it is necessary to know whether those differences are distinct or related to each other. One way of determining this is by evaluating which factors impact prediction of TT Status or TT subgroup.

Machine learning can help to untangle the relative impact of different related aspects on a particular outcome. For example, Yang et al. (2020) recently used a regression tree to examine the relative impact of demographic and phenotypic factors (Thermal taster, Sweet liker and PROP taster status) on orosensory responsiveness to specific orosensory stimuli. A similar approach could demystify the relative importance of different aspects influencing thermal taste, although the appropriate choice of machine learning method is essential.

#### **4.1.1 Using machine learning to investigate interrelated factors**

##### **4.1.1.1 Terminology**

In this section, a distinction is made between “categorical” and “numeric” variables, aligning with the terms used in the literature discussing Random Forest models (RFM) (Strobl et al., 2007; Touw et al., 2012). This computational categorisation differs slightly from the traditional categorisation of nominal, ordinal, interval, and ratio scales, focussing on how variables are measured. When considering the categorical/numeric binary, the distinction between nominal and ratio variables is quite clear: nominal variables are *always* categorical, and variables on a ratio scale are *always* numeric. However, for ordinal and interval variables, the distinction is less clear-cut. Here, the question is whether using variables in future calculations is sensible or routinely done in practice. For example, although Likert scales are by

definition ordinal, data collected on Likert scales are routinely used in further calculation to derive numeric scores such as Food Neophobia (Pliner & Hobden, 1992) or Food Involvement (Bell & Marshall, 2003), and the nine-point hedonic scale is routinely used to find average scores for product ratings (Lawless & Heymann, 2010). However, other ordinal variables, such as PROP taster status, cannot be sensibly used for further computation. Therefore, this section uses a simpler computational distinction between categorical and numeric variables based on whether the variable can be sensibly used in computation (numeric) or not (categorical). This choice intends to shift the focus from the categorisation conundrum to the relevant computational issues.

#### *4.1.1.2 Factors to consider when choosing a machine learning method*

When choosing a machine learning method, the first characteristic to consider is its suitability for the data intended for prediction. As alluded to in the preceding section, a mixture of categorical and numerical variables is of interest when considering the different potential factors influencing thermal taste. For example, oral processing behaviour subtype and sweet liker status (SLS) behaviour are measured nominally in non-ordered categories. In contrast, many aspects of other food-related behaviour (for example, Food Involvement and Food Neophobia) and personality (like the Big Five Inventory) have numeric (score) outcomes derived from responses on ordinal Likert scales. In addition, the method needs to handle a large number of predictors compared to the sample size, as the full dataset is limited by the number of participants that complete a full and complex study. A second desirable characteristic is a method that requires relatively little input from the researcher regarding hyperparameter tuning, which facilitates faster model fitting (Smilde et al., 2022).

Beyond the specific characteristics of the ideal machine learning method, the method

would ideally have a track record of being used in sensory science, including being the best-performing model or approach in relevant studies comparing this approach with other machine learning methods.

Random forest models (RFMs) fit all of the requirements detailed above. An RFM fits several randomly generated classification and regression trees (Breiman, 2001). For each tree, samples and variables are randomly selected, with the final model generalising the outcome of all trees in the forest (Breiman, 2001). RFMs have the advantage over other machine learning methods in that they have relatively few parameters to tune, they are simple to use and can deal with small sample sizes and a large number of predictors (Biau & Scornet, 2016), which can be a mixture of categorical and numeric variables (Touw et al., 2012).

The ability to accommodate a mixture of categorical and numeric predictor variables is a particularly relevant benefit of RFMs (Touw et al., 2012). However, mixing predictor types has the potential for bias as categorical variables with several levels are more likely to be selected (Strobl et al., 2007; Touw et al., 2012). (Note that this concern is valid for nominal and ordinal variables best encoded categorically.) Using a conditional RFM addresses this concern through algorithmically equalising selection probability (Strobl et al., 2007). In R, this approach is implemented in the `party` (Hothorn et al., 2006; Strobl et al., 2008; Strobl et al., 2007) and `partykit` (Hothorn & Zeileis, 2015) packages.

Due to the benefits of random forest models, there are numerous examples of their use in sensory science, including establishing predictors of sensory properties (Cardinal et al., 2020; Doyennette et al., 2019; Eggink et al., 2012; Granitto et al., 2007; Keller et al., 2017; Lu et al., 2020; Men et al., 2018; Vigneau et al., 2018), drivers of consumer

liking (Bi & Chung, 2011; Roberts & Spadafora, 2020; Rocha et al., 2020), to model consumer behaviour more generally (Eggink et al., 2012), and to establish quality control models of fresh fruit (Ribeiro et al., 2021). In general, in sensory studies comparing random forests with other predictive methods, random forests outperformed alternative methods, including linear discriminant analysis (LDA) (Granitto et al., 2007), discriminant partial least squares (Granitto et al., 2008; Granitto et al., 2007), penalised discriminant analysis (Granitto et al., 2008), linear models (Keller et al., 2017), ridge regression (Keller et al., 2017), support vector machines (Keller et al., 2017), k-nearest neighbours (Keller et al., 2017), boosting (Keller et al., 2017), probabilistic neural networks (Men et al., 2018) and partial least squares regression (Vigneau et al., 2018).

#### **4.1.2 Interpreting machine learning models**

Although random forest models, and machine learning in general, have many benefits, one drawback is the difficulty of model interpretation. These models can be considered “black box” models, as the prediction mechanisms and criteria are hidden (Lo Piano, 2020). In contrast, a simpler model, like a linear model, can be represented by an equation, or a decision tree can be represented using a single graphic. Black box models are difficult to interpret, and their lack of transparency has associated ethical issues (Lo Piano, 2020). The field of interpretable machine learning methods is growing, with much activity in the field (Azodi et al., 2020; Molnar et al., 2020).

Interpretable machine learning methods are commonly divided in two ways: by considering whether the tool is model-specific or model-agnostic and whether the tool explains the impact of variables on the model as a whole (global or model-level) or explains the impact on individual predictions (local or individual -level) (Biecek & Burzykowski, 2021; Panigutti et al., 2020). Model-specific methods only aid in

explaining a singular model type, and recent developments have favoured model-agnostic methods (Molnar et al., 2020), probably because the effort to develop and maintain facilitating software is better invested in more widely applicable tools. Model-specific methods also allow for fair comparison of different types of models and are immediately applicable to novel machine-learning methods as they do not need to be specifically developed for each modelling approach (Biecek & Burzykowski, 2021). (Note that this reflects the experience of the researcher. Although there are model-specific tools available for RFMs, the R packages dedicated to these tools were either outdated (no longer maintained), poorly documented or had not been adapted for use with conditional RFM.)

Global (model-level) tools consider a variable's importance and impact on a model's predictions. For example, variables highly correlated with the predicted variable would be considered important. Variable importance measures fall into two categories: predictive loss and partial dependence. Predictive loss measures determine how much the model's overall quality would decrease when removing each variable (Greenwell & Boehmke, 2020). However, predictive loss measures can give different results when using different outcome (success) measures (Greenwell & Boehmke, 2020) – for example, evaluating the same model using Mean Absolute Error (MAE) or Receiver Operating Characteristic (ROC) as success measures could give different variable importance results. In addition, predictive loss measures can be biased when a model contains a mix of categorical and numeric variables (Strobl et al., 2007). Partial dependence type approaches look at how the prediction would change by changing a level of a variable if everything else was kept the same (Biecek & Burzykowski, 2021). However, traditional partial dependence plots can produce erroneous results when variables are strongly correlated, as multivariate extrapolation can cause prediction

errors (Apley & Zhu, 2020). This issue was addressed by the development of Accumulated Local Effects (ALE) plots, which do not rely on extrapolation (Apley & Zhu, 2020).

Local (individual-level) approaches typically fall into three classes, although a formal naming or designation of the classes has not been established (Biecek & Burzykowski, 2021). The first is the “what-if” approach; Ceretius Parebius (CP) profiles look at how the prediction for an individual will change if everything else is kept the same. These form the backbone of the partial-dependence approaches described in the previous paragraph. However, CP profiles for complex models can be challenging to interpret (Biecek & Burzykowski, 2021). The second “local model” approach includes the Local Interpretable Model-agnostic Explanations (LIME) method. Here, the model is evaluated based on performance and interactions at the values of interest (Biecek & Burzykowski, 2021). The third group of methods follow the “variable attributions” approach; these methods involve breaking down each variable’s contributions to its final prediction and include Breakdown and Shapley Additive Explanations (SHAP) plots (Biecek & Burzykowski, 2021). Breakdown plots essentially break down the projection of a specific observation into its intercept (baseline for all individuals) and the additive contribution of each of these variables (Staniak & Biecek, 2019). SHAP values come from game theory and test different potential combinations of variables in the model, giving an outcome that shows a distribution of how variables contribute to the individual’s prediction (Lundberg & Lee, 2017).

From a practical perspective, global variable importance measures give an overall measure of variable importance, whereas local measures help contextualise how these measures fit together. However, as TT Status and TT Cluster membership are of

interest, the global predictive importance of specific variables will be evaluated.

#### **4.1.3 Aims**

Overall, the objective of this chapter was to establish the relationships between thermal taste and demographic, food behaviour, personality and orosensory factors. Both TT Status, as established in Chapter 2, and TT subgroups (TT Cluster), as established in Chapter 3, were of interest. This objective consisted of two parts: first, to profile a group of phenotyped participants on the specific factors, and second, to establish the relative importance of selected factors and their constituent measures in determining TT Status and TT Cluster. This approach was selected as a result of two overarching hypotheses. The first hypothesis was that TT are not a homogeneous group and that for some aspects of difference, differences between TT subgroups obscure differences between TT and TnT. The second was that, as demonstrated previously (Yang, 2015), specific items/subvariables will differ more between TT groups (TT Status or TT Cluster) than their aggregated measures in some cases. In addition, the selection of factors was driven by specific hypotheses.

In general, as explored in more detail in Chapter 1, factors selected for profiling that have not been previously profiled were selected based on the expectation that there would be differences between TT groups. This includes oral processing behaviour, food modification, spicy food liking, Private Body Consciousness (PBC) and fatty acid responsiveness. For factors where previous findings were inconsistent, these inconsistencies were expected to be driven by differences in TT subgroups (Food Neophobia) or the population used in the study (age, ethnicity). For the remainder of the factors, it was expected null findings (SLS, PROP) or differences between TT and TnT (Big Five Inventory, Food Involvement) would be duplicated.

These objectives and hypotheses led to specific aims for this chapter, which were to:

- 1) Profile the group of participants phenotyped and classified in Chapter 3 in terms of food behaviour, personality and orosensory responsiveness.
- 2) Use random forest models to predict TT Status.
- 3) Use random forest models to predict TT Cluster.
- 4) Use interpretable machine learning methods to determine the relative impact of selected variables on the prediction of TT Status and TT Clusters.

## **4.2 Materials and Methods**

This study combined online survey and chemosensory response data, and compared outcomes according to TT Status (obtained from RapCoTT - Chapter 2) and TT Cluster (obtained from the MHMM – Chapter 3). Note that unclassified participants were classified according to the TT Cluster to which the replicate in which they reported thermal taste had the most affinity. Random forest models (RFMs) were employed to predict TT Status and TT Cluster from the survey, oral response, and demographic data. The RFMs were used to establish the relative importance of demographic, survey, and chemosensory response data through interpretable machine learning methods in predicting TT Status and TT Cluster.

### **4.2.1 Recruitment, data collection & data processing**

Data were collected in June 2020 and November-December 2020.

An initial group of participants was recruited from the cohort completing the method validation study in July-October 2019 (Chapter 2), and the cohort phenotyped for TT Status by a different researcher (November 2019- January 2020) - as described in

Chapter 3. These participants completed an online survey in June 2020 and subsequently participated in an in-person taste response study in November/December 2020. Some participants previously completed sweet liker status (SLS) (section 4.2.3.1) and PROP taster status (section 4.2.3.2) portions as part of another researcher's phenotyping study, and so these portions were not repeated.

A further cohort of participants were recruited in November/December 2020 who completed all portions of the study (thermal taste phenotyping, online survey, and taste responsiveness) during November/December 2020.

All participants were compensated for their time in the form of shopping vouchers. The study received full ethics approval from Massey Human Ethics Committee (SOA 19/26, SOA 19/50, and SOA 20/40) (Appendix A and B).

#### **4.2.2 Online survey**

The online survey consisted of five portions: a recruitment screener (which also collected demographic data) and four surveys. Each participant received a unique identifier, and all participants completed the questionnaires in the same order: Oral processing (Jeltema et al., 2015), Food behaviour (Temperature behaviour, Food involvement (Bell & Marshall, 2003), Food Neophobia (Pliner & Hobden, 1992), food modification (Puputti, Hoppu, et al., 2019), spicy food liking (Lawless et al., 1985), Personality (Big Five Inventory (John et al., 2008; John & Srivastava, 1999)), Private Body Consciousness (Miller et al., 1981)) and Food Frequency. Food Frequency data was collected for use in a future study, described in Chapter 5.

The recruitment survey differed slightly between cohorts. For the June cohort, Qualtrics was used to distribute the recruitment screener; completion triggered the distribution of the remaining surveys. For the November/December cohort, the

recruitment screener was distributed using Compusense; scheduling in-person sessions triggered the distribution of the remaining surveys. Gender/sex and ethnicity were collected according to the New Zealand standard (Stats NZ, 2019); as an open text field and ethnicity using Check-All-That-Apply with the response categories “European”, “Māori”, “Pacific\_Peoples”, “Asian”, “Middle Eastern/ Latin American/ African” and “Other ethnicity”. Age was collected in the form of birth year.

#### *4.2.2.1 Oral Processing Survey (Survey 1)*

Participants were classified using the proprietary JBMB Mouth Behaviour Typing Tool, described in Jeltama et al. (2015) under a free academic license from the owner. The tool presented participants with four pictures of groups of foods with unambiguous oral processing behaviour and a corresponding statement. Each image represented a different oral processing group, namely “Crunchers” (crunchy foods like watermelon, crispy vegetables and crunchy cookies with the statement “I like foods that I can crunch”), “Chewers” (chewy foods like brownies, pineapple and chewy candies with the statement “I like foods that I can chew”), “Suckers” (foods that are usually sucked like hard candies and orange pieces to suck on with the statement “I like foods that I can suck on a long time and often suck on them until they dissolve”) and “Smooshers” (foods that are usually consumed by smooshing such as custard, ripe banana and soft serve ice cream with the statement “I like foods that I can smoosh. I even smoosh foods that I can chew”). Subsequently, participants indicated whether any groups/statements did not describe them. Respondents were categorised as either a “Chewer”, “Sucker”, “Smoosher”, or “Cruncher”.

As part of the questionnaire, participants also completed questions pertaining to oral health, but the data collected in this instance were not used in further analyses.

#### 4.2.2.2 Food behaviour Survey (Survey 2)

##### 4.2.2.2.1 Temperature behaviour

Participants answered a short questionnaire regarding their temperature preferences for a set of food and beverage items; however, this data was not used in further analysis in this thesis.

##### 4.2.2.2.2 Food involvement

Participants indicated their level of agreement with statements on the food involvement scale (FIS; Table 4-1) (Bell & Marshall, 2003) (“Strongly Agree”, “Agree”, “Somewhat agree”, “Neither agree nor disagree”, “Somewhat disagree”, “Disagree” and “Strongly disagree”).

*Table 4-1 Items on the food involvement scale (FIS) (Bell & Marshall, 2003). Reversal (R) is indicated, as well as the existing subscales (Bell & Marshall, 2003) (Preparation & Eating and Set & Disposal). and the additional subscales tested in this study (Cooking, Talking and Planning).*

Item statement	R	Subscale	
		Existing	This study
I don't think much about food each day	Y	Preparation & Eating	Talking
Cooking or barbequing is not much fun	Y	Preparation & Eating	Cooking
Talking about what I ate or am going to eat is something I like to do	N	Preparation & Eating	Talking
Compared with other daily decisions, my food choices are not very important	Y	Preparation & Eating	Planning
When I travel, one of the things I anticipate most is eating the food there	N	Preparation & Eating	Planning
I do most or all of the clean up after eating	N	Set & Disposal	
I enjoy cooking for others and myself	N	Preparation & Eating	Cooking
When I eat out, I don't think or talk much about how the food tastes	Y	Preparation & Eating	Talking
I do not like to mix or chop food	Y	Preparation & Eating	Cooking
I do most or all of my own food shopping	N	Preparation & Eating	Planning
I do not wash dishes or clean the table	Y	Set & Disposal	
I care whether or not a table is nicely set	N	Set & Disposal	

Due to the test setup in Qualtrics, all items were in reverse during data collection, with “Strongly agree” having a value of 1 and “Strongly disagree” having a value of 7. As a result, values not reversed in the test were reversed, and the item scores reversed in the test (1, 2, 4, 8, 9 and 11) were left as is (Table 4-1). In addition to the original subfactors/subscales proposed by Bell and Marshall (2003) describing Preparation & Eating (P & E) and Set & Disposal (S & D), the researcher divided the P&E into three

behaviours, namely “Cooking” (food preparation), “Talking” (discussion and intellectual engagement with food and eating) and “Planning” (planning behaviour around food) of food and eating (Table 4-1).

The total food involvement score was the sum of all transformed values, ranging from 12 to 84. Scores on the relevant subscales (P & E, S & D, Talking, Cooking and Eating) were calculated by summing the transformed scores related to each subscale.

#### 4.2.2.2.3 Food Neophobia

Food neophobia was measured using the food neophobia scale (FNS) as described in Pliner and Hobden (1992) (Table 4-2). Participants indicated on a seven-point scale how much they agreed with each item in Table 4-2.

*Table 4-2 Items of the Food Neophobia Scale (FNS) used to assess Food Neophobia. Reversal (R) is indicated, as well as loading on the Approach and Avoidance subscales (Nezlek & Forestell, 2019). Note that by default, items load to the approach and avoidance subscales as is, except where reversal is specifically indicated.*

<b>Item statement</b>	<b>R</b>	<b>Subscale</b>
I am constantly sampling new and different foods	Y	Approach
I don't trust new foods	N	Avoidance
If I don't know what is in a food, I won't try it	N	Avoidance
I like foods from different countries	Y	Approach
Ethnic food looks too weird to eat	N	Approach (Reversed)
At dinner parties, I will try a new food	Y	Approach
I am afraid to eat things I have never had before	N	Avoidance
I am very particular about the foods I will eat	N	Avoidance
I will eat almost anything	Y	Avoidance (Reversed)
I like to try new ethnic restaurants.	Y	Approach

Survey responses were reversed, and the same data processing strategy used for food involvement (4.2.2.2.2) was employed, resulting in an FNS score for each participant ranging from 10-70. Subfactors of Approach and Avoidance (Nezlek & Forestell, 2019) were also calculated by considering some items in Table 4-2 as Approach and others as Avoidance, resulting in three aggregated FNS scores per participant (Neophobia, Approach, Avoidance).

#### 4.2.2.2.4 Food modification

A similar approach to Puputti, Hoppu, et al. (2019) was used to assess the prevalence of intentionally modifying the taste of food and beverages. The original questionnaire (Puputti, Hoppu, et al., 2019), focussing mainly on bitterness masking (addition of sweeteners to foods/beverages or addition of dairy to tea or coffee) and salt/salty condiment use, was supplemented with questions about spicy additions (Table 4-3).

*Table 4-3 Questions asked for food modification. Questions were phrased in the format “how often do you?”*

<b>Item statement</b>	<b>Subscale</b>
add milk to coffee	Bitter
add cream to coffee,	Bitter
add sugar to coffee,	Bitter
add sweetener to coffee,	Bitter
add sugar or honey to tea,	Bitter
add sweetener to tea,	Bitter
add milk to tea; modifying taste with salt or condiments:	Bitter
add salt to water when cooking vegetables,	Salt
add salt to a meal when eating it,	Salt
add aromatic salt (mixture of salt and seasoning) to a meal when eating it,	Salt
add ketchup to a meal when eating it,	Bitter
add soy sauce to a meal when eating it	Salt
add chilli to a meal when preparing it	Heat
add hot chilli sauce to a meal when eating it	Heat
add pepper to a meal when eating it	Heat

Not all participants were directed to answer all questions. Only participants who confirmed that they prepared their own food were asked questions about food preparation (Table 4-3). In addition, only participants who confirmed tea/coffee consumption answered questions about tea/coffee additions. Participants responded how frequently (“Never”, “Rarely”, “Sometimes”, “Often” and “Always”) they performed each modification. This selective question presentation allowed for the omission of the “not applicable” category in Puputti, Hoppu, et al. (2019).

To accommodate for unequal response numbers across participants, response proportions per frequency response (“Never”, “Rarely”, “Sometimes”, “Often”, and “Always”) per modification group (masking bitterness, adding saltiness, and adding spiciness) were calculated for each participant. This calculation was repeated with three

collapsed categories, namely “No” (Never/Rarely), “Sometimes” (Sometimes) and “Yes” (Often/Always).

#### 4.2.2.2.5 Spicy food liking and consumption

Participants indicated their affinity for spicy foods through a questionnaire based on Lawless et al. (1985). Participants initially indicated how frequently they consumed all types of chilli pepper in foods (including Mexican, Indian, Chinese, and other foods that contain chilli pepper and cause tingling or burning) with the options (1 = once a year or less, 2 = less than once a month, 3 = 1–3 times per month, 4 = once a week, 5 = 3 – 4 times a week, 6 = every day, and 7 = more than once a day). Participants also indicated their overall liking for the taste of chilli pepper in their food and their enjoyment of the burn of chilli pepper (7-point hedonic scale: “Like extremely”, “Like a moderate amount”, “Like a little”, “Neither like nor dislike”, “Dislike a little”, “Dislike a moderate amount”, “Dislike extremely”). Finally, participants answered three TRUE/FALSE questions about spicy food (“I think spices make food taste better”, “Without hot spices, I find that food tastes bland”, and “I find it hard to appreciate the flavours of food when the food contains hot spices”).

A value for overall spicy food liking was calculated according to Lawless et al. (1985) as the sum of use frequency, hedonic ratings and a score out of 12 based on the TRUE/FALSE questions. For each response, participants either got a score of 4 or 0. TRUE responses for the questions relating to the liking of spicy food (“I think spices make food taste better” and “Without hot spices, I find that food tastes bland”) each received a score of 4, as did a FALSE response for the remaining item (“I find it hard to appreciate the flavour of food when it contains hot spices”) (Lawless et al., 1985).

### 4.2.2.3 Personality Survey (Survey 3)

#### 4.2.2.3.1 Personality – Big Five Inventory

Personality characteristics were assessed using the Big Five Inventory (BFI) (John et al., 2008; John & Srivastava, 1999), consisting of 44 items (Table 4-5), each loading to one of five factors, namely Openness to experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (Table 4-5). Participants indicated how much they agreed (“Disagree strongly”, “Disagree a little”, “Neither agree nor disagree”, “Agree a little”, “Agree strongly”) with each statement in the format of “I am someone who...”. Scores of items that were reversed (R) (Table 4-5) were transformed appropriately, and the score in each of the five factors was the sum of the items loading to each of the five factors, as indicated in Table 4-5.

#### 4.2.2.3.1 Private body consciousness

Private body consciousness (PBC) was assessed using the scale by Miller et al. (1981). Participants indicated how accurately each item (Table 4-4) described them (“Extremely accurately”, “Very accurately”, “Moderately accurately”, “Slightly accurately”, and “Not accurately at all”). Category responses were recoded to represent numeric responses ranging from one (“Not accurately at all”) to five (“Extremely accurately”). Each participant’s PBC score was the sum of their recoded responses.

*Table 4-4 Items used to assess Private body Consciousness (PBC), as detailed in Miller et al. (1981). Note that none of the items are reversed, and no subscale is present.*

<b>Item text</b>	<b>R</b>	<b>Subscale</b>
I am sensitive to internal bodily tensions.	N	N/A
I know immediately when my mouth or throat gets dry.	N	N/A
I can often feel my heart beating.	N	N/A
I am quick to sense the hunger contractions of my stomach.	N	N/A
I’m very aware of changes in my body temperature.	N	N/A

Table 4-5 Items of the Big Five Inventory (BFI) (John et al., 2008), used to assess personality. Reversal (R) is indicated, as well as factor to which each item loaded (Agreeableness, Conscientiousness, Extraversion, Neuroticism, Openness to experience). Each item was presented in the format: "I am someone who".

Item statement	R	Factor
Is talkative.	N	Extraversion
Tends to find fault with others.	Y	Agreeableness
Does a thorough job.	N	Conscientiousness
Is depressed, blue.	N	Neuroticism
Is original, comes up with new ideas.	N	Openness to experience
Is reserved	Y	Extraversion
Is helpful and unselfish with others.	N	Agreeableness
Can be somewhat careless.	Y	Conscientiousness
Is relaxed, handles stress well.	Y	Neuroticism
Is curious about many different things.	N	Openness to experience
Is full of energy.	N	Extraversion
Starts quarrels with others.	Y	Agreeableness
Is a reliable worker.	N	Conscientiousness
Can be tense.	N	Neuroticism
Is ingenious, a deep thinker.	N	Openness to experience
Generates a lot of enthusiasm	N	Extraversion
Has a forgiving nature	N	Agreeableness
Tends to be disorganised	Y	Conscientiousness
Worries a lot	N	Neuroticism
Has an active imagination	N	Openness to experience
Tends to be quiet.	Y	Extraversion
Is generally trusting.	N	Agreeableness
Tends to be lazy.	N	Conscientiousness
Is emotionally stable, not easily upset.	Y	Neuroticism
Is inventive.	N	Openness to experience
Has an assertive personality.	N	Extraversion
Can be cold and aloof	Y	Agreeableness
Perseveres until the task is finished.	N	Conscientiousness
Can be moody.	N	Neuroticism
Values artistic, aesthetic experiences.	N	Openness to experience
Is sometimes shy, inhibited	Y	Extraversion
Is considerate and kind to almost everyone.	N	Agreeableness
Does things efficiently.	N	Conscientiousness
Remains calm in tense situations.	Y	Neuroticism
Prefers work that is routine.	Y	Openness to experience
Is outgoing, sociable.	N	Extraversion
Is sometimes rude to others.	Y	Agreeableness
Makes plans and follows through with them.	N	Conscientiousness
Gets nervous easily.	N	Neuroticism
Likes to reflect, play with ideas.	N	Openness to experience
Have few artistic interests.	Y	Openness to experience
Likes to cooperate with others.	N	Agreeableness
Is easily distracted.	Y	Conscientiousness
Is sophisticated in art, music, or literature.	N	Openness to experience

## 4.2.3 Taste phenotyping

### 4.2.3.1 Sweet liker status

Sweet Liker Status (SLS) was assessed following the method in Yang et al. (2019). Participants tasted a set of sucrose solutions (3%, 6%, 12%, 24% and 36% w/v) prepared by dissolving A1 sucrose (Davis Food Ingredients, New Zealand) in filtered

water using a stirring plate (Thermo Scientific, USA). A dummy sample (water) was tasted first, and the remaining samples were randomised with the constraint that the strongest and weakest samples could not be consecutive, as described in Yang et al. (2019). Participants initially indicated their liking of each solution using the Labelled Affective Magnitude (LAM) scale. Next, participants rated the sweetness intensities of the solutions using the general line magnitude scale (gLMS).

Participants were clustered using Yang et al. (2019); Agglomerative Hierarchical Clustering (AHC) was performed using Ward's method and dissimilarity, using the Euclidean distance. The `NBClust` (Charrad et al., 2014) and `cluster` (Maechler et al., 2021) packages were used to establish the most suitable number of clusters and to perform clustering. Pearson's correlation coefficient was calculated to check validity of cluster groups. Any participants with a correlation coefficient lower than 0.6 to their assigned cluster were initially assigned to the cluster for which they had the highest affinity (highest correlation coefficient). With the new cluster assignment, the correlation coefficients were calculated, and any participants with a correlation coefficient of lower than 0.6 were assigned to an additional 'hard to fit' or Unclassified (UN) cluster.

#### *4.2.3.2 PROP taster status*

A 0.32 mM PROP solution was prepared by dissolving 6-n-propylthiouracil (Sigma Aldrich, New Zealand) in filtered water on a stirring hot plate (Thermo Scientific, USA) at  $40 \pm 1^\circ\text{C}$ . Using a soaked cotton bud, the solution was applied to participants' tongue's anterior (front) tip. Bitterness was rated using the gLMS. Responses were duplicated after a break of 3 minutes.

PROP taster status was determined using the approach in Lim et al. (2008);

participants who rated the level of bitterness lower than barely detectable were classified as PROP non-tasters (PNT), those scoring between barely detectable and moderate were PROP medium tasters (PMT), and those scoring above moderate were PROP strong tasters (PST).

#### *4.2.3.3 Fatty acid sensitivity*

Fatty acid (FA) detection thresholds were determined using a difference from control approach recommended by Lawless and Heymann (2010), as described in Lundahl et al. (1986). Typically, FA thresholds are determined using an ascending staircase method (Haryono et al., 2014), but this was impractical for phenotyping a large sample.

Samples were prepared as detailed in Haryono et al. (2014). In short, oleic acid (Sigma Aldrich, NZ) was blended with fat-free milk (Countdown, NZ), food-grade mineral oil (Pure Nature, NZ) and gum arabic (Hawkins Watts, NZ) using a Silverson blender and refrigerated at  $5 \pm 1^\circ\text{C}$  until serving.

Participants were presented with a control (no oleic acid) and four sets of fat-free milk samples containing increasing levels of oleic acid (7 levels of oleic acid - 0.5, 1, 2, 4, 8, 16, 32 mM/L). Participants indicated how different each sample was from the control on a nine point scale, anchored with “No difference” (0) and “Extreme difference” (8) (Lundahl et al., 1986). Thresholds were calculated following Marin et al. (1991) whereby the threshold is the estimated point at which the response slope first increases. Full mathematical details are shown in Appendix F.

#### **4.2.4 Data analysis**

All data analyses were performed using R version 4.2 (R. Core Team, 2021). The `tidyverse` (Wickham et al., 2019) framework facilitated data manipulation, and the

`targets` package (Landau, 2021) and the `tidymodels` framework (Kuhn & Wickham, 2020) enabled data analysis. Several non-core `tidymodels` packages were used; `themis` (Hvitfeldt, 2022) assisted upsampling during model fitting, `finetune` (Kuhn, 2022) facilitated model tuning using a racing method and `bonsai` (Falbel et al., 2022) allowed for the fitting of the random forest included in `partykit` (Hothorn & Zeileis, 2015). The packages `vip` (Greenwell & Boehmke, 2020), `stablelearner` (Philipp et al., 2018; Philipp et al., 2016) and `DALEX` (Biecek, 2018) were used to evaluate variable importance and model impact.

Two different data analysis approaches were taken. First, in the traditional approach (section 4.2.4.1), overall outcomes for all measures were compared between TT groups (TT Status groups and TT Cluster groups) using traditional univariate approaches. Note that, to avoid data-dredging, a form of p-hacking (Streiner, 2015), constituent variables were not compared. Second, two RFM ensembles (section 4.2.4.2) followed by interpretable machine learning (section 4.2.4.2.4) were used to establish the overall predictive impact of *all* variables for TT Status and TT Cluster, including constituents of aggregate measures (raw data) and aggregate measures found not to differ significantly using traditional statistics. The rationale for taking these two approaches was to allow for comparison with literature (traditional approach) and to establish the relative importance of different factors including their constituents in predicting TT groups (RFMs).

#### 4.2.4.1 *Participant characteristics*

Two explanatory data analysis approaches were taken to determine any differences between TT Status groups (TT, TnT) or between TT Clusters identified in Chapter 3 (Cluster 1, Cluster 2, Cluster 3) and TnT. For any variables where the primary outcome

was categorical (nominal or ordinal; for example, SLS cluster), Pearson's chi-squared test was used to determine any relationships between outcome categories and TT Status or TT Cluster. For any variables where the outcome was primarily numeric, the Kruskal-Wallis test was used to determine whether there was an overall effect. The Wilcoxon-test was used to perform pairwise comparisons, and these were corrected for the false discovery rate (fdr). Due to the criticisms on significance testing (Wasserstein & Lazar, 2016; Wasserstein et al., 2019), a dual-alpha approach was taken: p-values between 0.05 and 0.1 were reported as notable, and p-values below 0.05 were reported as significant. In addition, 0.1 was used as a threshold for an effect in the Kruskal-Wallis test worthy of multiple comparisons (a subsequent Wilcoxon-test). This approach was taken because a cutoff value of  $p < 0.05$  for reporting results is considered arbitrary, and more thoughtful reporting of results is recommended (Wasserstein & Lazar, 2016; Wasserstein et al., 2019).

#### *4.2.4.2 Random forests*

Random forest models (RFMs) were used to attempt to predict TT Status and TT Cluster from demographic, survey, and taste responsiveness data. It was hypothesised that raw data (such as responses to individual survey questions) could better predict TT Status/cluster than the associated constructs (e. g., neophobia, factors in the Big Five Inventory). Two sets of models were fitted – a binary classification model predicting TT Status (excluding NC) and a multiclass classification model predicting TT Cluster.

The same overall approach was followed for both the TT Status (binary) and TT Cluster (multiclass) models, as shown in Figure 4-1. Initially, data was prepared or pre-processed to ensure suitability for modelling (described in the next section). Next, three sequential RFMs were tuned and fitted—all three RFMs aimed to determine

variable importance. However, the first two RFMs reduced the overall number of variables by eliminating low-importance variables from the model, and the final model was interrogated using interpretable machine learning approaches to establish the actual impact of retained variables on final predictions.

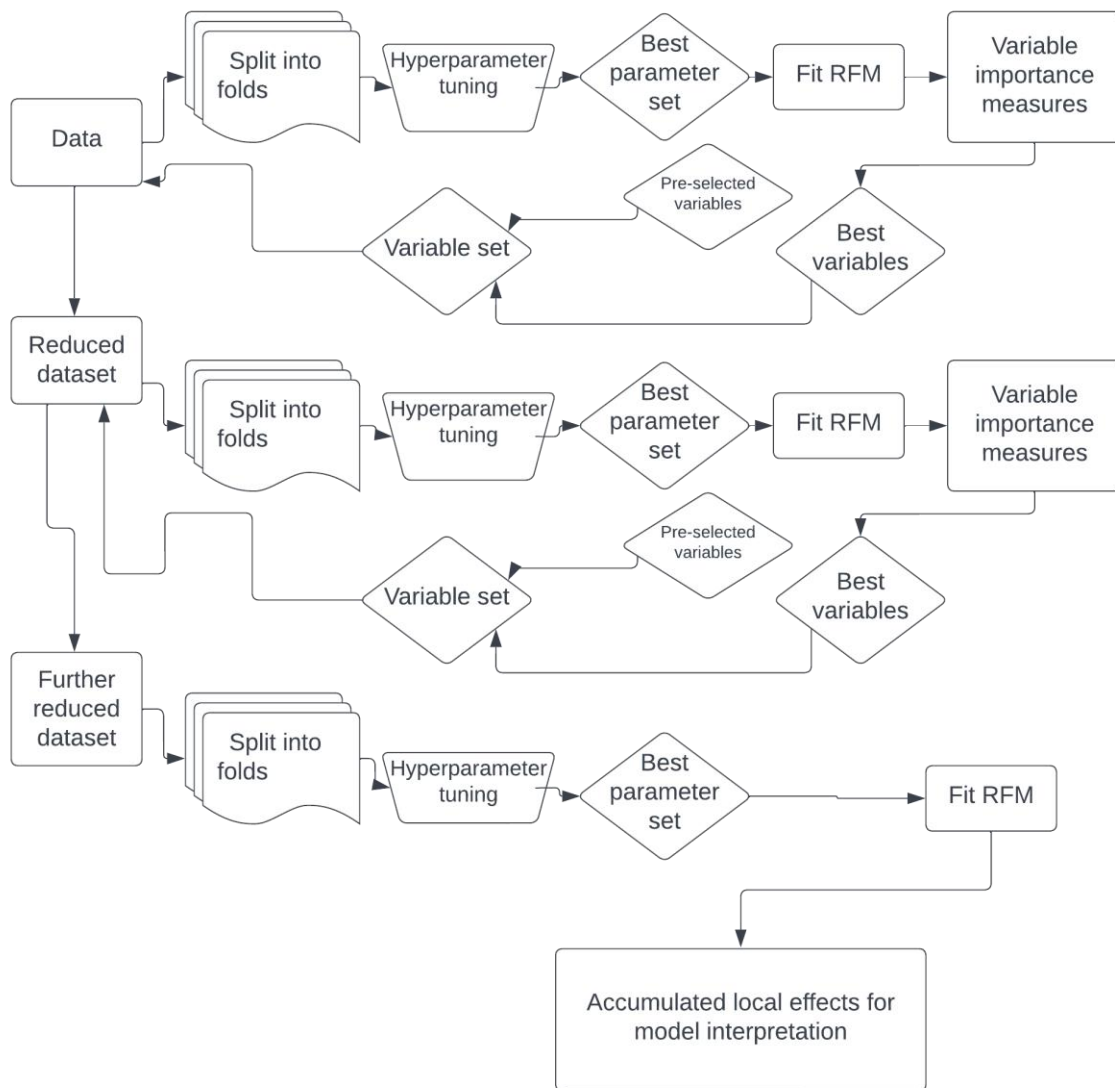


Figure 4-1 Flow chart showing the modelling approach for ensemble RFM fitting. Shapes are used to group similar steps or concepts, with all variables and parameters denoted as diamond shapes.

#### 4.2.4.2.1 Data pre-processing

Apart from the individual processing steps described in sections 4.2.2 and 4.2.3, additional pre-processing steps prepared the data for modelling. In initial test (model) runs, ethnicity data caused the model to fail, as some categories were highly

unbalanced. Using Principal component analysis (PCA) as a signal extraction technique solved this problem; PCA reduced the dimensionality (complexity) of the data without removing ethnicity variables (or participants from minority ethnic groups) from the data and retaining the nuance. The first two PC's were included in the modelling, which enabled modelling whilst preserving the complexity of the underlying ethnicity data.

Following signal extraction from the ethnicity data, highly correlated variables (higher than 0.9) were automatically removed from the combined dataset, retaining one variable from each correlated set. The retained variables were randomly chosen by the `recipes` package.

After initial pre-processing, data were randomly split into testing and training sets. The training set contained 80% of the observations, stratified by outcome (TT Status or TT Cluster). The training set was used for hyperparameter tuning, and the entire set was used for modelling.

#### 4.2.4.2.2 Hyperparameter tuning

The hyperparameters, number of trees (`ntrees`), number of variables selected for the best split at each node (`mtry`) and tree depth were tuned to find the parameters that led to the best fitting model (Probst et al., 2019; Touw et al., 2012). The same approach was taken for the initial (variable-selection) model and the final (interpretable) RFM. The overall hyperparameter tuning process is shown in Figure 4-2. The number of parameters sets (hypercube sizes) differed between tuning steps (Table 4-6), but were the same for both models (binary and multiclass).

*Table 4-6 Maximum entropy hypercube sizes used for hyperparameter tuning*

<b>RFM step</b>	<b>Hypercube size</b>
Step 1	9
Step 2	18
Step 3	27

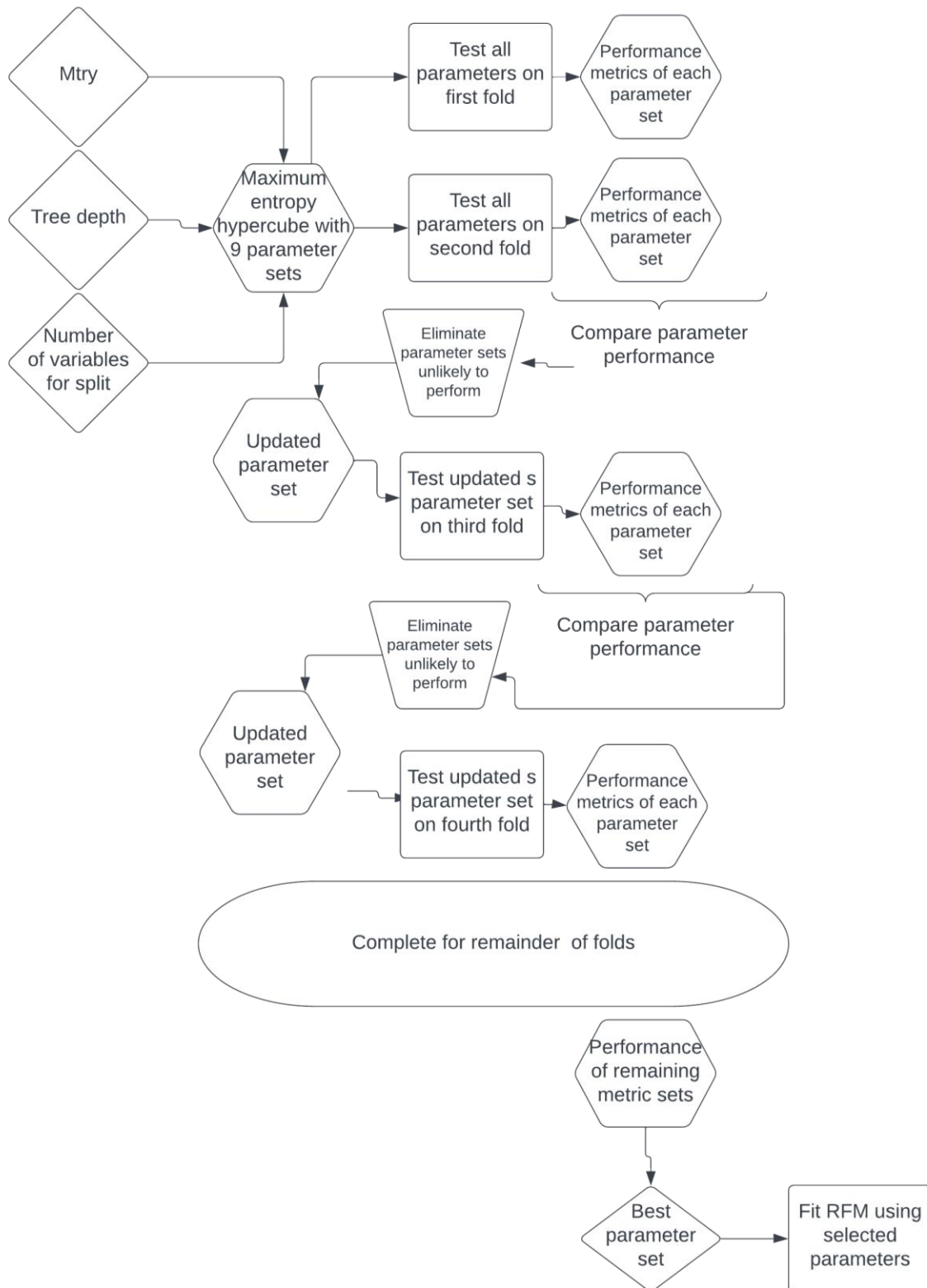


Figure 4-2 Flowchart explaining the approach taken for hyperparameter tuning during RFM fitting.

#### 4.2.4.2.3 Variable selection

The number of variables was systematically reduced by fitting an ensemble of models— with the first two RFMs aiming to reduce the number of variables requiring visual

evaluation. Variable selection was intentionally conservative, including an excess of variables rather than excluding a potentially important variable from future models. As the relative importance of variables in predicting TT subgroup (TT Status/ TT Cluster) was of interest, regardless of their performance using traditional univariate statistics, all variables were included at the start of modelling. Also note that both aggregate measures (for example, Food Neophobia) and the raw data used to calculate those measures (for example, the items on the Food Neophobia questionnaire) were included at the start of modelling.

The variable importance measures listed in Table 4-7 were calculated for each variable. Any variables that fell in the top 25% for any of the performance measures and any variables that were in the top 25% of their class (variables related to the same construct, for example, Big Five Inventory) were included. The threshold of 25% was arbitrary, but was chosen as small enough to reduce the number of variables in subsequent steps, but large enough to include all variables likely to be informative. Classes grouped the variables that measured similar constructs (i.e. from the same questionnaire/phenotyping test) with their aggregated measures (Figure 4-3).

*Table 4-7 Variable importance measures used for variable selection for subsequent RFM. (AOC = Area Under Curve)*

<b>Importance measure</b>	<b>Package</b>	<b>Reference</b>	<b>Explanation</b>
Importance to model	vip	Greenwell and Boehmke (2020)	Importance measured by mean accuracy loss from variable exclusion
Conditional importance to model	vip	Greenwell and Boehmke (2020)	Importance measured by mean accuracy loss from variable exclusion, compensating for correlated variables
Importance to model (AOC)	vip	Greenwell and Boehmke (2020)	Variable importance measured by effect of variable on AOC
Conditional importance to model (AOC)	vip	Greenwell and Boehmke (2020)	Variable importance measured by effect of variable on AOC compensating for correlations between variables
Stabletree frequency	stablelearner	Philipp et al. (2016)	The frequency of variable selection in a permuted RFM
Stabletree mean	stablelearner	Philipp et al. (2016)	The mean number of times that the variable was included in the stable tree model

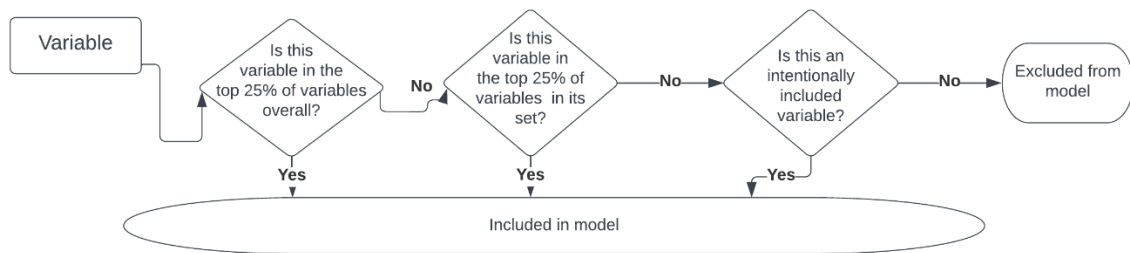


Figure 4-3 Flowchart showing decision process for inclusion of variables in subsequent RF models

Several variables were artificially included to ensure their representation in the final model (Figure 4-3). Artificially included variables were variables where the class consisted of only one (Age, Gender, and Oral Processing behaviour) or two (PROP status/intensity and Ethnicity) variables. Note that because the automatic variable selection process included both top variables overall and top variables by class, the automatic variable selection process would have included these variables in the final model, but their intentional inclusion ensured representation of variables from all classes in the final RFM. The artificial inclusion of these variables allowed for variables related to all hypotheses to be included in the final RFM and evaluated for their predictive impact.

#### 4.2.4.2.4 Obtaining insights from the models

Global variable impacts were assessed using accumulated local effects (ALE) plots (Apley & Zhu, 2020), which model each variable's impact across the variable's range. These plots also allow for a relative and intuitive measure of variable importance: variables that do not impact predictions will result in flat ALE plots, whereas those that do affect the prediction will have different impacts at different levels of the same variable. All plots are presented in Appendix H and Appendix I.

### 4.3 Results

#### 4.3.1 Participants

A total of 192 participants completed thermal taste phenotyping and demographic questionnaires, comprising 115 TT (Cluster 1: 64, Cluster 2: 29, Cluster 3: 40), 21 NC and 56 TnT. Of this cohort, 154 also completed the orosensory responsiveness part of the study, resulting in a dataset of 95 TT (Cluster 1: 54, Cluster 2: 24, Cluster 3:31), 13 NC and 46 TnT.

#### 4.3.2 Models

Five correlated variables were removed from the random forest before fitting (Table 4-8). Note that three of the correlated variables (regarding food modification) measure the same construct, as the middle category (Sometimes) was left as-is when collapsing the Food Modification data.

Table 4-8 Variables removed from the final RFM due to high correlation.

Variable removed	Correlated variables kept in model
Combined "Sometimes" salt addition	Raw "Sometimes" salt addition
Combined "Sometimes" heat addition	Raw "Sometimes" heat addition
Combined "Sometimes" bitter masking	Raw "Sometimes" bitter masking
Food involvement	Preparation & Eating Subscale
Total Spicy liking score	Liking of the taste of spices, Liking of the burn of spices

#### 4.3.3 Demographics

TT tended to be older than TnT, although not significantly ( $p = .087$  TT = 38.8, TnT = 35.1, Figure 4-4). However, TT Clusters did not differ in age, and no TT Cluster differed from TnT in age (Kruskal-Wallis  $p = .50$ ).

Six per cent of the sample identified with multiple ethnicities; notably, all Māori participants (2%,  $n = 3$ ) also identified as European. When taking into account only participants indicating a single ethnicity - mainly European (41%,  $n = 78$ ) or Asian (36%,  $n = 68$ ) - there was no relationship between either TT Status ( $p = .87$ ) or TT

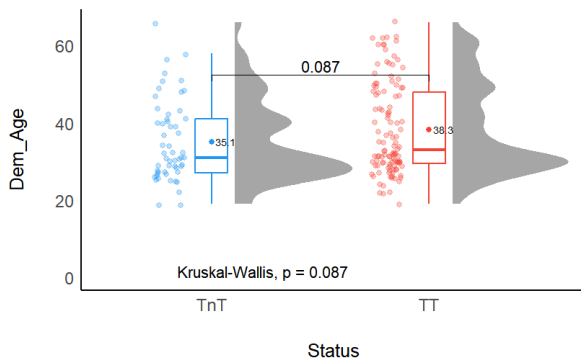


Figure 4-4 Distribution of the Age data according to TT Status showing boxplots and means, as well as summary statistics for the Kruskal-Wallis test.

Cluster ( $p = .70$ ). Gender (male  $n = 50$ , female  $n=140$ , all others  $n=1$ ) did not impact TT Status ( $p = .57$ ) or TT Cluster ( $p = .64$ ).

Although all demographic variables (Age, Gender, Two Ethnicity PCs) were intentionally included in both final random forest models (RFMs) none were useful in predicting TT Status or TT Cluster. These variables were intentionally included because they represented different constructs consisting of only one (Age, Gender) or two (Ethnicity) representative variables.

#### 4.3.4 Food behaviour (Survey 1 +2)

##### 4.3.4.1 Oral processing behaviour

Most participants were either “Chewers” ( $n = 68$ , 38%) or “Crunchers” ( $n = 60$ , 33%). “Smoothers” were less common ( $n = 41$ , 23%), and there were comparatively few “Suckers” ( $n = 11$ , 6%). These proportions align roughly with those found in previous studies (“Chewers” = 24 – 43%; “Crunchers” = 33 – 46%; “Smoothers” = 13 – 22%; “Suckers” = 2 - 8%) (Franks et al., 2020; Jeltama et al., 2014; Wilson et al., 2018). Note the comparatively high proportion of “Smoothers” aligns with the proportion previously found in a New Zealand population (22%: Wilson et al., 2018). There was no relationship between oral processing subtype and TT Status ( $p = .79$ ) or TT Cluster ( $p = .94$ ).

Oral processing behaviour subtype was intentionally included in the final RF models, as it was a singular variable representing a construct of interest. Oral processing impacted TT probability; being a cruncher slightly increased the likelihood of being a TT (Figure 4-5). The ALE plot in Figure 4-5 shows the impact of different Oral Processing behaviour subtypes on the probability of a participant being a TT if all other factors are kept the same, and shows the probability of a participant being a TT (0.658) based on sample size alone. In contrast to TT Status, oral processing behaviour subtype did not impact TT Cluster probability.

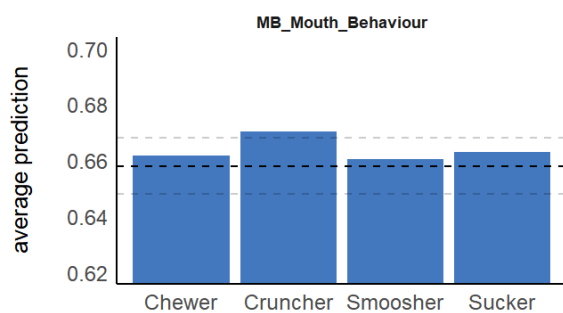


Figure 4-5 Accumulated local effects (ALE) plot for oral processing behaviour included in the final binary (TT Status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

#### 4.3.4.2 Food involvement

There were no differences between TT Status groups in overall food involvement ( $p = .70$ ), or in their scores on the Preparation and Eating (P&E) ( $p = .85$ ) or Set & Disposal (S&D) ( $p = .73$ ) subscales. In addition, there were no differences between TT Clusters (Overall  $p = .14$ ; P&E  $p = .20$ ; S&D  $p = .39$ ).

Based on their performance in the preceding RFMs (Figure 4-1, Figure 4-3), six individual items from the food involvement questionnaire were included in the final RFM, predicting TT Status (Table 4-9). Food Involvement overall was automatically removed prior to modelling due to its high correlation with the P&E subscale (Table

4-8). However, neither of the dimensions proposed by Bell and Marshall (2003) or this study (refer to Table 4-1) were included in the final TT Status RFM (Table 4-9). In contrast, seven items from the food involvement questionnaire were selected for the final RFM predicting TT Cluster (Table 4-9), as well as the Set & Disposal subscale from Bell and Marshall (2003), and the Talking subscale proposed by the present study.

*Table 4-9 Variables and constructs from the Food Involvement questionnaire included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs. Note that the Talking, Cooking and Planning constructs all form part of the Preparation & Eating subscale.*

<b>Construct</b>	<b>Binary model (TT Status)</b>	<b>Multiclass model (TT Cluster)</b>
Set & Disposal	<ul style="list-style-type: none"> <li>• I care whether or not a table is nicely set</li> <li>• I do most or all of the clean up after eating</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Set &amp; Disposal subscale</i></li> <li>• I care whether or not a table is nicely set</li> <li>• I do most or all of the clean up after eating</li> </ul>
Talking	<ul style="list-style-type: none"> <li>• When I eat out, I don't think or talk much about how the food tastes</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Talking subscale</i></li> <li>• Talking about what I ate or am going to eat is something I like to do</li> </ul>
Cooking	<ul style="list-style-type: none"> <li>• I enjoy cooking for others and myself</li> </ul>	<ul style="list-style-type: none"> <li>• I enjoy cooking for others and myself</li> <li>• I do not like to mix or chop food</li> </ul>
Planning	<ul style="list-style-type: none"> <li>• Compared with other daily decisions, my food choices are not very important</li> </ul>	<ul style="list-style-type: none"> <li>• Compared with other daily decisions, my food choices are not very important</li> </ul>

None of the selected variables were useful in predicting TT Status. In contrast, three variables were useful in predicting TT Cluster (Figure 4-6). This ALE plot is interpreted similarly to Figure 4-5; each coloured line denotes the probability of a participant forming part of a particular TT Cluster based in different levels of the specified variable, if all other variables are kept the same.

The Set and Disposal (S&D) dimension had the largest overall impact on group classification (Figure 4-6); participants scoring below 14 on S&D were substantially more likely to be part of Cluster 1 and less likely to be TnT. (A similar pattern was observed for Clusters 2 and 3). Responses to the item “I do most or all of the clean up after eating” (S&D) had a similar but smaller effect on class probability than S&D (Figure 4-6). In addition, participants who rated food choice as important (Disagreeing with “Compared with other daily decisions, my food choices are not very

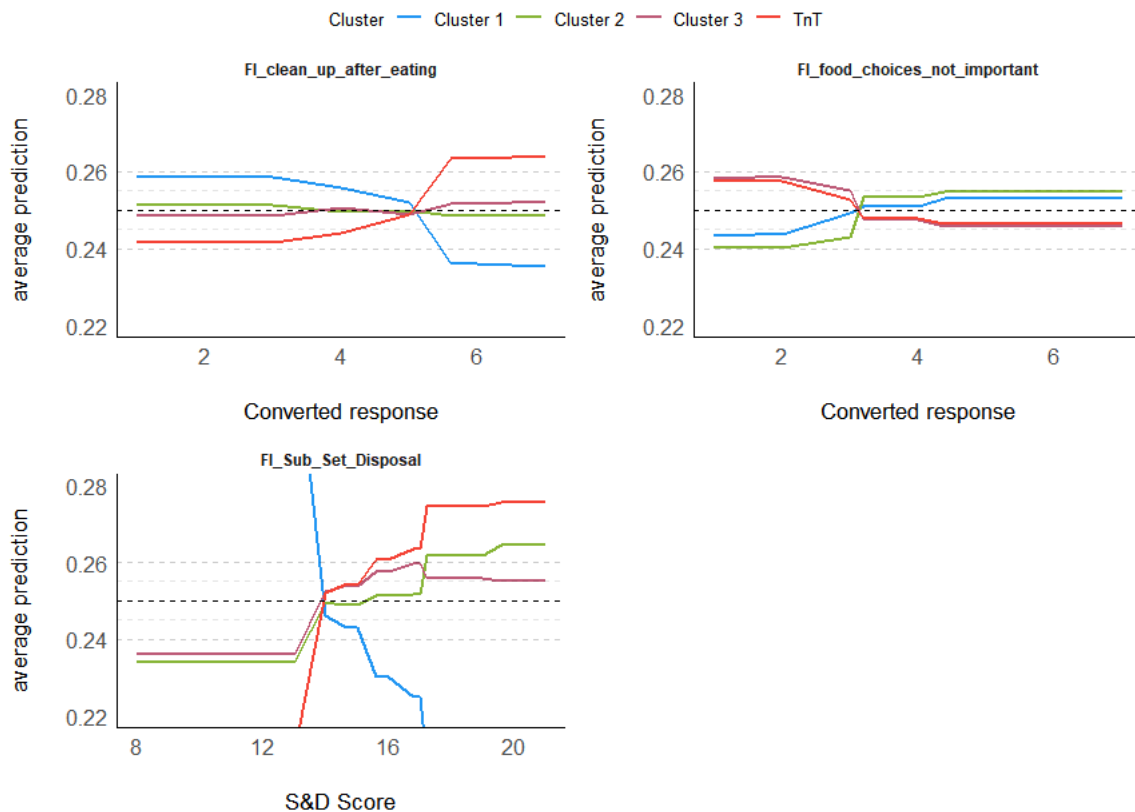


Figure 4-6 Accumulated local effects (ALE) plots for items from the food involvement scale (FIS) impacting the final multiclass (TT Cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability.

important”) were more likely to be classified as TnT or Cluster 3 and less likely to be included in Cluster 1 or 2.

#### 4.3.4.3 Food Neophobia

There were no differences in Food Neophobia scores between TT and TnT. ( $p = .84$ ). TT and TnT also did not differ in their scores on the Approach ( $p = 1$ ) or Avoid ( $p = .68$ ) subscales. In addition, there were no overall differences in Food Neophobia between TT Clusters (Neophobia:  $p = .43$ ; Approach:  $p = .70$ , Avoid:  $p = .18$ )

Based on their performance in the preceding RFMs (Figure 4-1, Figure 4-3), the overall Food Neophobia score and two items from the questionnaire (related to the avoidance construct) were automatically selected for the final TT Status RFM (Table 4-10). In addition to the Food Neophobia scores, both Approach and Avoidance constructs and

five items related to the Avoidance construct were selected for the final TT Cluster RFM (Table 4-10).

Table 4-10 Variables and constructs from the Food Neophobia scale included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.

Construct	Binary model (TT Status)	Multiclass model (TT Cluster)
Neophobia	• Neophobia	• Neophobia
Approach		• Approach subscale
Avoid	• If I don't know what is in a food, I won't try it • I will eat almost anything	• Avoidance subscale • If I don't know what is in a food, I won't try it • I will eat almost anything • I am afraid to eat things I have never had before • I am very particular about the foods I will eat

In the TT Status model, the overall FNS score, as well as responses to the “If I don’t know what is in a food, I won’t try it”, were useful predictors – participants who disagreed with this statement were more likely to be TT. Similarly, less neophobic participants (Neophobia score below 25) were more likely to be TT (Figure 4-7). Responses to the item “If I don’t know what is in a food, I won’t try it” also affected TT Cluster probability: participants disagreeing with this item were less likely to be classified as TnT and more likely as Cluster 2, consistent with the binary classification RFM (Figure 4-8). In addition, agreeing with the statement “I am afraid to eat things I have never had before” slightly increased Cluster 1 class probability: disagreeing slightly increased Cluster 2 probability and decreased Cluster 1 and TnT probability.

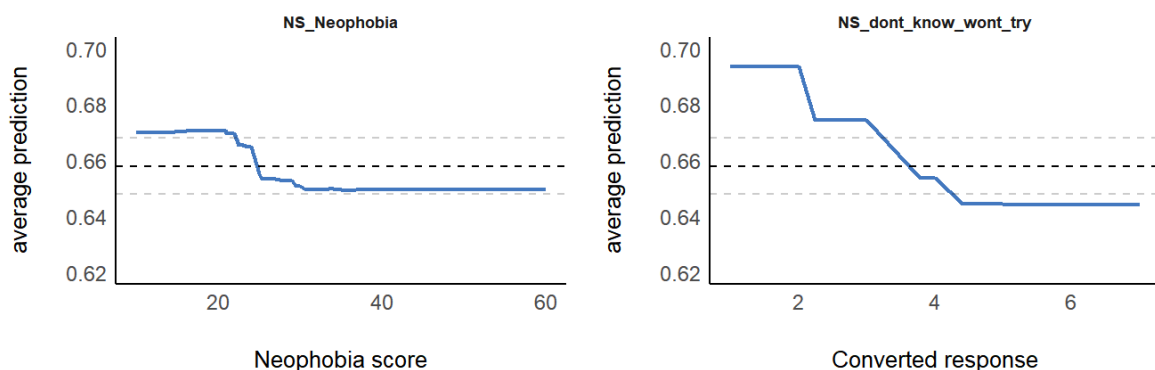


Figure 4-7 Accumulated effects (ALE) plots for items related to the food neophobia scale (FNS) included in the final binary (TT Status) RFM which impact TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept +/- 0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

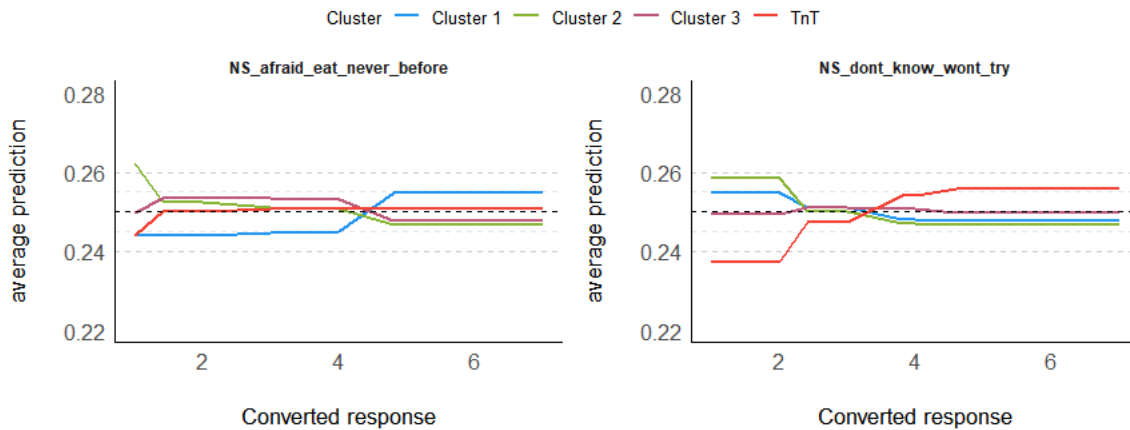


Figure 4-8 Accumulated local effects (ALE) plots for items and dimensions from the food neophobia scale (FNS) included in the final multiclass (TT Cluster) RFM which impact cluster probability. These plots show chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

#### 4.3.4.1 Food modification

Salt addition was the only food modification behaviour which differed between TT and TnT. TT indicated, on average, that they sometimes added salt to significantly more items than TnT ( $p = .029$ , Figure 4-9). In addition, although not significant, TnT never added spices or seasonings to increase the heat level more often than TT ( $p = .068$ , Figure 4-9). In contrast, there were no differences in food modification behaviour overall between TT Clusters.

Based on their performance in the preceding RFMs (Figure 4-1, Figure 4-3), four calculated indexes from the food modification questionnaire were automatically included in the final TT Status RFM (Table 4-11) two related to heat (spice) addition, one to bitterness masking and one to salt addition. Nine of the raw scores and six of the combined scores were included in the TT Cluster RFM (Table 4-11). Five items related to bitter masking, five to salt addition and four to heat addition were included in the TT Cluster RFM.

Heat and salt addition were useful predictors of TT Status. Participants who always added heat to their food were significantly more likely to be TT (Figure 4-10). In

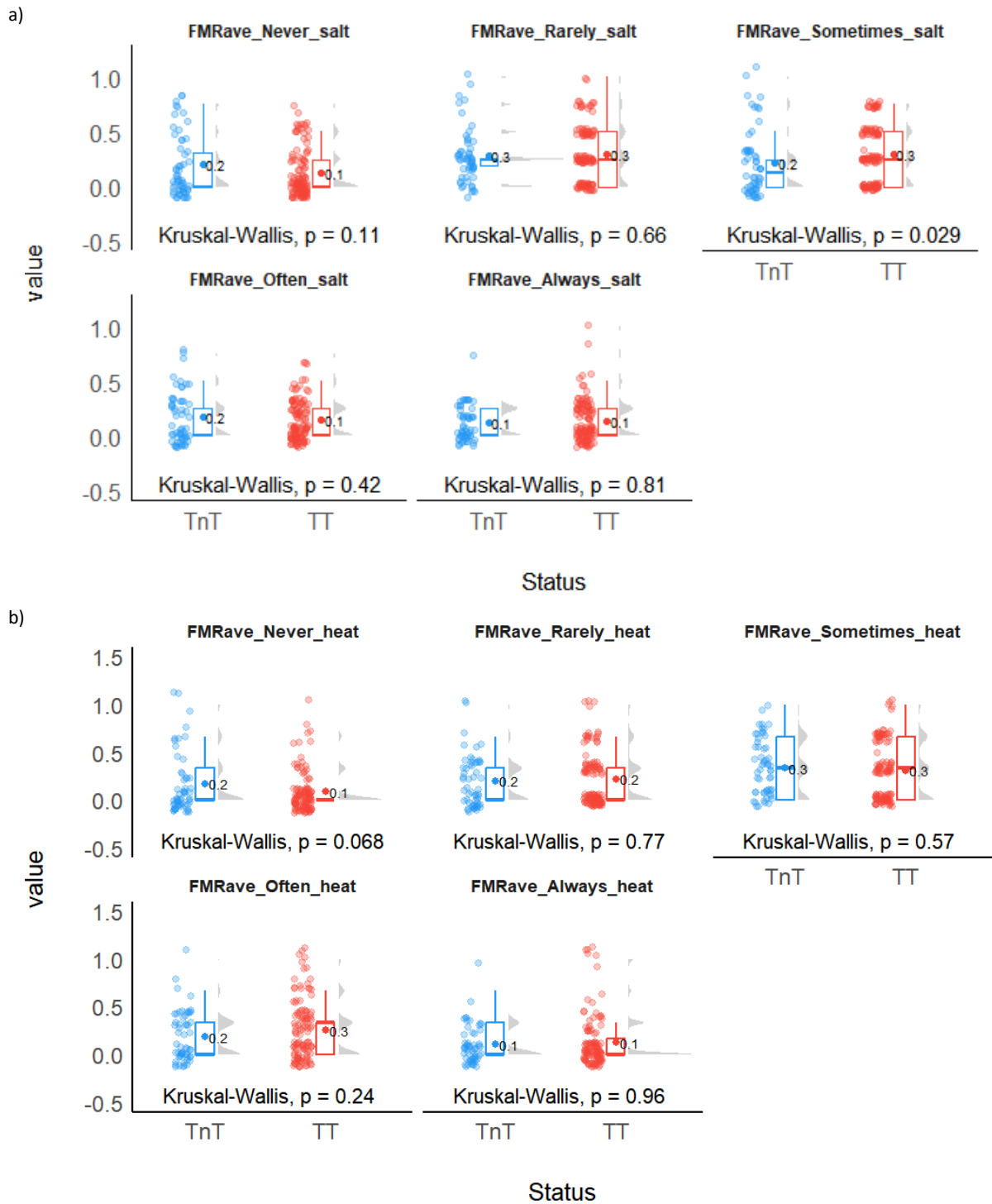


Figure 4-9 Distribution of the food modification raw data (indicated by prefix FMRave) regarding a) salt addition and b) heat addition according to TT Status, showing boxplots, means, summary statistics for the Kruskal-Wallis test.

Table 4-11 Calculated mean Food Modification responses included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.

Construct	Binary model (TT Status)	Multiclass model (TT Cluster)
Bitter masking	<ul style="list-style-type: none"> <li>• Mean “Never” bitter masking</li> </ul>	<ul style="list-style-type: none"> <li>• Mean combined “No” bitter masking</li> <li>• Mean combined “Yes” bitter masking</li> <li>• Mean “Sometimes” bitter masking</li> <li>• Mean “Rarely” bitter masking</li> <li>• Mean “Always” bitter masking</li> </ul>
Salt addition	<ul style="list-style-type: none"> <li>• Mean “Rarely” salt addition</li> </ul>	<ul style="list-style-type: none"> <li>• Mean combined “No” salt addition</li> <li>• Mean “Never” salt addition</li> <li>• Mean “Rarely” salt addition</li> <li>• Mean “Often” salt addition</li> <li>• Mean “Always” salt addition</li> </ul>
Heat addition	<ul style="list-style-type: none"> <li>• Mean combined “Yes” heat addition</li> <li>• Mean “Never” heat addition</li> </ul>	<ul style="list-style-type: none"> <li>• Mean combined “No” heat addition</li> <li>• Mean combined “Yes” heat addition</li> <li>• Mean “Rarely” heat addition</li> <li>• Mean “Always” heat addition</li> </ul>

addition, participants who rarely added salt to more than 25% of the items in question were more likely to be TT (Figure 4-10).

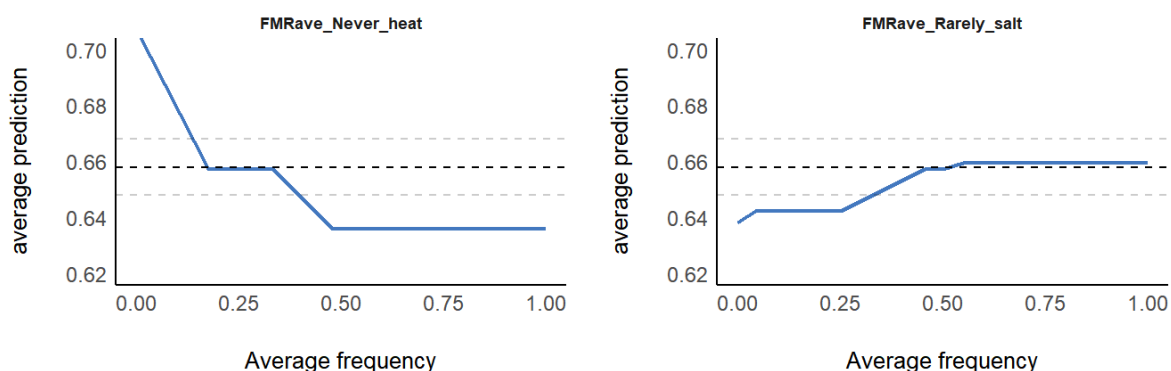


Figure 4-10 Accumulated local effects (ALE) plots for calculated indexes from the food modification questionnaire in the final binary (TT Status) RFM which impact TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. The prefix “FMR” indicates raw values (scale items not concatenated).

The response frequency to “Rarely” had the largest impact on TT Cluster probability. Participants who rarely added salt to fewer than 25% of consumed items were more likely to be a TnT and less likely to be part of Cluster 1 (Figure 4-11).

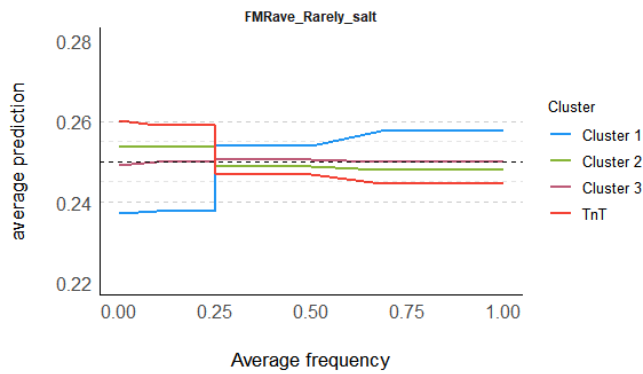


Figure 4-11 Accumulated local effects (ALE) plot for calculated raw (not concatenated) indexes from the food modification questionnaire included in the final multiclass (TT Cluster) RFM which impact TT Cluster probability. This plot shows chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

#### 4.3.4.1 Spicy food liking

There was no significant difference in overall spicy food liking between TT Status groups ( $p = .30$ ) or TT Clusters ( $p = .52$ ). Based on their performance in the preceding RFMs (Figure 4-1, Figure 4-3), three items from the spicy food liking questionnaire were included in the final TT Status RFM, and five in the final TT Cluster RFM (Table 4-12). Spice consumption frequency and the liking of the burn of spices were included in both final models. Note that liking the burn of spices was highly correlated to the overall spicy liking score, which was removed prior to RFM fitting (Table 4-8).

Table 4-12 Responses from the Spicy Food Liking questionnaire included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.

Construct	Binary model (TT Status)	Multiclass model (TT Cluster)
Spice consumption	• Spice consumption frequency	• Spice consumption frequency
Spice liking	• Liking of spice burn	• Liking of spice burn • Liking of spice taste
Spice opinions (T/F)	• "I think spices make food taste better"	• "Without hot spices, I find that food tastes bland" • "I find it hard to appreciate the flavours of food when the food contains hot spices"

Hedonic responses to spicy foods were useful predictors of TT Status – participants who enjoyed the burn of spicy foods and those who agreed that chilli makes food taste better were more likely to be TT (Figure 4-12). None of the selected items had an impact on TT Cluster probability.

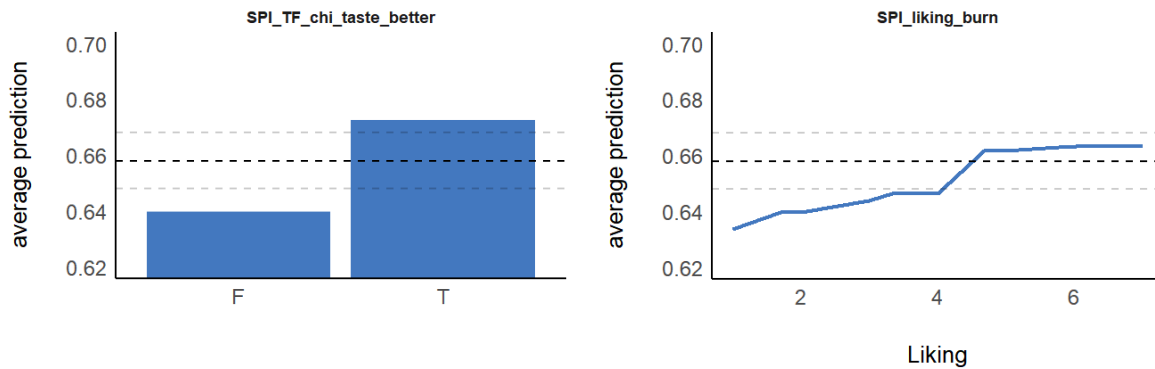


Figure 4-12 Accumulated local effects (ALE) plots for the items from the spicy food questionnaire included in the final binary (TT Status) RFM which impact TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

### 4.3.1 Personality (Survey 3)

#### 4.3.1.1 Big five inventory

TT were significantly more extraverted than TnT ( $p < .01$ ). TT were also more open to new experiences (Openness to experience  $p = .084$ ) (Figure 4-13). TT Clusters differed in Agreeableness ( $p = .063$ ), Conscientiousness ( $p = .072$ ) and Extraversion ( $p = .099$ ) (Figure 4-14). Participants in Cluster 1 were less Agreeable than Cluster 3 ( $p = .013$ ) and TnT ( $p = .073$ ). Cluster 3 was more Conscientious than all other clusters (Cluster 1  $p = .02$ , Cluster 2  $p = .037$ , TnT  $p = .033$ ). TnT were less extraverted than Cluster 1 ( $p = .033$ ) and Cluster 2 ( $p = .044$ ).

Based on their performance in the preceding RFMs (Figure 4-1, Figure 4-3), four variables from the BFI were automatically selected for the final TT Status RFM: Extraversion, two items related to Extraversion and one relating to Agreeableness (Table 4-13). In contrast, four of the five dimensions of the BFI (Agreeableness, Conscientiousness, Extraversion, and Openness to experience) were selected in the TT Cluster RFM and twenty-eight of the 44 items in the BFI (Table 4-13). Of these 28 items, four related to Agreeableness, five to Conscientiousness, six to Extraversion, seven to Neuroticism and six to Openness to experience.

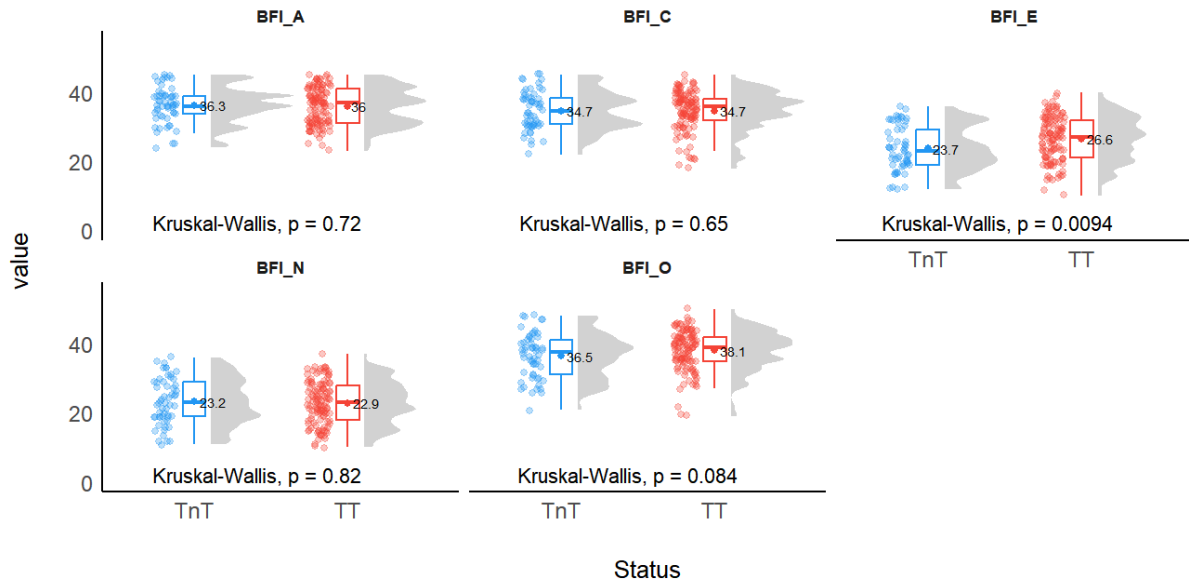


Figure 4-13 Distribution of five factors of the BFI (BFI\_A = Agreeableness, BFI\_E = Extraversion, BFI\_C = Conscientiousness, BFI\_N = Neuroticism, BFI\_O = Openness to experience) according to a) TT Status showing boxplots and means, as well as summary statistics for the Kruskal-Wallis test.

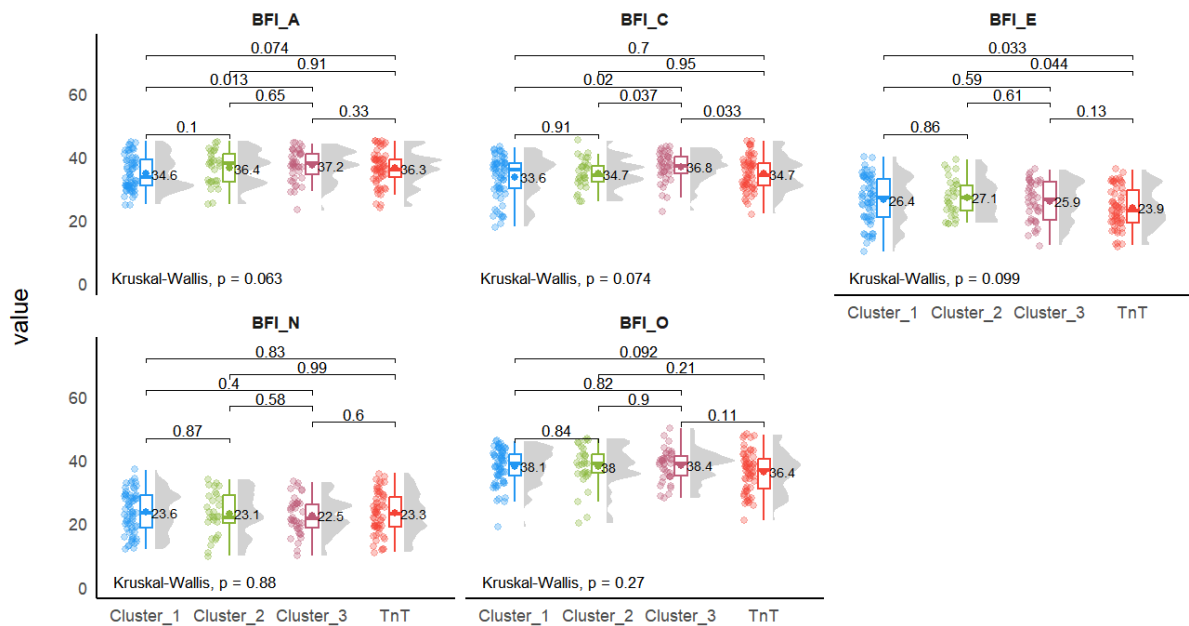


Figure 4-14 Distribution of five factors of the BFI (BFI\_A = Agreeableness, BFI\_E = Extraversion, BFI\_C = Conscientiousness, BFI\_N = Neuroticism, BFI\_O = Openness to experience) according to ) TT Cluster showing boxplots and means, as well as summary statistics for the Kruskal-Wallis test. and pairwise comparisons (results of the Wilcoxon-test) corrected for the false discovery rate (fdr)

Table 4-13 Variables and constructs from the Big Five Inventory personality questionnaire included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.

Construct	Binary model (TT Status)	Multiclass model (TT Cluster)
Agreeableness	<ul style="list-style-type: none"> <li>• Tends to find fault with others.</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Agreeableness</i></li> <li>• Is considerate and kind to almost everyone.</li> <li>• Likes to cooperate with others.</li> <li>• Starts quarrels with others.</li> </ul>
Conscientiousness		<ul style="list-style-type: none"> <li>• <i>Conscientiousness</i></li> <li>• Can be somewhat careless.</li> <li>• Is easily distracted.</li> <li>• Does a thorough job.</li> <li>• Makes plans and follows through with them.</li> <li>• Perseveres until the task is finished.</li> </ul>
Extraversion	<ul style="list-style-type: none"> <li>• <i>Extraversion</i></li> <li>• Has an assertive personality.</li> <li>• Tends to be quiet.</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Extraversion</i></li> <li>• Has an assertive personality.</li> <li>• Tends to be quiet.</li> <li>• Is full of energy.</li> <li>• Generates a lot of enthusiasm</li> <li>• Is outgoing, sociable.</li> <li>• Is sometimes shy, inhibited</li> </ul>
Neuroticism		<ul style="list-style-type: none"> <li>• Is depressed, blue.</li> <li>• Can be moody.</li> <li>• Gets nervous easily</li> <li>• Is relaxed, handles stress well.</li> <li>• Remains calm in tense situations.</li> <li>• Can be tense.</li> <li>• Worries a lot</li> </ul>
Openness to experience		<ul style="list-style-type: none"> <li>• <i>Openness to experience</i></li> <li>• Has an active imagination</li> <li>• Is ingenious, a deep thinker.</li> <li>• Is inventive.</li> <li>• Prefers work that is routine.</li> <li>• Is sophisticated in art, music, or literature.</li> <li>• Values artistic, aesthetic experiences.</li> </ul>

Only the items related to Extraversion were useful in predicting TT Status. Participants who were more extraverted were more likely to be TT (Figure 4-15). In addition, participants who described their character as assertive were more likely to be TT, and those who described themselves as quiet were less likely to be TT (Figure 4-15).

Of the 28 items automatically selected for the TT Cluster RFM, only four were useful predictors of TT Cluster, including the Conscientiousness dimension, an item from Agreeableness and two from Extraversion (Figure 4-16). Participants scoring lower in Conscientiousness were less likely to form part of Cluster 1. Participants disagreeing with the statement “I am someone who is sometimes rude to others” were more likely

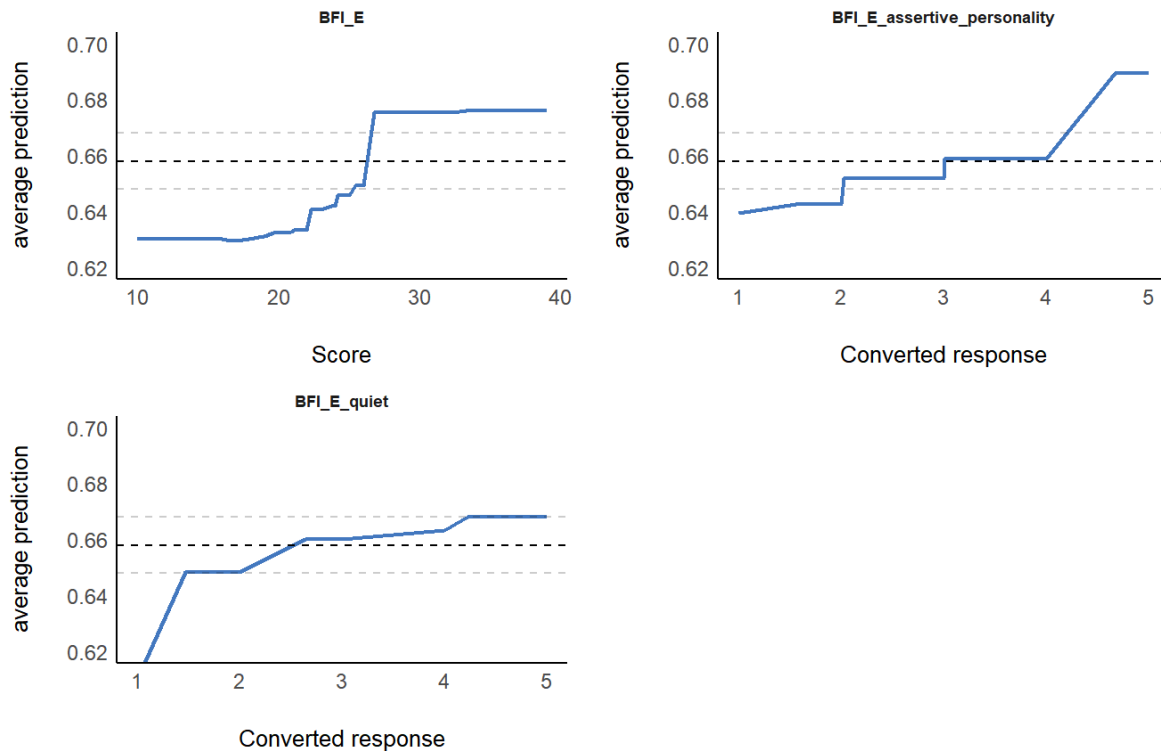


Figure 4-15 Accumulated local effects (ALE) plots for personality factors and items from the Big Five Inventory (BFI) included in the final binary (TT Status) RFM which impact TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. Note that individual items are prefixed by the factors to which they contribute.

to be part of Cluster 1 and less likely to be part of Cluster 3 or to be TnT. The items loading to Extraversion (responses to “I am someone who is sometimes shy, inhibited” and “I am someone who tends to be quiet”), impacted class probability, followed a similar pattern; participants disagreeing with either statement were more likely to be TnT and less likely to form part of Cluster 1 (Figure 4-16).

#### 4.3.1.1 Private body consciousness

There was no difference in Private Body Consciousness between TT Status groups ( $p = .30$ ) or TT Clusters ( $p = .32$ ).

Based on their performance in the preceding RFMs (Figure 4-1, Figure 4-3), all five items from the Private Body Consciousness questionnaire and Private Body Consciousness (PBC) overall were included in the TT Status RFM, and four of the five items were included in the TT Cluster RFM (Table 4-14).

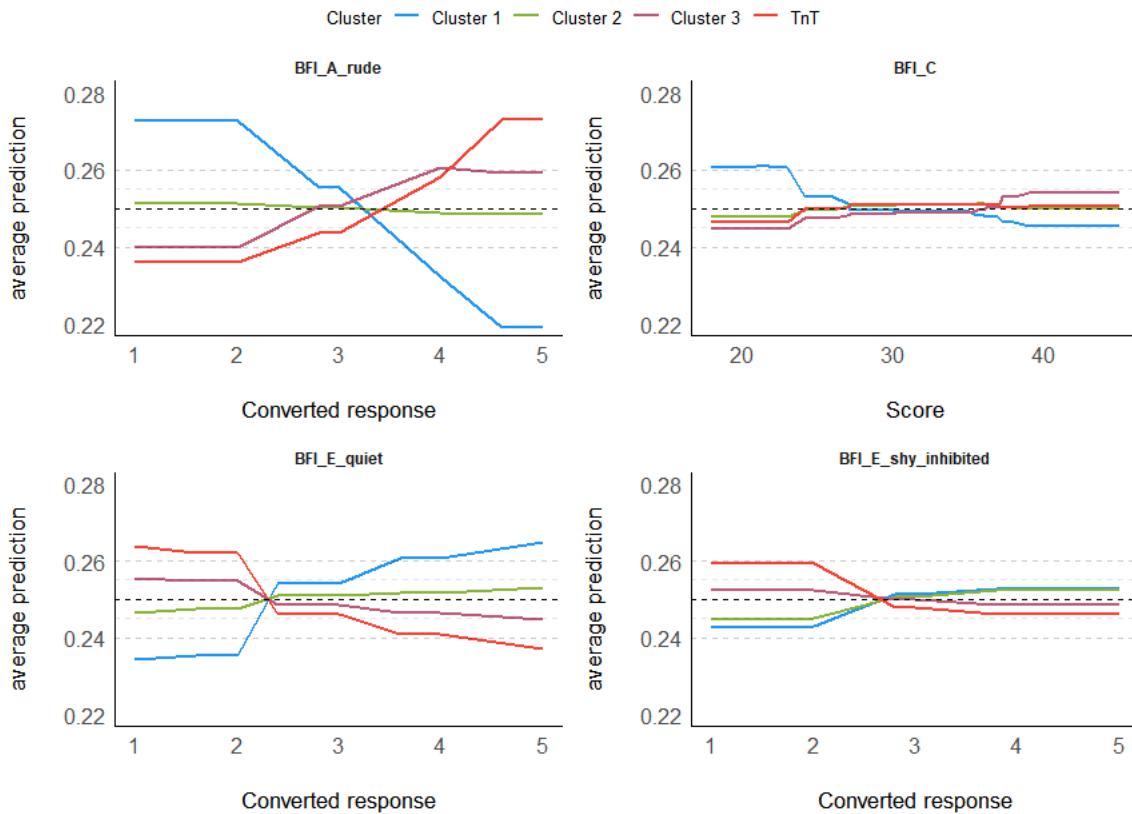


Figure 4-16 Accumulated local effects (ALE) plots for personality items and factors from the Big Five Inventory (BFI) included in the final multiclass (TT Cluster) RFM impacting cluster probability. These plots show chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that individual items are prefixed by the factors to which they contribute.

Table 4-14 Variables from the Private Body Consciousness questionnaire included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.

Construct	Binary model (TT Status)	Multiclass model (TT Cluster)
Private body consciousness	<ul style="list-style-type: none"> <li>• Private body consciousness</li> <li>• I am quick to sense the hunger contractions of my stomach.</li> <li>• I can often feel my heart beating.</li> <li>• I can often feel my heart beating.</li> <li>• I know immediately when my mouth or throat gets dry.</li> <li>• I am sensitive to internal bodily tensions.</li> </ul>	<ul style="list-style-type: none"> <li>• I am quick to sense the hunger contractions of my stomach.</li> <li>• I can often feel my heart beating.</li> <li>• I know immediately when my mouth or throat gets dry.</li> <li>• I am sensitive to internal bodily tensions.</li> </ul>

Scoring higher in PBC overall, being sensitive to internal body tensions and being quick to sense hunger contractions were predictors of TT Status (Figure 4-17), but this effect was small compared to other items in the model. None of the items from the PBC questionnaire were useful in predicting TT Cluster.

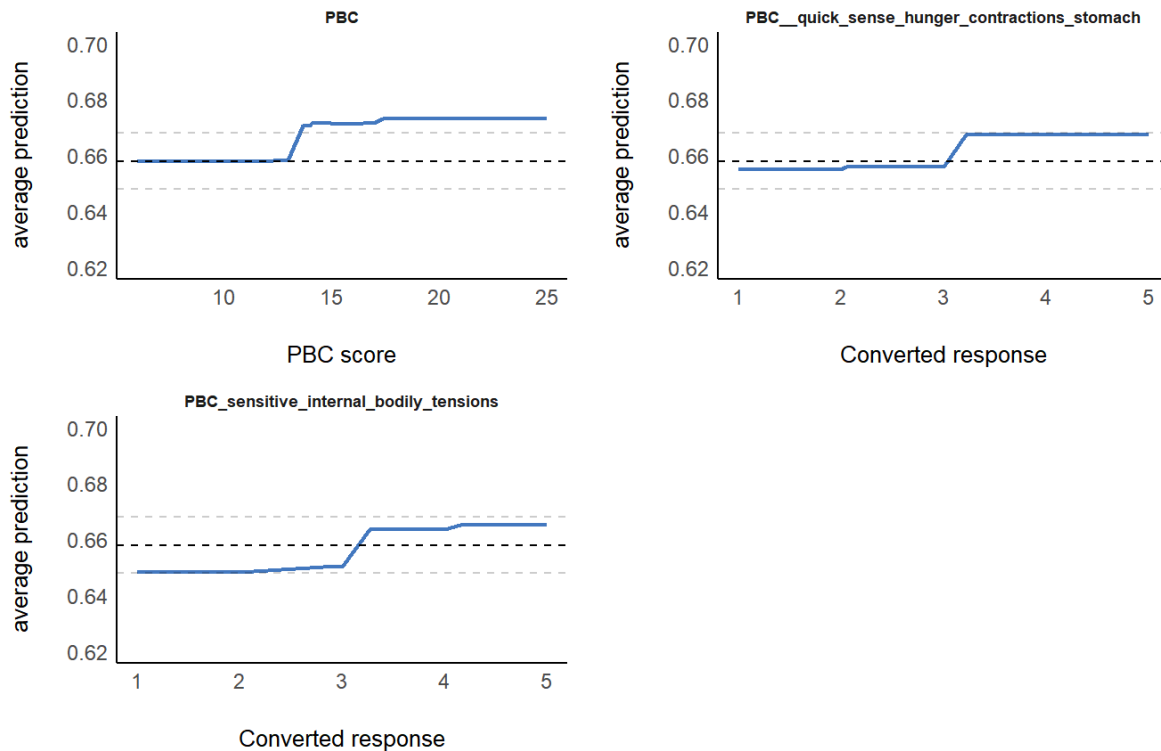


Figure 4-17 Accumulated local effects (ALE) plots for Private body consciousness (PBC) and the items from the PBC questionnaire included in the final binary (TT Status) RFM impacting TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

### 4.3.2 Orosensory response (Taste Phenotyping)

#### 4.3.2.1 Sweet liker status (SLS)

Participants exhibited three distinct patterns of sweet liking, resulting in four SLS clusters: one cluster per liking pattern and a fourth cluster of participants who were inconsistent in their responses and were difficult to classify. The sucrose-liking pattern of SLS Cluster 1 (Medium sweet likers or MSL,  $n = 46$ ) followed an inverted U, with their optimal sucrose concentration falling at approximately 10 g/L. SLS Cluster 2 (Low Sweet Likers, LSL;  $n = 74$ ) preferred lower sucrose concentrations. SLS Cluster 3 (High Sweet Likers, HSL;  $n = 28$ ) preferred higher sucrose concentrations. In SLS Cluster 4 (Unclassified or UN;  $n = 13$ ), the participants with no affinity for any specific cluster showed no clear pattern in sucrose liking.

No relationship existed between SLS cluster and TT Status. However, participants in TT Cluster 1 (which reported predominantly metallic and salty thermal taste) were more likely to be HSL/Cluster C ( $p = .043$ ).

Based on their performance in the preceding RFMs (Figure 4-1, Figure 4-3), hedonic responses to two sucrose concentrations were the only two sucrose response items automatically selected for the final RFM for TT Status (Table 4-15). In contrast, SLS Cluster, the hedonic response to all five sucrose solutions, the sweetness intensity of three solutions, and the maximum liking of sucrose were included in the final model for TT Cluster (Table 4-15).

*Table 4-15 Variables from Sweet Liker Status phenotyping (including hedonic and intensity responses) included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.*

<b>Construct</b>	<b>Binary model (TT Status)</b>	<b>Multiclass model (TT Cluster)</b>
Sweet liking	<ul style="list-style-type: none"> <li>• Liking 24% sucrose</li> <li>• Liking 6% sucrose</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Sweet liking cluster</i></li> <li>• Liking 12% sucrose</li> <li>• Liking 24% sucrose</li> <li>• Liking 3% sucrose</li> <li>• Liking 36% sucrose</li> <li>• Liking 6% sucrose</li> <li>• Most liked sucrose concentration</li> </ul>
Sweetness intensity		<ul style="list-style-type: none"> <li>• Intensity 12% sucrose</li> <li>• Intensity 24% sucrose</li> <li>• Intensity 36% sucrose</li> </ul>

Both items selected for the TT Status RFM were useful for prediction. Participants who scored the 6% solution higher than 55 on the LAM were much more likely to be TT, and those who scored this solution less than 55 were less likely to be TT. Similarly, participants who scored the 24% sucrose solution higher than 27 on the LAM were more likely to be TT, and those who scored less than 27 were less likely to be TT. However, the 24% solution was less useful in predicting TT Status than the 6% solution (Figure 4-18).

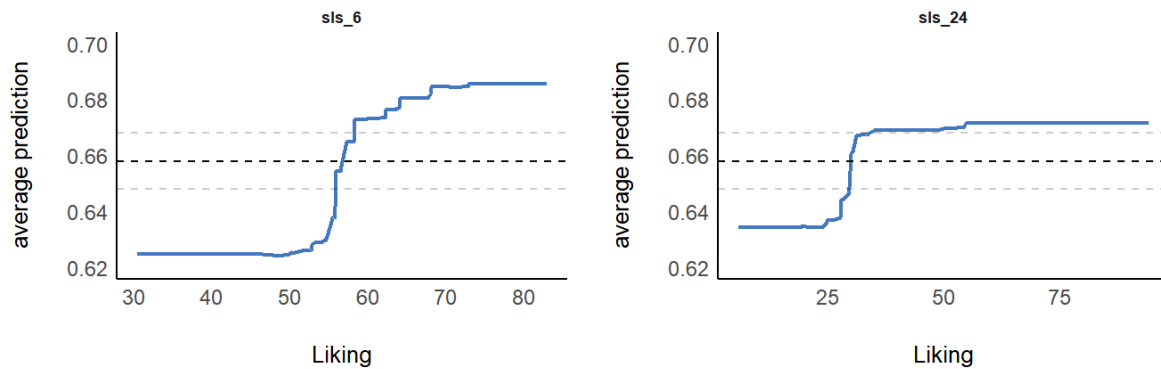


Figure 4-18 Accumulated local effects (ALE) plots for sweetness response variables (liking, intensity, SLS cluster) included in the final binary (TT Status) RFM impacting TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm 0.01$ ), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. Note that the prefix “sls” indicates liking on the LAM scale.

Hedonic response to the 36% sucrose solution was the only variable significantly impacting cluster prediction. Scoring this solution higher than 50 on the LAM increased Cluster 1 probability – the inverse was true for liking below 50 (Figure 4-19).

#### 4.3.2.2 PROP responsiveness

Twenty participants were PNT, 93 were PMT and 45 were PST. Neither TT Status ( $p = .64$ ) nor TT Cluster ( $p = .71$ ) was related to PROP status classification. In addition, there was no difference in PROP intensity rating based on TT Status or TT Cluster.

PROP Status and reported Bitterness intensity from PROP were intentionally included in both final RFMs, as these were the only two variables representing a construct of interest. Bitterness intensity reported from PROP exposure (Figure 4-20) impacted TT probability; participants who scored PROP bitterness lower than the first quartile on the intensity scale were less likely to be TT. PROP taster status showed a similar trend; participants in the PMT and PST categories were slightly more likely to be TT, although this effect was relatively small (overall increase in probability less than 0.1; Figure 4-20). In contrast, neither PROP intensity nor PROP taster status impacted TT Cluster probability.

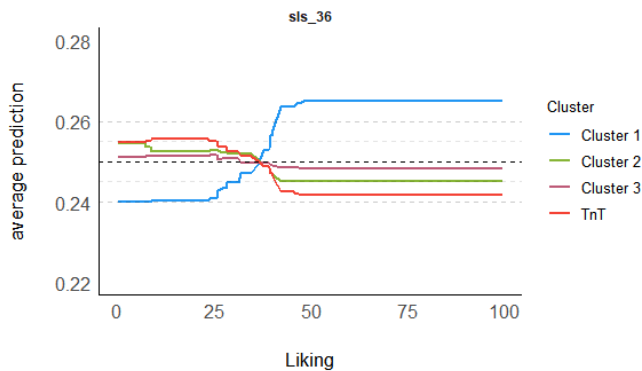


Figure 4-19 Accumulated local effects (ALE) plots for sweetness response variables (liking, intensity, SLS cluster) included in the final multiclass (TT Cluster) RFM impacting cluster probability. This plot shows chance probability (0.25; baseline cluster probability), four guides (chance  $\pm 0.01$  and chance  $\pm 0.005$ ), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that the prefix “sls” indicates liking on the LAM scale.

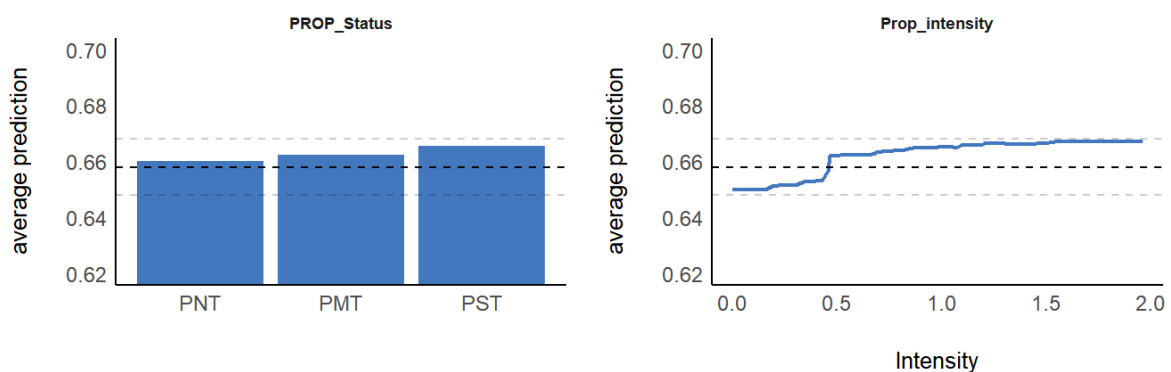


Figure 4-20 Accumulated local effects (ALE) plots for PROP responsiveness variables included in the final binary (TT Status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm 0.01$ ), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

#### 4.3.2.3 Fatty acid responsiveness

Neither TT Status ( $p = .97$ ) nor TT Cluster ( $p = .80$ ) affected fatty acid detection threshold.

Based on their performance in the preceding RFMs (Figure 4-1, Figure 4-3), responses to four of the fatty acid samples (level 1, level 4, level 5, level 6). were automatically selected for the final TT Status RFM. In addition, responses to six of the eight fatty acid levels were included in the final TT Cluster RFM (Table 4-16 ).

Table 4-16 Variables (including responses to specific oleic acid concentrations) related to fatty acid responsiveness included in the final binary (TT Status) and multiclass (TT Cluster) RFMs. Variables were conservatively selected based on their importance in the two previous RFMs.

Construct	Binary model (TT Status)	Multiclass model (TT Cluster)
Response bias		• Response Blind control
Fatty acid response	<ul style="list-style-type: none"> <li>• Response 0.5 mM/L oleic acid (L1)</li> <li>• Response 4 mM/L oleic acid (L4)</li> <li>• Response 8 mM/L oleic acid (L5)</li> <li>• Response 16 mM/L oleic acid (L6)</li> </ul>	<ul style="list-style-type: none"> <li>• Response 0.5 mM/L oleic acid (L1)</li> <li>• Response 4 mM/L oleic acid (L4)</li> <li>• Response 8 mM/L oleic acid (L5)</li> <li>• Response 16 mM/L oleic acid (L6)</li> <li>• Response 32 mM/L oleic acid (L7)</li> </ul>

Responses to the lowest concentration of oleic acid (level 1) did not affect TT probability. In contrast, the responses to the remainder of the samples all followed a similar pattern; participants who rated the sample zero (no difference) were less likely to be TT, whereas those who scored anything higher than zero were slightly more likely to be TT. This effect is most pronounced for Level 4 (Figure 4-21).

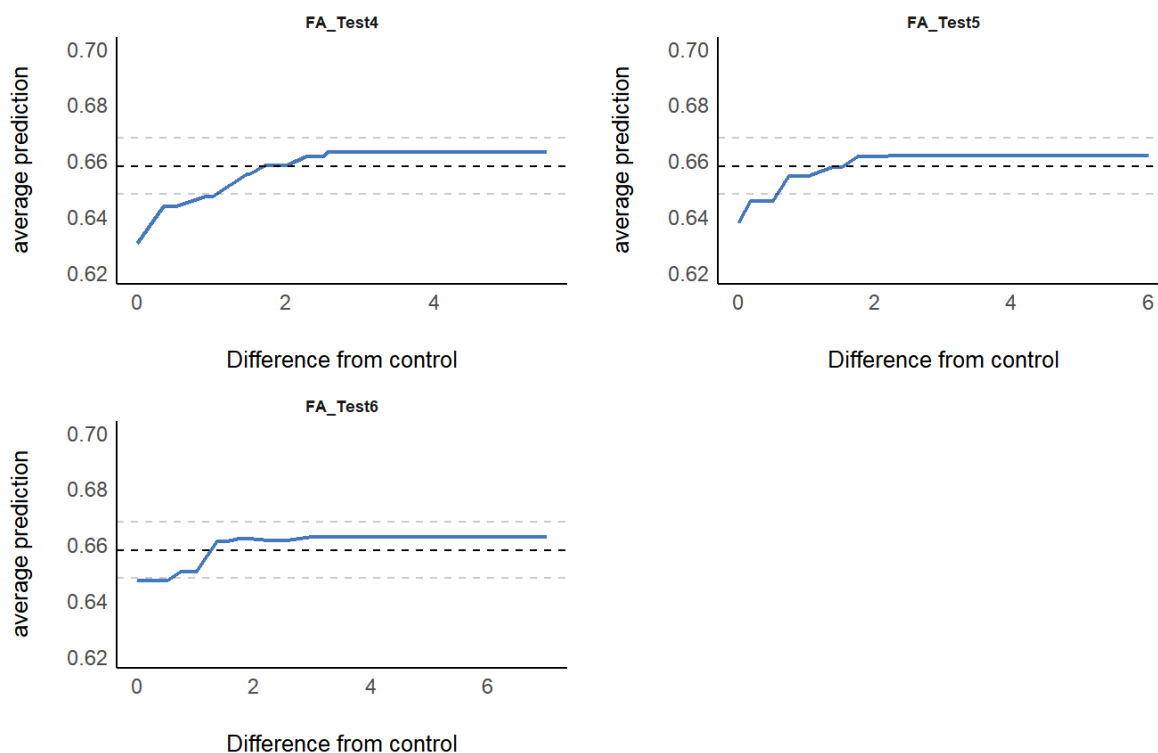


Figure 4-21 Accumulated local effects (ALE) plots for Fatty acid response variables included in the final binary (TT Status) RFM, impacting TT probability. These plots show the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

Responses to level 6 and level 7 of the FA samples had the greatest effect on class prediction (Figure 4-22). Participants who responded, “no difference” (0) to level 6

were more likely to be Cluster 3 or TnT and less likely to be Cluster 2, although the participant response to this sample did not affect Cluster 1 probability. In contrast, participants who responded, “no difference” (o) to sample 7 were less likely to be classified as Cluster 1.

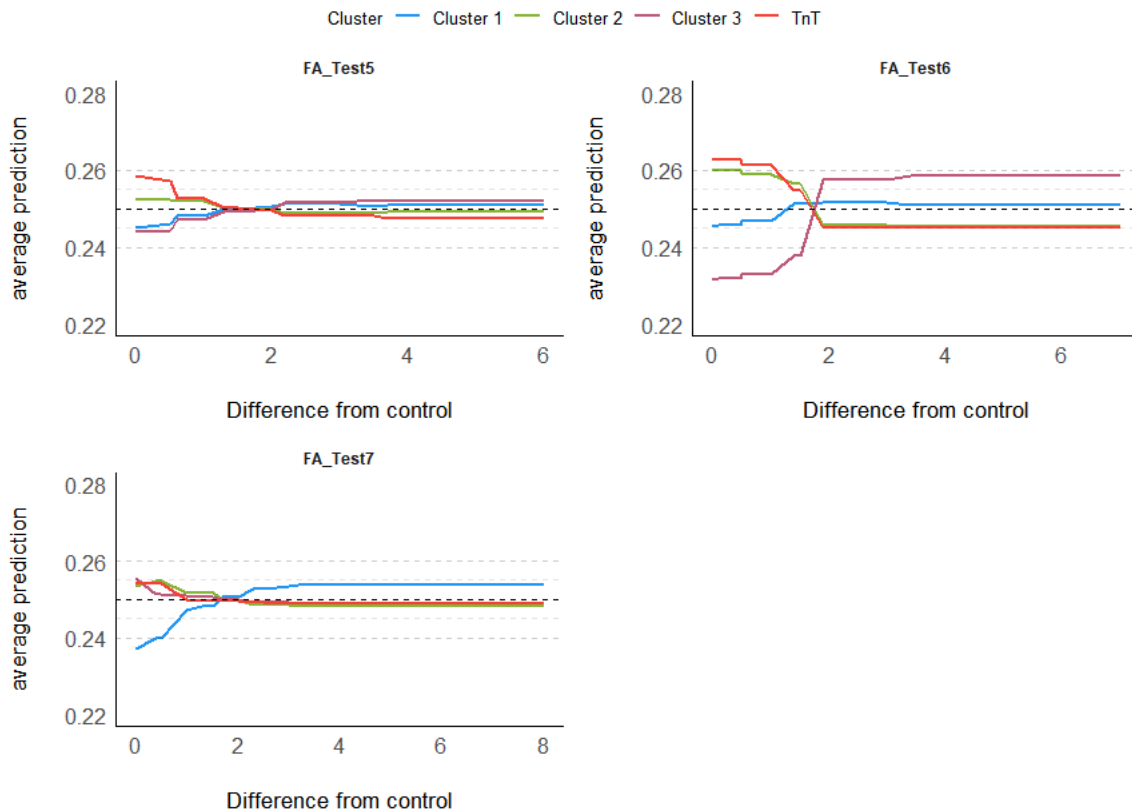


Figure 4-22 Accumulated local effects (ALE) plots for Fatty acid response variables included in the final multiclass (TT Cluster) RFM impacting cluster probability. These plots show chance probability (0.25; baseline cluster probability), four guides (chance +/- 0.01 and chance +/- 0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

#### 4.3.1 Results summary

Overall, small differences were found between TT and TnT in various aspects, with some similarities between specific TT clusters and TnT. TT tended to be slightly older than TnT, but age did not predict TT status using the ensemble RFM. Food behaviour differences included oral processing habits and food neophobia, with TT more likely to add salt and spices to food. TT clusters differed in food involvement, specifically cleaning behaviour after eating. TT tended to be more extraverted and open to new

experiences, with differences in agreeableness and conscientiousness among TT clusters. Orosensory responsiveness, such as liking for sucrose concentrations and bitter response to PROP, was predictive of TT status. TT clusters showed differences in hedonic responses to sucrose concentrations and fatty acid sensitivity.

#### **4.4 Discussion**

Overall, in addition to confirming relationships to factors previously found regarding TT Status, this chapter discovered new relationships and identified nuanced relationships between TT subgroups. In addition, the process of employing an RFM allowed for the contextualisation of different effects. This discussion first explores the RFM generally, and then explores the specific aspects explored experimentally.

##### **4.4.1 Modelling discussion**

Overall, there were several differences between the findings of the traditional statistical tests and the RFMs. These discrepancies took two forms. Firstly, there were some overall measures that were significantly different between TT groups but were not important predictors in the RFM. Secondly, there were some variables which were not significantly different between TT groups but were important predictors in the RFM.

The first case, namely the lack of inclusion of significantly different predictor variables is likely because of a better predictive effect of the variables that constitute that measure. For example, Agreeableness on the BFI was significant for TT Cluster, but this measure was not included in the model. However, an item loading to that scale (relating to rudeness) was a key predictor in the RFM, indicating a stronger relationship between that item and TT Cluster. Although the overall measures were not significantly different, this effect was demonstrated by other variables in the form

of Food Neophobia and PBC for TT Status. Note that this study specifically chose not to compare the individual variables between groups, as the vast number of comparisons would be data-dredging, considered a form of p-hacking (Streiner, 2015).

The second case, where predictors are important in the RFM without being significantly different was likely a result of more complex patterns being present in the data than that detected by significance testing. For example, for Food Neophobia and TT Status, the overall scores were not significantly different between groups. However, the variable was a notable predictor in the RFM. When considering the data distribution (Figure G-1, Appendix G), it is evident that there is a more complex pattern. In this case, TnT had a bimodal distribution, with a group of “high Neophobia” TnT and a group of “low Neophobia” TnT. Similarly, for Private Body Consciousness, and TT Status, although there were no significant differences in the means, there was a difference in the shape of the distribution; TnT had a relatively flat distribution, with TT having a much clearer mode, as indicated in (Figure G2, Appendix G).

Overall, these discrepancies between traditional statistics highlight some of the strengths of combining an RFM with interpretable machine learning over traditional approaches. Firstly, the use of an RFM allowed for the contextualisation of effects overall. Secondly, the use of a unified (ensemble) machine learning model allowed for the evaluation of all potential factors, without overreliance on p-values and becoming vulnerable to their pitfalls, including the arbitrary nature of choosing a cut-off value (Wasserstein et al., 2019). Finally, the approach taken is more sensitive to patterns in data overall, not just taking into account central tendency.

#### **4.4.2 Relationships between variables of interest and thermal taste**

##### *4.4.2.1 Demographics (Age, ethnicity, gender)*

Although not significant, TT in this study tended to be older than TnT, in contrast with Thibodeau et al. (2019), who found TT to be slightly younger than TnT. However, the cohort in Thibodeau et al. (2019) was younger than those included in the present study, with mean ages of TT and TnT of 24.8 and 27, respectively, compared to 38.8 and 35.1 in the present study, suggesting a complex age-dependent interaction between TT Status and age. One explanation of the difference in age-related findings between Thibodeau et al. (2019) and the present study is age-related differences in working memory. Working memory capacity declines with age (Rypma et al., 2001), and the lower working-memory requirements of RapCoTT (as explored in Chapter 2) could mean that in the present study, older TT were more likely to be identified, whereas in previous studies these TT may have been unclassified.

Methodological differences, however, do not explain why, in the present study, TT were older than TnT. A possible explanation is that trigeminal learning happens throughout the lifetime, and repeated exposure to trigeminal stimuli enhances the ability to perceive thermal taste. Typically, taste sensitivity decreases with age (Jilani et al., 2022; Mojet et al., 2001; Mojet et al., 2003; Sato et al., 2022; Yang et al., 2022), but this phenomenon does not occur with pungent stimuli (Yang et al., 2022). In addition, exposure to trigeminal stimuli can increase sensitivity to them (Oleszkiewicz et al., 2018), which means that an aspect of the thermal taste response might be learnt through exposure to relevant stimuli.

The present study found no relationship between TT and gender/sex, in agreement with previous studies looking at TT and gender (Bajec & Pickering, 2008; Pickering &

Kvas, 2016; Pickering et al., 2016; Thibodeau et al., 2019; Yang, 2015). Gender is a complex interaction between biological and social factors impacting brain development (Rippon, 2019). Interestingly, the present study found differences between TT subgroups in aspects of food involvement, some of which reflect gendered behaviours (Yang, 2015). In addition, as thermal taste has been thought of as a type of crossmodal interaction between temperature and taste (Spence, 2020), gender differences in crossmodal interaction (Sijben et al., 2021) would lead to expected gender differences between TT and TnT. However, a lack of gender difference between TT and TnT indicates that the effects of any correlates with gender are less important in influencing thermal taste.

This study failed to replicate (calculated effect size  $V = 0.0$  for both potential relationships) the strong relationship (calculated effect size  $V = 0.24$ ) previously found between Asian ethnicity and TT Status (Yang et al., 2020). These differences could be related to differences between the populations in the present study and Yang et al. (2020). Yang et al. (2020) conducted their research in the United Kingdom, in contrast with the present research conducted in New Zealand. The Asian populations in the two studies likely have two different socio-ethnic and genetic makeups and different degrees of underlying diversity. In other words, a specific subgroup of the highly diverse group which is covered by Asian ethnicity (Heyer et al., 2009) more likely to be TT may have been better represented by the Asian sample in Yang et al. (2020), but underrepresented in this study. Notably, the majority of the Asian sample in Yang et al. (2020) was Chinese (54 out of 67 participants). Although more detailed ethnicity data was not collected in this study, higher diversity can be inferred, as roughly only 35% of the NZ Asian population is Chinese (Stats NZ, 2020).

#### 4.4.2.2 Food Behaviour

##### 4.4.2.2.1 Oral processing behaviour

This was the first study to explore relationships between oral processing behaviour and thermal taste, finding no significant overall relationships between oral processing behaviour subgroups and TT Status or TT Cluster. However, being a part of the “Cruncher” oral processing behaviour subgroup predicted TT Status, a finding which agrees broadly with the previous finding (Bajec & Pickering, 2010) that TT liked “mushy” foods significantly less than TnT. Although oral processing behaviour subgroups are not directly related to the liking of specific foods and textures but rather the modification of food textures to suit preferred behaviour (Jeltema et al., 2015; Jeltema et al., 2020; Jeltema et al., 2014), it is difficult to modify mushy items to suit the preferred behaviour of crunching, validating the overall agreement. This relationship between thermal taste and texture preference indicates the importance of further investigation into thermal taste and texture sensitivity and preference, especially as two mechanoreceptors were linked with thermal taste in Chapter 3, namely TRPA1 (Castiglioni & García-Añoveros, 2007) and TRPV4 (Suzuki et al., 2003).

##### 4.4.2.2.2 Food involvement

Although there were no significant differences in food involvement or any of the subscales between TT and TnT, scores on the S & D subscale differentiated TnT from Cluster 1, and was, notably, the strongest predictor in this study. This predictive pattern mirrors the item “I do most of the cleaning up after eating”, suggesting a surprising link between cleaning after eating and a specific (large) group of thermal tasters. One aspect of cleaning up after eating is dishwashing, and the relationship may be related to differences in (non-taste) sensory experiences of dishwashing between

certain TT and TnT. This experience may be more unpleasant to some TT due to potential differences in wetness sensation. Moistness or wetness is associated with negative valence (Thibodeau, 2016), and the sensation of wetness is a combination of thermosensation and mechanosensation (Kumar & Kumar, 2021). Although a previous study showed TT not to be more sensitive to uniform temperature exposure on the palm (Green & George, 2004), TT were more sensitive to the temperature treatment of the warming (traditional) phenotyping treatment (Bajec & Pickering, 2008; Hort et al., 2016; Thibodeau et al., 2019). Different temperature sensitivities might extend to different wetness (and dishwashing) experiences, making some TT less likely to clean up after eating. However, as cleaning after eating does not only involve handwashing dishes, this aspect of difference requires further exploration.

#### 4.4.2.2.3 Food Neophobia

Although TT did not differ from TnT for FNS score, low food neophobia predicted TT Status. The tendency to sample new foods was also a predictor of TT Status (binary model) and a negative predictor of TnT group membership (multiclass model). These findings agree broadly with previous findings of no overall significant differences in Food Neophobia between TT and TnT (Bajec & Pickering, 2010; Yang, 2015) and with the trend found by Yang (2015) that TT tend to be less food neophobic than TnT. Yang (2015) also found TT significantly more likely to sample new foods, agreeing with the present study's findings. In addition, Pickering and Pickering (2022) found TT more likely to be "foodies", one aspect of which is sampling new foods. Overall, this indicates a tendency towards food experimentation in TT, which is counterintuitive given the heightened orosensitivity of TT. However, TT could experience heightened enjoyment from heightened orosensation, making them more open to potential pleasant food experiences.

#### 4.4.2.2.4 Food modification

There were subtle differences in food modification behaviour between TT and TnT, and the addition of salt to food was predictive of TT Status and TT Cluster – rarely adding salt to half of the food items was predictive of TT Status, and rarely adding salt to less than 25% of the items made participants more likely to be part of Cluster 1 and less likely to be TnT. In other words, Cluster 1, who were more likely to report Salty thermal taste, were less likely to add salt to their food. Although no relationship was previously found between specific taste sensitivity and reported thermal taste (Thibodeau et al., 2020), this behavioural construct indicates an interaction between the nature of thermal taste reported and the overall food experience.

The present study found no difference between groups in bitterness-masking behaviour (through the addition of sweetening agents in general or milk to tea/coffee), and no aspect of this behaviour was predictive of group membership. Puputti, Hoppu, et al. (2019) found that bitterness-masking behaviour was more prevalent in groups with higher orosensitivity, suggesting that TT Status as measured through RapCoTT may not be an overall predictor of orosensitivity. As previously mentioned, it is possible that the memory load requirement for traditional TT phenotyping only caused higher-intensity thermal taste sensations to be retained (Liang et al., 2018), self-selecting a group with higher orosensitivity as thermal tasters.

#### 4.4.2.2.5 Spicy food liking (including heat addition)

Spice addition to food was predictive of TT Status, as participants who did not refrain from adding spices to any items were more likely to be TT. In addition, the enjoyment of the burn of spices and agreeing with the statements that adding spices makes foods taste better was predictive of TT Status. Traditionally phenotyped TT are more sensitive to capsaicin (Yang et al., 2014), and infrequent users rate the burn of

capsaicin more intensely (Nolden & Hayes, 2017). However, the intensity effect found by Nolden and Hayes (2017) does not extend to cross cultural comparisons taking differences in chilli consumption into account (Berry & Simons, 2020). In addition, chilli users are more sensitive to thermal stimuli (Ludy & Mattes, 2012), aligning with previous evidence (Thibodeau et al., 2019) that TT tend to be more sensitive to oral heat. It is possible that, although TT perceive capsaicin more intensely, they experience more crossmodal enhancement of taste (Han et al., 2021; Wang et al., 2022), aroma (Petit et al., 2007; Spencer & Dalton, 2019) and texture (Lyu et al., 2021) qualities from spices overall.

Overall, this finding further supports the idea that an aspect of the thermal taste response is an aspect of chemosensory/trigeminal learning. Neural correlates have recently been found with spicy food consumers, with increased grey matter areas likely reflecting sustained and clearer mental imagery of pungent sensations and spicy flavours (Han et al., 2020; Zhou et al., 2023). These differences may extend to the ability to recognise and report thermal taste.

#### *4.4.2.3 Personality*

##### *4.4.2.3.1 Big Five Inventory*

The finding that TT scored higher in Extraversion overall contrasts with Yang (2015), who found no relationship between Extraversion and TT Status. The relationship between Extraversion and TT Status in the present study might result from the experimental setup. Having an assertive personality and not being quiet were predictive of TT Status, suggesting that participants who were assertive and comfortable expressing themselves (not quiet) may have been more comfortable responding to the RapCoTT/TCATA paradigm in real-time. Yang (2015) required

participants to respond with taste sensations after completing the trial, allowing introverted participants to reflect on their experiences and respond. This finding points to an area for additional improvement in the RapCoTT protocol - adding additional steps or measures to ensure *all* participants, regardless of temperament, are comfortable responding in real-time.

TT also scored higher overall in Openness to experience, which agrees with previous research (Yang, 2015), with one cluster (C3) in particular that scored higher in the Openness dimension than TnT. This link, in combination with the findings related to Extraversion, could be connected to the differing abilities of participants to form mental images. Although Chapter 3 presented evidence that thermal taste is biologically induced, differences in translating that biological sensation into a taste experience may have similar neural underpinnings as the ability to imagine other sensory sensations. When comparing ratings on the Plymouth Sensory Imagery Questionnaire (Andrade et al., 2014) with the BFI, participants who were better at imagining sensory sensations scored higher on Openness to experience, Extraversion and Agreeableness (Hitsuwari & Nomura, 2023). This connection further supports the idea that thermal taste is both biological and neurocognitive in nature.

#### 4.4.2.3.2 Private Body Consciousness

Although there was no significant difference in Private Body Consciousness (PBC) between TT and TnT, PBC and self-reported sensitivity to internal bodily tensions contributed to the prediction of TT Status, indicating that some differences in PBC exist between TT and TnT. The PBC questionnaire measures interoceptive sensibility (the awareness of bodily sensations), which is distinct from interoceptive accuracy (the ability to sense bodily sensations accurately) and interoceptive awareness (confidence in awareness of bodily sensations) (Garfinkel et al., 2015). Participants with higher

interoceptive sensitivity may be more confident in their interoceptive accuracy and, therefore, more likely to report thermal taste when perceived.

There may be a direct neural link between interoceptive accuracy and thermal taste, as traditionally phenotyped thermal tasters showed higher activity in the insula when exposed to cold, sweet samples (Eldeghaidy et al., 2021) – a region where higher functional connectivity has been linked to higher interoceptive accuracy (Chong et al., 2017). Note, however, that this same region has not been linked with interoceptive sensibility (Smith et al., 2022), although different questionnaires of interoceptive sensibility measure different constructs (Vig et al., 2022).

Although Yang (2015) investigated differences in alexithymia, the inability to perceive and interpret emotions, in TT and TnT (using the Toronto Alexithymia Scale), the present study is the first to investigate links between interoception and thermal taste. Interestingly, a meta-analysis (Barańczuk, 2019) showed alexithymia to have an opposite relationship to the five-factor personality than the relationship between TT and the BFI in the present study. In addition, alexithymia is considered a generalised interoception deficit (Brewer et al., 2016; Murphy et al., 2018). Therefore, the potential relationship between thermal taste, five-factor personality, different measures of interoception and alexithymia is an interesting avenue for future research.

#### *4.4.2.4 Orosensory response*

##### *4.4.2.4.1 Sweet liker status*

The finding that thermal taster status is independent of SLS is consistent with Yang et al. (2020). However, increased liking of two sucrose concentrations (6% and 24%) predicted TT Status. The thresholds at which TT probability increases correspond to the mean liking scores of LSL for these two solutions, suggesting that TT are less likely

to be LSL. Kavaliauskaite et al. (2023) recently reported LSL to rate sweetness intensity higher than HSL. The increased orosensitivity of the group that may be less likely to be TT potentially calls into question previous findings linking thermal taste with increased orosensitivity.

The relationship between elevated liking of specific sucrose concentrations and TT Status may also indirectly result from the relationship between HSL and the largest TT Cluster (Cluster 1). Participants in Cluster 1 were more likely to be HSL than expected by chance, and higher hedonic response to the 35% sucrose solution was predictive of cluster membership. This phenomenon may be linked to differences in specific thermal taste mechanisms: Chapter 3 proposed that participants in Clusters 2 (Sour/Savoury) and 3 (Sweet/Bitter) report traditional thermal taste in TRPM5-dependent mechanisms, and TRPM5 is linked to sweetness perception (Talavera et al., 2005). (A link between Metallic thermal taste and TRPM5 also exists, although this link was less direct). Thermal taste mechanisms may interact with chemosensory taste, resulting in different sweetness experiences in participants who experience TRPM5-dependent thermal taste.

#### 4.4.2.4.2 PROP

This study found no relationship between TT Status and PROP taster status, consistent with previous findings (Bajec & Pickering, 2008; Bering et al., 2013; Thibodeau et al., 2019; Yang et al., 2014; Yang et al., 2020). In addition, no relationships between any TT clusters were found, further underlining the independence of these phenotypes. However, the present study found higher PROP responsiveness to predict TT Status, mirroring the finding of Green and George (2004) that TT are more responsive to PROP, likely due to an interaction between the two phenotypes. As thermal tasters exhibit increased orosensitivity (Bajec & Pickering, 2008; Bajec et al., 2012; Green et

al., 2005; Green & George, 2004; Thibodeau et al., 2019; Yang et al., 2014), it is possible that TT participants who perceive PROP experience the sensation of PROP more intensely partially due to the heightened orosensitivity stemming from TT Status.

#### 4.4.2.4.3 Fatty acid responsiveness

No differences were found in fatty acid thresholds between TT and TnT or between TT subgroups. However, responses to specific concentrations of oleic acid affected predictions in both models. Not distinguishing levels 4 (4 mM/L), 5 (8 mM/L) and 6 (16 mM/L) from the control increased the probability of participants being TnT, suggesting a relationship between TT Status and fatty acid sensitivity. A similar effect was seen for TT subgroups; not distinguishing level 6 from the control increased the probability of being classified in Cluster 2 or TnT. Overall, this suggests an interaction between specific drivers of oleogustus and thermal taste - potentially due to differences in the expression of TRPA1 and TRPM5. TRPA1 (Motter & Ahern, 2012; Startek et al., 2019) and TRPM5 (Liu et al., 2011; Sclafani & Ackroff, 2022) are both involved in fatty acid sensation and were linked to thermal taste sensation in Chapter 3.

## 4.5 Limitations

One limitation of the research reported in this chapter is that it primarily used participant self-report, which can limit the accuracy of underlying construct measurement. For example, Private Body Consciousness measures interoceptive sensibility (the awareness of bodily sensations), which has been shown to be independent of the implicitly measured construct of interoceptive accuracy (Garfinkel et al., 2015). However, Corneille and Gawronski (2024) recently argued that self-reports have an advantage over implicit measures by being better at predicting

behaviour and more reliable than implicit measures. Nevertheless, the inclusion of more implicit measures would have strengthened the possible conclusions.

In addition, this study only investigated orosensory sensitivity for a limited set of stimuli. The use of an expanded set of stimuli, more similar to that used in previous studies would have allowed the research to further clarify the relationships and interactions between orosensory sensitivity, thermal taste and thermal taster subgroups.

The use of the clusters obtained in Chapter 3 represents a further limitation. In Chapter 3, both thermal taste responses from participants were clustered independently, and for a proportion of participants (35%), their clustering assignment was inconsistent. This inconsistent clustering represents a potential source of error in the models, as these participants may not have shared characteristics with their assigned cluster, and may have had characteristics more similar to their secondary cluster. In addition, these participants, as a group, may have shared characteristics, which was not explored. However, the fact that the inconsistently classified participants tended to fit well into one cluster and poorly into another cluster (as opposed to moderately well into both clusters) indicates that this was a prudent choice. However, this chapter could have, alternatively, explored the inconsistently classified participants as their own group – either as a TT Status group or as a fifth TT Cluster.

A final limitation relates to the interpretation of ALE plots. ALE plots are intended to be interpreted visually, and the ability to illustrate the effect of a particular variable across its range is a strength of the approach. However, this introduces the potential for researcher error due to the subjective judgements required. Although the judgement was simple for some variables (for example, Cleaning up after eating in TT

Cluster prediction), variables where there were smaller effects were more difficult to judge, potentially introducing researcher bias.

#### **4.6 Conclusion**

In this chapter, a group of participants were successfully profiled for food behaviour, personality and orosensory responsiveness, and RFMs coupled with interpretable machine learning was successfully used to establish the relative importance of specific measures and variables in predicting TT Status and TT Cluster. Overall, it was found that individual variables were better predictors of TT groups than aggregated measures.

This study found that TT were older than TnT. TT also differed from TnT in salt and heat addition behaviours. TT also tended to be more extraverted and open to new experiences. Interpretable machine learning highlighted differences in PROP sensitivity, food neophobia, oral processing behaviour, aspects of spicy food liking and private body consciousness between TT and TnT. In addition, differences in sweet liking and fatty acid sensitivity were highlighted by the RFM.

This study also found significant differences between TT subgroups. Specifically, TT clusters differed in sweet liking and personality as measured by the BFI. The RFM further highlighted differences in food involvement (specifically Set & Disposal), food neophobia and fatty acid sensitivity.

These findings highlight the complex relationships between thermal taste, personality, food behaviour and orosensitivity. How these different factors interact with thermal taste in influencing food choice is explored further in Chapter 5.

## **Chapter 5 *The path to enlightenment*: The relative impact of factors mediating food choice behaviour, including Thermal Taste**

### **5.1 Introduction**

Prior research on thermal taste has yielded inconclusive findings concerning its overall impact on food choice behaviour (Pickering et al., 2010; Pickering et al., 2009). Nevertheless, there are indications that TT and TnT differ in their preference for carbonated beverages (Hort et al., 2016), creamy items (Pickering & Klodnicki, 2016), and food textures (Bajec & Pickering, 2010; Pickering et al., 2016) and that there is an interaction between TT Status, beverage temperature and liking (Yang, 2015). More discriminant methods, such as emotional response, have exhibited preference differences between TT and TnT, where traditional methods did not show differences (Q. Yang et al., 2018).

There is growing evidence that TT differ from TnT in overall food behaviours. Yang (2015) found small differences in Food Neophobia between TT and TnT, mirrored in research in this thesis presented in Chapter 4. In addition, TT are more likely to be “foodies” (Pickering & Pickering, 2022), and although different measures were applied, Chapter 4 revealed differences in specific aspects of food involvement between TT and TnT and between subgroups of TT. Recently, differences in disgust sensitivity between TT and TnT have also been evidenced (Thibodeau et al., 2023). These observations make it interesting to explore not only *how* thermal taster status influences food choice directly but also how thermal taster status may *interact* with other mediating factors affecting food choice, such as orosensory sensitivity, personality, and other food-related behaviours such as oral processing behaviour.

Notably, Chapter 3 of this thesis identified four TT groups (three TT Clusters and TnT), and Chapter 4 revealed differences between TT and TnT and between the TT groups

for several food-related behaviours, including sweet liking, salt addition, and spice addition. Consequently, exploring not only *how* thermal taster status is related to food choice but whether the different characteristics demonstrated by the TT groups, namely, tastes reported and time/temp at which it is reported, also impacts food choice behaviour is warranted.

### **5.1.1 Potential mechanisms of thermal taste's influence on food choice**

One possible mechanism of thermal taste's influence on food choice is through the induction of (thermal) taste via the thermal properties of food, which either synergistically intensifies food sensory properties or induces new sensory properties, potentially suppressing the original properties. Yang (2015) demonstrated this phenomenon by inducing thermal taste through different water temperatures. In addition, Nachtigal and Green (2020) used water instead of a thermode to induce thermal taste. However, the impact of food temperatures should be interpreted with caution, as consuming a hot beverage has been shown to increase oral temperature by a maximum of 2°C, and consuming a cold beverage to decrease oral temperature by 2.2°C (Newman & Martin, 2001). However, Newman and Martin (2001)'s study afforded several methodological limitations, including minimum temperature measurement capability and not controlling (or reporting) the temperatures of the beverages. Therefore, before directly interpreting the actual effect of thermal taste temperatures on induced taste in a consumption setting, it is necessary to thoroughly investigate the effects of food and beverage temperature on oral temperature.

A second possible mechanism of influence is through an indirect relationship with receptor sensitivity. It is possible that reporting thermal taste at a specific temperature indicates a generalised enhanced sensitivity of receptors activated at that temperature, resulting in a higher responsiveness to specific textures and tastes. This higher

responsiveness is related to the overall heightened orosensitivity previously linked to thermal taste (Bajec & Pickering, 2008; Bajec et al., 2012; Green et al., 2005; Green & George, 2004; Thibodeau et al., 2019; Yang et al., 2014) and could, in turn, result in different food choices. Note, however, that this is probably not a direct sensitivity to a specific taste, as reporting a specific thermal taste is not directly linked to increased sensitivity to that specific taste (Thibodeau et al., 2020). Instead, this is a heightened sensitivity of temperature-sensitive receptors at thermal taste induction temperatures. These receptors include temperature-sensitive receptors specifically linked to thermal taste (TRPM5) (Talavera et al., 2007), temperature-sensitive receptors that are both taste receptors and mechanoreceptors like ENaC (Drummond, 2007; Talavera et al., 2008), TRPV1 (Birder et al., 2002; Talavera et al., 2007), and TRPA1 (Castiglioni & García-Añoveros, 2007), temperature-sensitive mechanoreceptors like TRPV4 (Suzuki et al., 2003) and temperature-sensitive receptors not yet directly linked to taste or mechanoreception like TRPC5 (Zimmermann et al., 2011) and TRPM8 (Voets et al., 2007).

### **5.1.2 Influence on food choice of factors potentially related to thermal taste**

In isolation, the demographic, food behavioural, personality and orosensory factors detailed in Chapter 1 likely impact food choice. In addition, due to their potential relationship to thermal taste (as explored in Chapter 4), these factors likely interact with thermal taste regarding food choice behaviour. Previous studies investigating food choice typically looked at these (demographic, food behavioural, personality and orosensory) factors in isolation or in combination with a limited number of other factors. Therefore, it is necessary to explore the impact of these factors on food choice in conjunction with thermal taste. The following section reviews the known effects of these factors on food choice.

#### *5.1.2.1 Demographics (Age, gender/sex, and ethnicity)*

Demographic factors are expected to influence food choice through socio-cultural and physiological differences. Through different social and behavioural backgrounds, different ethnic groups eat different foods and have different oral processing behaviours (Ketel et al., 2020); there is also evidence that through genetic mediation, ethnicity can impact taste sensitivity (Yang et al., 2020). Age is related to differences in taste sensitivity (Jilani et al., 2022; Mojet et al., 2001; Mojet et al., 2003; Sato et al., 2022; Yang et al., 2022), oral processing behaviour (Ketel et al., 2021), and the ability to recognize tastes (Doty et al., 2017). There is some evidence of gender/sex differences in taste sensitivity (Jilani et al., 2022) and that women report olfactory sensations as more intense than men (Chao et al., 2022). Gender also related to attitudes and engagement with foods (Rodrigues et al., 2020). For example, there are gender differences in fatty meat consumption (Spinelli et al., 2020) and spicy food liking, although the latter is mediated by personality (Byrnes & Hayes, 2015; Spinelli et al., 2018).

#### *5.1.2.2 Food Behaviour*

##### *5.1.2.2.1 Oral processing behaviour*

Although the study in Chapter 4 found no differences in oral processing behaviour between TT and TnT, it is still an essential potential mediating factor in overall food choice. Texture plays a crucial role in food aversions (Scott & Downey, 2007), and texture preference is related to food neophobia overall (Laureati et al., 2020). However, oral processing subgroups do not translate directly to texture preference (Kim & Vickers, 2020). Although different oral processing behaviour groups will choose different products (Jeltema et al., 2016), textures that fit the preferred oral processing behaviour (foods that can be manipulated to fit the preferred behaviour)

drive this choice, and not the texture itself (Jeltema et al., 2015; Jeltema et al., 2020; Jeltema et al., 2014). Although single-point texture ratings are unaffected by oral processing behaviour (Zhou et al., 2021), overall texture experiences differ based on oral processing behaviours (Jeltema et al., 2016; Jeltema et al., 2020). This change in experience extends to other sensory properties, demonstrated during the melting of ice cream (Doyennette et al., 2019), chewing gum (Kim et al., 2019) and chocolate consumption (Li & James, 2021). However, although oral processing can affect satiation (James, 2018), different oral processing subgroups do not differ in their portion size selection (Nguyen et al., 2020).

#### 5.1.2.2.2 Food involvement

The food involvement scale (FIS) (Bell & Marshall, 2003) measures engagement with foods and is directly linked to home cooking, as participants more involved with food are more likely to cook for themselves (Davison et al., 2015). In addition, people with higher food involvement tend to eat healthier diets overall (Schnettler et al., 2019), consuming fewer calories (Marshall & Bell, 2004), with fewer calories from fats (Marshall & Bell, 2004) and more vegetables overall, which could be partly a result of a lower consumption of junk food (Davison et al., 2015) or restricting the consumption of less healthy foods (Jezewska-Zychowicz et al., 2020). Interestingly, participants with higher food involvement tend to consume more spices and milk (Marshall & Bell, 2004). In a study investigating hedonic responses to different cheeses, participants with higher food involvement tended to like cheese more overall (Lahne et al., 2014). Similarly, the enjoyment of coffee of participants with higher coffee involvement was less affected by situational factors (Kim et al., 2016) – although this effect has not yet been shown to extend to all food and beverage categories.

#### 5.1.2.2.3 Food neophobia

Food neophobia (Pliner & Hobden, 1992), correlated to Big Five Inventory (BFI) personality characteristics (Knaapila et al., 2011; Nezlek & Forestell, 2019), is related to differences in overall food choice (Jaeger, Rasmussen, et al., 2017). Food neophobics (people who score low on the Food Neophobia Scale) tend to score the hedonic attributes of foods lower (Fenko et al., 2015; Henriques et al., 2009; Laureati et al., 2018; Prescott et al., 2022), an effect which is more pronounced with foods with dominant “warning” orosensations (Laureati et al., 2018). Specifically, food neophobics dislike spicy foods (Spinelli et al., 2018) and choose less bitter/astringent foods (De Toffoli et al., 2019) in food choice experiments. Neophobics dislike fruit and vegetables (Tornwall et al., 2014), likely resulting in lower consumption (Hazley et al., 2022). Food neophobics also self-report a higher effect of food appearance, odour, texture, temperature, oral processing, and aftertaste on consumption (Prescott et al., 2022) and food neophobia interacts with texture preference (Laureati et al., 2020). While food neophilics prefer more complex foods, neophobic consumers prefer bland foods (Olabi et al., 2015) and tend to get more of their calories from sugar (Hazley et al., 2022). Therefore, food neophobia can be thought of as an interaction of personality (Knaapila et al., 2011; Nezlek & Forestell, 2019) and sensitivity to food characteristics (Prescott et al., 2022), influencing hedonic experiences (Prescott et al., 2022) and food choice motives (Eertmans et al., 2005), likely resulting in overall differences in diet (Jaeger, Rasmussen, et al., 2017), which tend to be less healthy overall (Hazley et al., 2022).

#### 5.1.2.2.4 Food modification

Modifying the sensory properties of foods through condiment addition showed a stronger relationship with orosensitivity than the recalled pleasantness of foods

(Puputti, Hoppu, et al., 2019). In addition, Chapter 4 showed differences between TT subgroups in their food modification behaviour. It is, therefore, necessary to contextualise this behaviour amongst all the other potential factors of food behaviour.

#### 5.1.2.2.5 Spicy food liking

Although several studies have evaluated the effect of spicy food liking on sensory sensitivity (Li et al., 2017; Ludy & Mattes, 2012; Su et al., 2022) and the interrelationship between spicy food liking and personality (Byrnes & Hayes, 2013, 2016; Scott et al., 2019), few studies have examined the effects of spicy food liking on overall food choice. Nevertheless, chilli users tend to consume more calories than chilli nonusers (Choi & Chan, 2015), potentially explaining the higher overall BMI of chilli users (K. Yang et al., 2018).

#### 5.1.2.3 Personality (BFI/PBC)

Behavioural factors can also influence food choice, some directly linked to food attitudes and others indirectly linked to foods. The five-factor personality model has been linked to sensory sensitivity (Ferentzi et al., 2017), alexithymia (Barańczuk, 2019), disordered eating (MacLaren & Best, 2009), and acceptance of novel foods (Ardebili & Rickertsen, 2020). Differences in emotional responses to wines also correlate to aspects of the Big Five Inventory (Mora et al., 2019).

Specific, somewhat consistent eating patterns are linked to aspects of the BFI. Openness to experience correlates positively with the consumption of plant-based foods (Keller & Siegrist, 2015; Lunn et al., 2014; Pfeiler & Egloff, 2020), and conscientiousness has been linked to healthy eating overall (Keller & Siegrist, 2015; Lunn et al., 2014), which includes the consumption of plant-based foods (Pfeiler & Egloff, 2020). Neuroticism associates negatively with plant-based food consumption

(Pfeiler & Egloff, 2020) and agreeableness has been linked to lower meat consumption (Keller & Siegrist, 2015). However, studies that have tried to link the BFI to the consumption of specific foods have failed to find patterns (Nystrand et al., 2021).

The awareness of bodily sensations, measured simplistically through the Private body consciousness (PBC) questionnaire (Miller et al., 1981), is expected to affect food selection, despite some inconclusive studies on the effect of PBC on food choice (Byrnes & Hayes, 2013). However, consumers who tended to eat vegetables with higher bitterness, sourness and astringency scored higher in PBC (Pierguidi et al., 2023). In addition, differences in PBC showed differential effects on the influence of different types of information on food choice (Jaeger et al., 1998; Solheim & Lawless, 1996). Participants with higher PBC also tend to rate emotional words regarding food higher (Jaeger & Hedderley, 2013), and PBC has been linked to higher emotional arousal (Jaeger et al., 2022).

#### *5.1.2.4 Taste phenotyping (orosensitivity)*

##### 5.1.2.4.1 PROP taster status

Since its discovery, it has been postulated that PROP taster status likely affects food choice. PROP taster status has been found to be related to orosensitivity overall. Generally, the evidence around PROP taster status is inconclusive, with many studies having null findings (see Hayes & Keast, 2011; Reed, 2008). However, as it is a commonly explored factor, it makes sense to explore its relative effect in the context of other potential factors.

##### 5.1.2.4.2 Sweet liking

Differences exist in hedonic responses to sweet stimuli, with people typically separated into low sweet likers/sweet dislikers, medium sweet likers and high sweet likers

(Pangborn & Giovanni, 1984). The liking of sucrose solutions generalises to other sweet stimuli (Looy et al., 1992) but does not impact the hedonic responses to other nonsweet stimuli (Looy et al., 1992). Generally, higher hedonic responses to sweet stimuli predict sugar intake (Holt et al., 2000; Lim et al., 2020; Pangborn & Giovanni, 1984), including sugar-sweetened beverages (Peng et al., 2023). Although sweet likers tend to rate the liking of sweet foods higher, this does not always translate into higher consumption of these specific foods (Armitage et al., 2023). In contrast, medium sweet likers consume more phenol-rich vegetables and tend to consume less added sugar (Spinelli et al., 2021). Sweet likers have been found to have a higher proportion of body fat than sweet dislikers (Iatridi et al., 2020), although sweet liking was not directly linked to obesity (Armitage et al., 2021). Interestingly, sweet likers performed better at interoception tasks and were more mindful and intuitive eaters (Iatridi et al., 2021), possibly accounting for some counterintuitive relationships between liking, consumption and body composition.

#### 5.1.2.4.3 Fat taste sensitivity

Fatty acids are perceived as both a taste and a texture (Laugerette et al., 2007), and sensitivity towards them is genetically correlated (Graham et al., 2021; Laugerette et al., 2005), although there is evidence that repeated exposure (Tucker & Mattes, 2013) and the consumption of a high-fat diet (Costanzo et al., 2018; Keast, 2016; Tanaka et al., 2022) also affects sensitivity. Conversely, fat taste sensitivity can affect diet (Karmous et al., 2022) and BMI (Daoudi et al., 2015), although a meta-study has shown that body composition is related to by the enjoyment of fatty acids rather than sensitivity (Cox et al., 2016), explaining null results between fatty acid sensitivity and diet (Zhou et al., 2016).

### **5.1.3 Specific considerations around thermal taste**

#### *5.1.3.1 Previous research*

Previous research into food choice and thermal taste has mainly highlighted differences in texture liking (Bajec & Pickering, 2010; Pickering et al., 2016), and differences in liking driven by trigeminal (Hort et al., 2016) and temperature (Yang, 2015) differences. However, Chapter 4 highlighted differences in food-related behaviours and other factors between TT groups (TT and TnT and between TT subgroups) that may interact with thermal taste to influence food choice.

In particular, Chapter 4 highlighted differences in oral processing behaviour, salt and spice addition behaviour and enjoyment of the burn of spices between TT and TnT. TT also tended to be lower in food neophobia and scored higher in the Openness and Extraversion dimensions of the BFI. Chapter 4 also suggests differences in fatty acid sensitivity between TT and TnT. In addition, TnT differed from the largest TT subgroup in their responses to the items on the Set & Disposal subscale of the FIS. These findings highlight the necessity to investigate relationships and interactions between thermal taste and other mediating factors.

#### **5.1.4 Modelling background**

Path modelling is an approach in which the interrelationships between different data blocks are tested to predict an outcome variable or set of variables. For example, this approach has been used to link the physical properties, sensory properties and finally, hedonic judgements of foods (Pages & Tenenhaus, 2001). It has also been used to understand the effect of personality on food choice (Keller & Siegrist, 2015), the relative effect of product and consumer characteristics on satiety, liking and portion sizes (Nguyen et al., 2020), and the relative effect of different aspects of the Food

Choice Questionnaire on interest in low sugar products (da Veiga et al., 2021).

However, path modelling presents two specific challenges: the requirements for unidimensionality and model specification (Nguyen et al., 2020). Unidimensionality means that each piece of data included in the model needs to reduce to a single principal component or factor. If all data blocks (sets of related variables like personality) do not reduce to a single factor, a possible approach is first to perform data preprocessing/factor analysis and to include singular factors in the model individually (Nguyen et al., 2020). In addition, the researcher needs to specify the model (order of components) based on a theoretical framework or a hypothesis. It is also possible to test several potential paths to discover interrelatedness between factors (Nguyen et al., 2020). However, as the potential number of components increases, the potential number of paths increases exponentially.

An alternative approach is using a supervised modelling approach specifically intended for multiblock methods, which automates the order of data block choice and can handle blocks with different underlying dimensions. Smilde et al. (2022) details several partial least squares (PLS) approaches which can handle these differences in underlying dimensionality that reduce predictors to a smaller set of uncorrelated components but recommends Response Oriented Sequential Alternation (ROSA) (Liland et al., 2016) for automated block selection. ROSA is a type of sequential PLS, where (data) blocks are included in the predictive model in a stepwise manner, chosen algorithmically based on the extracted components from the block which will most increase the overall prediction (decrease the residuals) of the model (Smilde et al., 2022).

Another benefit of the ROSA approach is the potential to allow the algorithm to choose

from distinct components (components extracted from a single data block) or distinct and common components (Smilde et al., 2022). When the ROSA algorithm considers common components, components derived simultaneously from a combination of two or more data blocks can be included, allowing for the consideration of common underlying features from different blocks. The following sections will clarify this aspect of ROSA models in more detail.

#### *5.1.4.1 Technical details of the ROSA algorithm*

Three key concepts are necessary for understanding ROSA models and their underlying algorithm. First, it is necessary to differentiate between supervised and unsupervised component extraction. Second, steps in the ROSA algorithm and how the algorithm chooses blocks/components need clarification. Third, the difference between common and distinct components needs to be clarified.

##### *5.1.4.1.1 Supervised vs unsupervised components*

In unsupervised (exploratory) methods like PCA and its multiblock extensions (for example, MFA), the goal is to simplify a complex data space (X) either for further interpretation or modelling (feature extraction). In contrast, in supervised (predictive) methods like PLS and its multiblock extensions (e.g. ROSA), the goal is to predict a variable block (Y) from a complex data space (X). Therefore, unsupervised methods only have an X (explanatory) block (or blocks), while supervised methods require both X and Y (target) blocks (Smilde et al., 2022).

The difference in the relationship between X and Y between unsupervised and supervised methods also informs their component extraction approaches. As unsupervised learning (e.g., PCA and MFA) has no Y space, the algorithm extracts components sequentially to explain the maximum variation in the X space. In contrast,

in supervised learning (e.g. PLS and ROSA), the algorithm extracts components based on their ability to predict the Y space rather than simply their ability to simplify the X space (Smilde et al., 2022).

#### 5.1.4.1.2 ROSA component fitting

As mentioned, the ROSA algorithm extracts PLS components sequentially while choosing blocks/components automatically based on their predictive ability (Liland et al., 2016). In contrast to single-block PLS, the ROSA algorithm fits and evaluates several candidate components before adding components to the model (Table 5-1). The ROSA algorithm extracts the component from each (X) data block in each component fitting step or iteration, predicting the maximum variation in the Y space. The algorithm then compares these candidate components and selects the one which predicts the maximum variation in the Y space (Table 5-1) (Smilde et al., 2022). Although the algorithm selects components, this is referred to as block selection, in keeping with the practice of “block specification” in similar methods, and likely because of the transience of candidate components. In the next iteration, new candidate components are extracted from all (X) data blocks to fit the remaining variation in the Y space. This process continues with as many iterations as components specified by the researcher, as there is no “natural stopping” of the algorithm (Smilde et al., 2022).

*Table 5-1 Steps in each iteration of the ROSA algorithm, focussed on component extraction and component/block selection*

<b>Step</b>	<b>Process</b>	<b>Result</b>
1. Candidate component fitting	Algorithm extracts exponents from all data blocks in the X space (distinct component model) or blocks and combinations (common component model) based on their ability to predict Y	Candidate Components for each block and/or block combination
2. Block choice	Block (s) chosen based on the best fitting component in Step 1	Best fitting component
3. Model update	Selected component incorporated in model	Reduction in the unexplained variance considered in the Y space

#### 5.1.4.1.3 Common vs distinct components

The idea of common and distinct components can be better understood using an unsupervised example and PCA. A typical situation encountered in sensory science is that two or more datasets (or blocks), such as sensory (block A) and chemical (block B) data, are available to explain a set of products (Smilde et al., 2017). Although it is possible to choose a block (e.g. A, sensory data) as a Y space, this approach is only suitable if the end goal is predictive modelling instead of exploration. An alternative approach is to perform PCA on all the data blocks separately (PCA-A and PCA-B) and explore relationships between the blocks – either algorithmically (i.e. MFA) or manually. However, a third approach is often used in industry (as regularly encountered by the researcher in their tenure in industry) in that the blocks are combined or concatenated prior to PCA into a new space (block AB). The difference between this resulting model (PCA-AB) and those extracted from the separate blocks (PCA-A and PCA-B) is that PCA-AB includes components which explain variation common to both of the original blocks (Smilde et al., 2017). These components may not be related to the original components extracted by PCA-A and PCA-B, but rather explain the space common to the new concatenated (A+B) dataset.

In multiblock PLS methods, distinct components relate to one explanatory block. In other words, the algorithm extracts these components to explain the Y while only considering one block at a time. The algorithm also considers components extracted from concatenated blocks when fitting common components, explaining the common space between two or more blocks. Common components do not represent interactions in the strict mathematical sense but represent a relationship between blocks in modelling the outcome (Smilde et al., 2022). This relationship can be better understood as a latent predictive element in the combined space of two blocks, which

does not exist when the blocks are not combined. If this predictive element existed without the combination of blocks, it would have, by definition, been a distinct component.

#### *5.1.4.2 Interpretation of ROSA models*

Although the application of ROSA to predict multiple outcome variables has been recommended by Smilde et al. (2022) and the `multiblock` R package (Liland, 2022) includes relevant functionality, there have been no published examples of multiple-outcome ROSA models to use as examples for interpretation. However, Liland et al. (2016) notes that similar interpretation approaches to regular PLS models can be used. However, ROSA, specifically multiple response ROSA, presents interpretation challenges and opportunities.

The automatic block selection of the ROSA algorithm based on prediction ability presents an opportunity for interpretation based on block selection order (Smilde et al., 2022). As blocks are selected sequentially based on the ability of their extracted components to predict variability in the Y space, blocks selected earlier can be considered more important for prediction. For example, this interpretation has been applied to compare different spectral data preprocessing methods (Dirks et al., 2023; Karami et al., 2024). In addition, as the ROSA algorithm can select the same block multiple times (Liland et al., 2016), selection frequency can be used as a secondary measure of block importance.

Two of the properties of ROSA, which make it uniquely suited for complex analysis problems, also complicate the use of some typical PLS interpretation approaches. The fact that ROSA can handle multiple X blocks with different dimensionalities from which different numbers of components can be extracted can complicate the typical

approach of using scores and loading plots for a complex model (Smilde et al., 2022). However, some of the variable selection procedures explored in Mehmood et al. (2012) and Mehmood et al. (2020) can be used to simplify interpreting a complex ROSA model.

The first approach is to use variable influence on projection (VIP) (Wold et al., 1993) to establish which variables most significantly influence the overall projection of a ROSA model. Although more sophisticated methods have been proposed, a general guideline is to use a VIP cutoff value of 1 (Mehmood et al., 2012), an approach previously used with ROSA (Strani et al., 2022). Secondly, interpretation of regression coefficients can be used (Mehmood et al., 2012), particularly useful for the multiple response case of ROSA in the presence of a regression coefficient for each input (X) and output (Y) variable. Generally, larger regression coefficients indicate more important variables, and the sign indicates the direction (positive or negative) of influence (Azadi, 2021), an approach used in the interpretation of ROSA models (Azadi, 2021; Strani et al., 2022).

#### **5.1.5 Aims**

The research in this chapter aimed to explore the relative importance of thermal taste in the context of food choice behaviour. This work was primarily exploratory, meaning no specific hypotheses were tested, and all data blocks were included based on their potential relationship to food choice. However, it was expected that the nuances of thermal taste (temperature at which thermal taste is reported, thermal taste cluster or the identity of thermal taste) would be more influential than thermal taste itself, as different thermal taste experiences (temperatures and tastes experienced) likely interact with orosensory experiences during eating in a complex manner.

Specifically, the chapter set out to:

- 1) Use ROSA models to explore the relative importance of thermal taste, demographic, food behaviour, personality and orosensory responsiveness data in predicting food choice behaviour.
- 2) Explore the relative importance of different aspects of thermal taste in predicting food choice using component selection order of a distinct component ROSA model (no common predictive relationships considered) as well as interactions between thermal taste and relevant aspects using the selection order of the common component ROSA model (common predictive relationships considered).
- 3) Explore predictive relationships between thermal taste and other data types in a distinct component model by interrogating the relative loadings to components extracted from the model.
- 4) Determine the relative importance of different aspects of thermal taste (TT cluster, nature of thermal taste reported, temperature at which thermal taste was reported) in predicting food choice behaviours compared to other factors measured.
- 5) Evaluate the specific impact of the most important aspect of thermal taste identified in 4) on the consumption prediction of individual food items.

## **5.2 Materials and methods**

This chapter explored thermal taste's relative influence on food choice using the consumption frequency of common food items and an exploratory approach in modelling the relative influence of different aspects of food behaviour and non-food behaviour. Specifically, data collected during Chapter 3 (thermal taste response and

subgroup) and Chapter 4 (Demographics, Food behaviour, Personality and Orosensory responsiveness) were used. The ROSA algorithm, resulting in two separate models, was employed to explore the relative influence of these factors on food choice behaviour.

### **5.2.1 Data Collection**

#### *5.2.1.1 Participants*

The participant cohort described in Chapter 4 also provided the data for this study.

#### *5.2.1.2 Explanatory data*

Data collected in Chapters 3 and 4, structured as fourteen data blocks (Table 5-2), were used as explanatory data to predict eating behaviour collected using a Food Frequency Questionnaire (FFQ). From Chapter 3, three categorical data blocks were used: TT cluster membership, the identity of reported thermal taste and the time (temperature) at which thermal taste was reported. All RapCoTT data was combined into a singular on/off (taste/no taste) response per 5s/5°C per participant. A similar concatenation approach to the one described in Chapter 3 was used. For participants not reporting thermal taste, their cluster membership was “TnT”, and taste identity and reporting time were indicated as zeroes.

Raw data from each measure in Chapter 4 were included as a separate block. For example, all the data from the BFI constituted one block. A specific difference existed in discriminating the data for ROSA modelling. As PROP responsiveness data existed in a single dimension (one variable indicating response), extracting a principal component from this “block” was impossible. Therefore, all the sensory response data (PROP, sweetness intensity, fatty acid response) was combined into a single block (“Orosensory Response”). However, sucrose liking responses and SLS cluster

membership were contained in a separate block (“Sweet liking”) (Table 5-2).

*Table 5-2 Data blocks for input into food choice prediction model. Size refers to the dimensions of the data block, which refers to either the number of recoded categories or the number of variables in the data block. The category (Cat) refers to the data block categorisation in this chapter; Thermal Taste (T), Unrelated to thermal taste or food behaviour (N), Direct measures of food behaviour (D) or Indirect measures of food behaviour (I).*

Subgroup	Data block	Cat	Size	Notes
Thermal taste	TT Cluster	T	3	Binary data for the three identified clusters
	TT Taste	T	6	Binary data for the six reported thermal tastes
	TT Temperature	T	13	Binary data describing whether each participant reported thermal taste in each of the 13 temperature blocks
Demographics	Demographics <sup>1</sup>	N	2	Age (numeric) and gender (Binary)
	Ethnicity <sup>1</sup>	N	6	Binary data for six most prominent Ethnicity groups
Food Behaviour	Food Involvement <sup>2</sup>	I	12	Numeric values for each of the food involvement items
	Food modification <sup>3</sup>	D	15	Average food modification score for each behaviour
	Oral processing behaviour <sup>4</sup>	D	4	Oral processing subgroup membership (Binary)
	Food Neophobia <sup>5</sup>	I	10	Raw Food neophobia responses
	Spicy food liking <sup>6</sup>	D	6	Raw response data from the Spicy Food Liking questionnaire
	Sweet liking	D	9	Optimum sucrose liking, liking and intensity of 5 sucrose levels, encoded SLS cluster membership
Personality	Big Five Inventory <sup>7</sup>	N	44	Rating of each of the items on the Big Five Inventory
	Private Body Consciousness <sup>8</sup>	N	5	Rating of each of the items on the Private Body Consciousness questionnaire
	Sensory response	N	15	Intensity of 5 sucrose levels, raw response data on eight levels of oleic acid, and PROP intensity data

<sup>1</sup>Stats NZ (2019); <sup>2</sup>Bell and Marshall (2003); <sup>3</sup>Puputti, Hoppu, et al. (2019); <sup>4</sup>Jeltema et al. (2015); <sup>5</sup>Pliner and Hobden (1992), <sup>6</sup>Lawless et al. (1985), <sup>7</sup>John et al. (2008); John and Srivastava (1999) <sup>8</sup>Miller et al. (1981)

For further interpretation, data blocks were subdivided based on their relationships to thermal taste and food behaviour, resulting in four categories (Table 5-2). The first category, “Thermal Taste” (T) related to whether the data block measured thermal taste. The second category, “Unrelated” (N), contained all the data blocks that contained neither thermal taste nor food behavioural information. The remaining data blocks were all related to food behaviour and were divided into “Direct” (D) and “Indirect” (I) measures. Direct measures were all data blocks containing questions or information about the preference of foods or where participants directly reported on their food behaviour. Indirect measurements were the data blocks which only contained questions on attitudes and general food-related behaviours.

### 5.2.1.3 Prediction data (Food frequency)

A modified version of the NZ short food frequency questionnaire (FFQ) described in

Sam et al. (2020) was used to collect food consumption data. The questionnaire consisted of 57 food items commonly consumed in New Zealand, shortened to 51 items (Table 5-3 and Table 5-4) (P. Skidmore, personal communication, 10 June 2020).

*Table 5-3 Dairy, Fruit, Vegetable and Protein items included in the Food Frequency Questionnaire, including their serving sizes and examples as presented to participants*

<b>Category</b>	<b>Item</b>	<b>Serving</b>	<b>Example</b>
Dairy	Milk	1 cup/ 250mL	Include milk in hot drinks, cereals, creamed soups etc. e.g. Cow, Soy
	Ice cream	1 cup	NA
	Yogurt	1 pottle (150ml)	NA
	Low fat Cheese	¼ cup	Cottage, Ricotta, Low Fat Cheddar
	Cheese	¼ cup	Cheddar, Edam, Tasty, Mozzarella, Brie, Camembert
	Cream based dairy	¼ cup	Cream, Sour cream, Cream cheese
Fruit	Banana	1	NA
	Apples/Pears	1	NA
	Citrus fruit	1	Oranges, Mandarins, Grapefruit, Lemons
	Stone fruit	1	Apricots, Plums, Nectarines, Peaches
	Berries	½ cup	Strawberries, Blueberries
	Dried fruit	1 Tbsp	Raisins, Sultanas, Prunes
	Other fruit	1 serving	Kiwifruit, Grapes, Feijoa, Pineapple, Mango, Rhubarb, Tamarillos, Guava, Pawpaw, Melon
Vegetable	Tomatoes	1 small/ ½ cup	Fresh, canned, tomato based sauce
	Legumes	½ cup	Green beans, Runner beans, Baked beans, Lentils, Chickpeas
	Salad Greens	½ cup	Lettuce, Cucumber, Celery, Rocket
	Other Greens	½ cup	Broccoli, Cauliflower, Spinach, Silverbeet, Cabbage, Brussel Sprouts, Bok choy, Chinese cabbage, Watercress, Puha
	Onion, Leeks	½ cup	Including in cooking
	Potatoes, Kumara, Pumpkin	½ cup	NA
	Other Root Vegetables	½ cup	Carrot, Beetroot, Parsnip, Turnips, Yams
	Other Vegetables	½ cup	Corn, Mushrooms, Asparagus, Courgette, Eggplant, Capsicum, Peas, Coleslaw
Protein	Eggs	1 egg	NA
	Sausages, Hotdogs	1 sausage/ hotdog	NA
	Processed meat	½ cup	NA
	Red meat	½ cup	Mince, Roast, Steak, Stew, Casserole, Lasagne, Frozen dinners, etc
	Meat pie	1 serving	NA
	Poultry	½ cup	NA
	Fatty Fish	½ cup	NA
	Fried fish	1 serving	NA
	Other Fish and Seafood	½ cup	Cod, Sole, Hoki, Gurnard, Shrimp, Mussels, Oysters, Crayfish

*Table 5-4 Cereals, Fast Food and Sweets, Baked Goods and Miscellaneous items included in the Food Frequency Questionnaire, including their serving sizes and examples as presented to participants*

<b>Category</b>	<b>Item</b>	<b>Serving</b>	<b>Example</b>
Cereals	High Fibre Cereals	¾ cup	Porridge, Muesli, Bran Flakes, All Bran
	Other Cold Breakfast Cereal	¾ cup	Light' n' Tasty, Special K, Weetbix
	White Bread	2 slices/ 1 roll	Sliced, Tortillas, Pita etc.
	Wholemeal bread	2 slices/ 1 roll	Sliced, Tortillas, Pita etc
	Crackers, Crispbread	2 crackers	Vitawheat, Cruskits
	Pancakes	1 serving	NA
	Brown Rice/ Wholemeal Pasta	1 cup	NA
	White rice/ Couscous	1 cup	NA
	Other Pasta	1 cup	Spaghetti, Spirals, Instant noodles, Tinned
	Fast Food	Hot chips/ French fries	1 cup
Chips shelf stable		1 cup	NA
Pizza/ Hamburgers		2 slices/ 1 medium burger	NA
International Takeaway		1 serving	Chinese, Thai, Turkish, Indian, Japanese etc
Sweets Baked Goods Misc	Chocolate/ Chocolate bars	5 squares (25g)	NA
	Sweets, Lollies	Small handful	NA
	Biscuits	2	NA
	Cake/Pastries	1 serving	NA
	Jam/Syrup	1 Tbsp	NA
	Nuts	1/8 cup	include Peanut butter), seeds
	Oils	1 Tbsp	Vegetable oil, Olive oil, Mayonnaise, Salad dressing, include frying
	Fats	1 tsp	Butter or Margarine, used as spreads or in cooking, excluding baking

Beverage items were excluded to shorten the overall questionnaire to ensure completion, as the online questionnaire included the questionnaires described in Chapter 4. For each item, participants indicated how often they consumed the item, using eight frequency options ranging from “never or less than once per month” to “4–6 times a day” (“never or less than once per month”, “1-3 per month”, “Once per week”, “2-4 per week”, “5-6 per week”, “1 every day”, “2-3 times a day”, “4–6 times a day”) (Sam, 2012; Sam et al., 2020).

### 5.2.2 Data analysis

All data analysis was performed using R version 4.2. As in Chapter 4, the `targets` framework (Landau, 2021) and `tidyverse` (Wickham, 2017) and `tidymodels` (Kuhn & Wickham, 2020) meta packages facilitated modelling. The package `multiblock` facilitated explanatory modelling (ROSA) (Liland, 2022), and the `plsVarSel` (Mehmood et al., 2012) package was used to establish overall variable importance.

#### 5.2.2.1 Food frequency data processing

Food frequency data was converted to weekly consumption frequency, using the conversion strategy outlined in Sam (2012). Typically, when using FFQ data for nutritional analysis, values are adjusted for total calorie consumption. However, as no nutritional calculations were done, a double normalisation strategy was employed. Initially, to compensate for the variation in calories consumed between participants, the total number of servings consumed per week was used as a proxy for total weekly calorie consumption, and the number of servings per item was expressed as a proportion of the total number of weekly servings. Next, to compensate for typical variation in diet, the data was normalised per product, resulting in consumption data of all products having unit standard deviation.

#### 5.2.2.2 ROSA predictive models (model selection)

Two ROSA (Liland et al., 2016) models were fitted to predict the food frequency block (described in section 5.2.2.1) through stepwise extraction of predictive components from the other data blocks (described in section 5.2.1.2). When fitting these models, the ROSA algorithm sequentially extracted predictive components from all available blocks, basing their selection on the block (or combination of blocks) that most

increased the predicted variation in overall food choice. The two models differed in their component extraction strategies; the first (distinct component) considered components from one data block at a time, whereas the second (common component) considered components relating to a single data block (distinct) or the space encompassed by two different data blocks (common).

The two models served slightly different purposes, which informed choices made during modelling. The distinct component model was intended for interpretation based on its component selection order, X-loadings and regression coefficients. The common component model was intended to explore where thermal taste might interact with the other aspects of interest (as encoded in the other data blocks) in determining food choice. The more conservative interpretation of the common component model was partially driven by the fact that, although Smilde et al. (2022) provides theoretical justification for using common components with ROSA, no published examples exist.

For both models, a two-step leave-one-out cross-validation strategy (Mevik & Cederkvist, 2004) was employed, independently cross-validating the block selection order and the predictions. As the research aimed to explore relationships between predictive variables rather than build a definitive predictive model, the number of components was chosen using an “optimistic” approach. To ensure the inclusion of the optimum components in the model, 50 components were fitted per model, a choice made by the researcher based on their experience testing the ROSA algorithm. (Note that as the model is sequential, fitting an arbitrarily large number of components does not affect the final model based on preceding components.)

The final number of components in both models was kept the same and was chosen

based on the point at which the unadjusted  $R^2$  (explained variance of the training data) decreased for any item in the distinct component model. However, each item's cross-validated  $R^2$  was examined to evaluate the model's suitability.

#### *5.2.2.3 ROSA model interpretation*

The distinct component model was primarily used for further interpretation. Data block importance and interactions were established by considering the selection order in both (common and distinct) models. In addition, the importance of specific variables was determined using VIP and by examining the residuals of individual items and variable combinations. Loadings plots were used to aid in interpreting relationships between variables and components.

##### 5.2.2.3.1 Relative block importance (Block selection order and frequency)

As the ROSA algorithm sequentially selected blocks based on the maximum improvement in prediction, block selection order was used as a block importance measure, with earlier selection indicating greater importance. In addition, as the ROSA algorithm could select each block an unlimited number of times, the frequency of block selection was used as a secondary measure of importance. Finally, selection patterns in the common component model were inspected to evaluate the presence of common components between thermal taste blocks and any other blocks, as common components point to interactions between components in food choice determination.

##### 5.2.2.3.2 Overall variable importance

The overall importance of each variable in the total model was determined using VIP (Biancolillo et al., 2016) by applying the built-in functionality in the `plsVarSel` package. A VIP threshold of 1 (Mehmood et al., 2012) was used for further discussion, explicitly focusing on the ten variables having the highest VIP scores and, therefore,

impacting prediction.

#### 5.2.2.3.3 Component and variable relationships

Plots of loading weights and x-loadings were used to aid in the interpretation and contextualisation of block selection and the interpretation of different importance measures. Loading weights are used to calculate the relevant loadings from the variables present in the selected data block and can directly indicate how the variables in the selected data block relate to the extracted component. Note that in ROSA, loading weights for all other blocks are set to zero (Smilde et al., 2022). In contrast, x-loadings are projections of all X variables on the selected components. X-loadings can be used to determine relationships between variables that do not form part of the selected data block and those used to calculate the components, further contextualising the relationship between thermal taste and other predictive aspects.

For brevity, loadings weights plots were only presented for any thermal taste related components in the distinct component model. In addition, X-loadings plots were presented for all components of the distinct component model. Loadings plots were filtered only to contain the top 30 variables per block, established by choosing the variables with the largest absolute values per component. Only loadings related to the data block from which the component was extracted, and loadings related to thermal taste were discussed.

#### 5.2.2.3.4 Importance of variables for specific items

The impact of individual variables on consumption prediction was evaluated by plotting and examining regression coefficients produced by the model. The approach of examining regression coefficients in PLS models is outlined in Mehmood et al. (2012) and applied in ROSA by Azadi (2021). Absolute coefficients exceeding 0.25

were considered to impact the overall prediction of a specific item.

#### *5.2.2.4 Impact of thermal taste on food consumption*

The impact of important aspects of thermal taste identified in the overall interpretation of the ROSA model (component selection, VIP, regression coefficients) was established by examining regression coefficients in detail. Clustering was used to facilitate interpretation and discussion, allowing for items with similar prediction patterns to be grouped for discussion. Items were clustered by applying AHC using Ward's method to the regression coefficients, facilitating pattern description and discussion. The optimal number of clusters was determined using the `NBClust` (Charrad et al., 2014) package. The `NBClust` function used 23 published indexes (methods) examining optimal clustering to test the optimal number of clusters between 2 and 22. The indexes used included those detailed by Dunn (1974) and Rousseeuw (1987), but also included 21 other indexes, as detailed in Charrad et al. (2014). The number of clusters recommended by most indexes was chosen as the final number. However, the clustering structure was also visually examined. Relevant regression coefficients were plotted and examined for each item.

### **5.3 Results**

The results of the models predicting food choice from the X blocks (presented in section 5.2.1.2) were evaluated from three different perspectives, informing the structure of the results section. First, results are considered from a model perspective, where the overall fit of the two models is presented, as well as the data block selection order of both models and the top VIP values of the distinct component model. Second, results are presented from a data block perspective. This section presents more in-depth results per data block, including x-loadings, and regression coefficients and

revisiting selection orders and VIP values. The data block section is arranged with any thermal taste-related data blocks presented first, followed by the remaining data blocks in order of first selection. Third, results are presented with a focus on thermal taste, including clustering of thermal taste-related regression coefficients (clustering items based on the relative impact of thermal taste on their consumption) and exploring predictive patterns of thermal taste on food choice.

### 5.3.1 Overall model results

#### 5.3.1.1 Model fitting and selection (distinct vs common components)

The first time the distinct component model encountered a reduction in prediction  $R^2$  (unadjusted  $R^2$  considering explained variance of the training data) was when fitting the 21<sup>st</sup> component. At this point, 27 food items showed a slight reduction in explained variance, with the reductions ranging from  $1.07 \times 10^{-14}$  (*Wholemeal bread* and *Cream-based dairy*) to  $2.27 \times 10^{-12}$  (*Nuts*) (Table J-1, Appendix J). At this point, several items also displayed a plateau in their explained variance by the distinct component model (Figure 5-1 a.), indicating a suitable point for terminating model fitting.

After fitting 20 components, the distinct component model explained between 14.5% (*Other Root Vegetables*) and 47.8% (*Processed meat*) (Figure 5-1a) of difference in consumption of specific items between participants (item variation), whereas the common component model explained between 10.6% (*Other cold breakfast cereal*) and 55.2% (*Processed meat*) of item variation (Figure 5-1b). (Appendixes K and L contain expanded versions of Figure 5-1, showing the patterns of explained variance for all items for both models).

Although the cross-validation results were not used to choose the final number of components, the cross-validation  $R^2$  was compared with the unadjusted  $R^2$  to establish

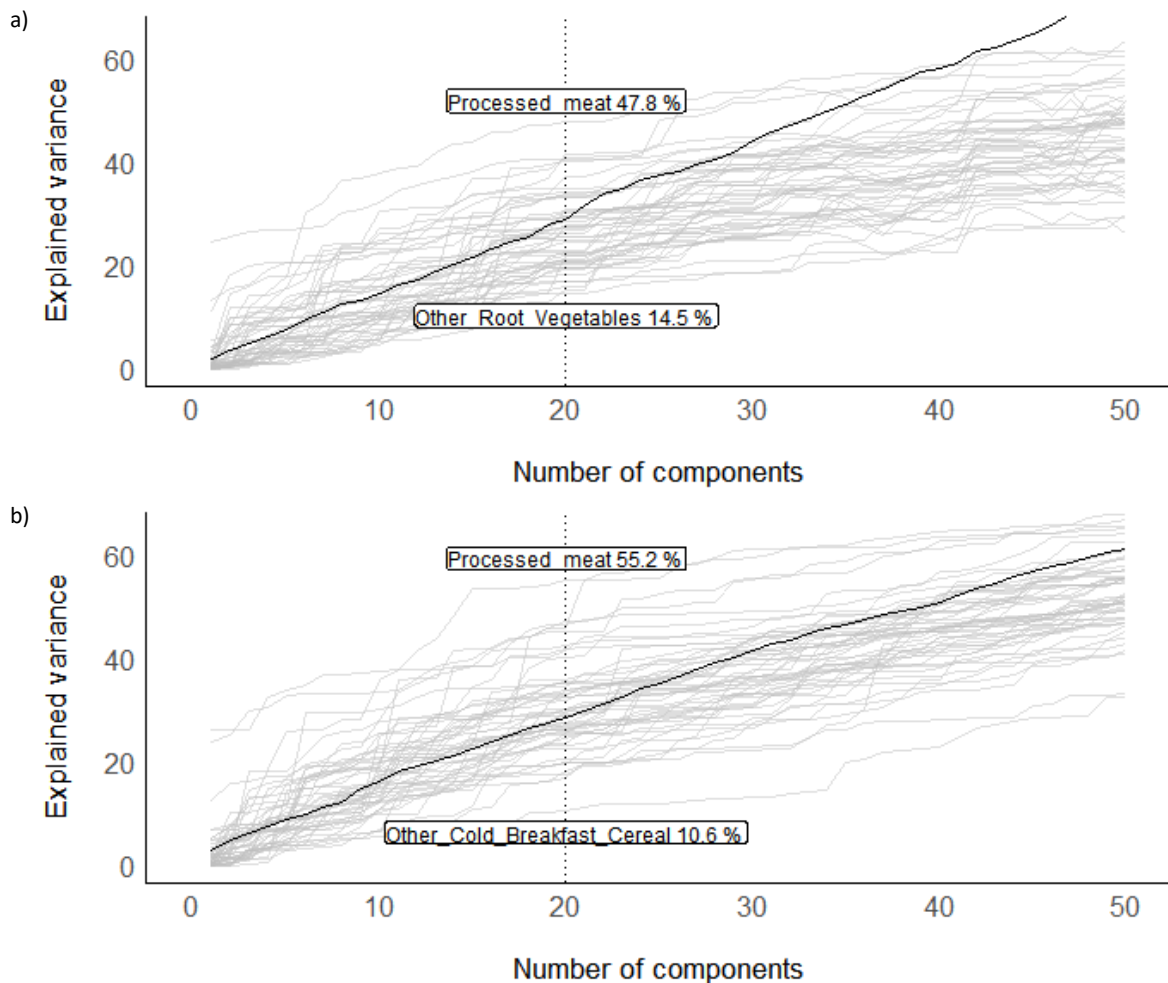


Figure 5-1 Graphical representation of the a) distinct component and b) common component ROSA models. Each grey line indicates the prediction trajectory (increased explained variance with each fitted component) of a single food item. The black line indicates total explained variance of the X space (blocks used for prediction). The final fitted component prior to destabilisation is highlighted with a vertical line, as are the items with the highest and lowest proportion of explained variance at that point. For figures detailing the prediction trajectory of each item, see Appendix K and Appendix L.

patterns of fit and overfit in the model (Figure 5-2). In the associated figures, if the red line is below zero, it indicates a negative cross-validation  $R^2$ , indicating overfitting. Figure 5-2 reveals a couple of different patterns. For some items, for example, *Fats*, *Bananas* and *Processed meat*, the model fits the data up to the stopping point, indicating that the model is more effective at explaining the consumption of these items. For some other items, for example, *Biscuits*, *White rice*, *Low-fat cheese*, *Chocolate/Chocolate bars* and *Tomatoes*, the model fits for smaller component numbers but fit decreased below zero by the 20<sup>th</sup> component. For other items, such as *Poultry*, *Fatty fish*, *Other vegetables*, and *Citrus fruits*, although the model was

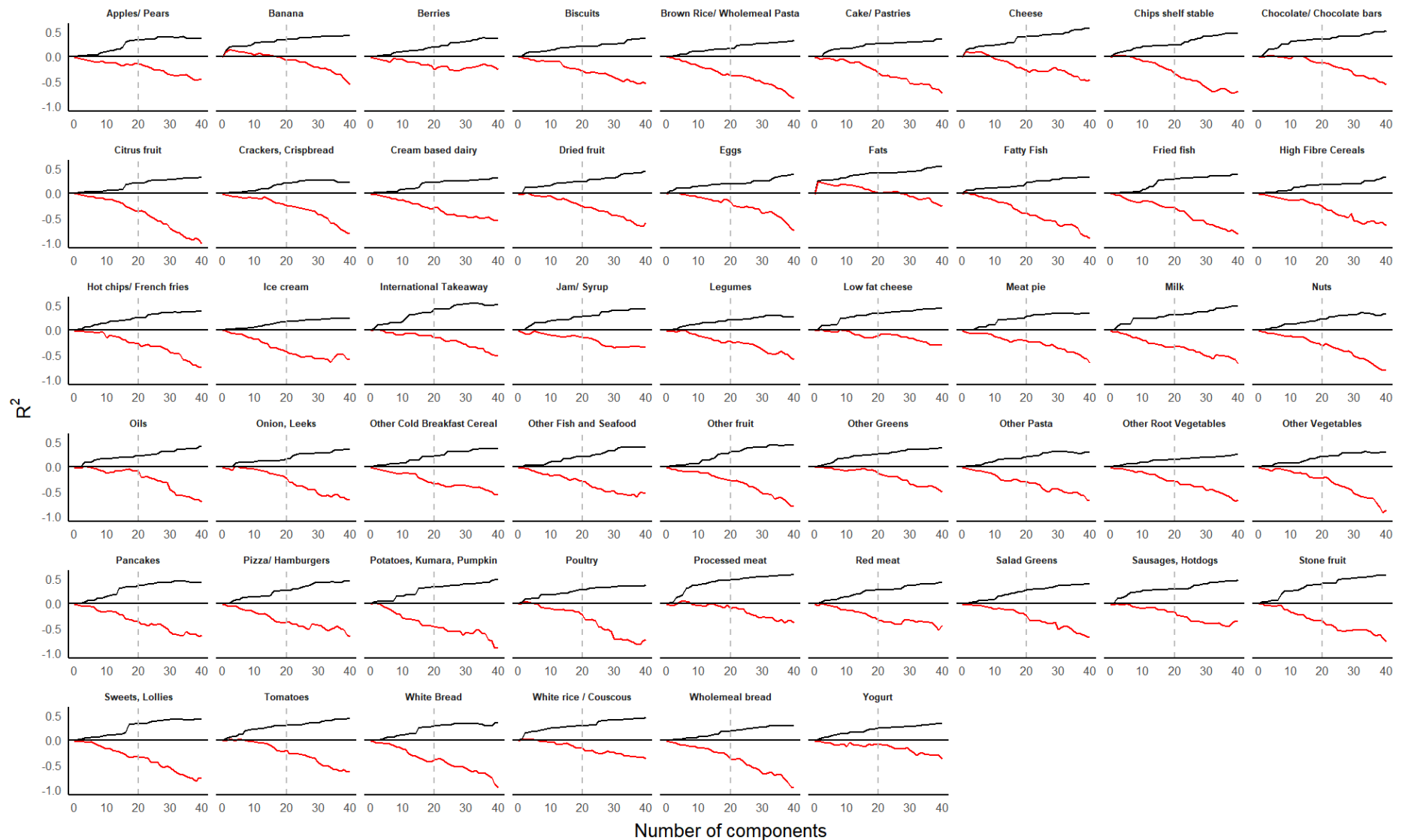


Figure 5-2 Plot of unadjusted  $R^2$  (black line) and cross-validated  $R^2$  (red line) for all items using the distinct component model, with the final number of components (20) chosen highlighted with a vertical dashed grey line. The higher the unadjusted  $R^2$ , the more variation in the data explained, whereas the cross-validated  $R^2$  indicates model fit. Negative cross-validated  $R^2$  indicates an overfitted model.

overfitted, the fit worsened sharply in the 20th component region. Overall, these patterns indicate that the number of components chosen using the unadjusted  $R^2$  was reasonable, as no simpler (fewer components) model fitted all items well, and there were no sharp reductions in fit prior to the 20<sup>th</sup> component.

#### *5.3.1.2 Relative importance of data blocks in predicting food choice*

This section briefly overviews the overall component selection patterns of the distinct and common component models. In addition, VIP results are presented briefly. Section 5.3.2 revisits these aspects of the model by data block to facilitate the overall interpretation of each data block's role in the model.

##### 5.3.1.2.1 Block selection order and patterns

Block selection order was used to establish the relative importance of blocks in the model overall, as the ROSA algorithm selects blocks sequentially based on their ability to predict the outcome variables. In addition, the number of times a block is selected also suggests its importance. Finally, as the selected components (extracted from the blocks) are orthogonal (Liland et al., 2016), it is interesting to consider instances where components from the same block are selected in succession, as this suggests a complex, non-linear predictive relationship. In such cases, the selection patterns imply a sub-model where independent latent factors from the same block have predictive effects of similar importance.

When fitting the distinct component model, the ROSA algorithm selected Demographics (Age/Gender) first, followed by Orosensory responsiveness, Spicy Food Liking (selected twice in succession), Oral Processing and then TT Cluster (Table 5-5). TT Cluster was the only thermal taste-related data block selected by the distinct component model, and it was selected three times (positions 6, 13 and 19, Table 5-5).

Note that for this model, the algorithm tended to select several consecutive (orthogonal) components from the same data block (Spicy Food Liking: positions 3,4; Sweet Liking: positions 10, 11 and 12; Big Five Inventory: positions 15, 16), suggesting sub-models related to these data blocks. In addition, although non-behavioural aspects were selected first and consisted of half of the components in the model (10 in total), direct measures of food behaviour were selected more frequently than indirect measures, with indirect measures only selected twice (Food Involvement: position 8; Food Neophobia: Position 17) (Table 5-5).

*Table 5-5 Component selection order by the distinct and common component model. Component number indicates the position in the model that the component was selected in, with lower numbers (earlier selection) indicating higher importance to the model. In the common component model, where two data blocks are listed, this indicates the selection of a component from the joint space of these two data blocks. Category (Cat) refers to whether the data block is related to thermal taste (T), Direct measures of food behaviour (D), indirect measures of food behaviour (I) or unrelated to either thermal taste or food behaviour (N).*

Component	Distinct component model		Common component model	
	Data block	Cat	Data block	Cat
1	Demographics	N	Demographics X Ethnicity	NxN
2	Orosensory Responsiveness	N	<b>TT Cluster X Demographics</b>	<b>TxN</b>
3	Spicy Food Liking	D	Food Modification X Neophobia	DxI
4	Spicy Food Liking	D	Food Involvement X Sweet Liking	IxD
5	Oral Processing	D	Demographics X Food Involvement	NxI
6	<b>TT Cluster</b>	<b>T</b>	Big Five Inventory X Food Modification	NxD
7	Food Modification	D	Spicy Food Liking	D
8	Food Involvement	I	Oral Processing X Spicy Food Liking	DxD
9	Oral Processing	D	<b>TT Cluster X Demographics</b>	<b>TxN</b>
10	Sweet Liking	D	<b>TT Temperature X Sweet Liking</b>	<b>TxD</b>
11	Sweet Liking	D	Food Modification X Sweet Liking	DxD
12	Sweet Liking	D	Oral Processing X Spicy Food Liking	DxD
13	<b>TT Cluster</b>	<b>T</b>	<b>TT Taste X TT Temperature</b>	<b>TxT</b>
14	Orosensory Responsiveness	N	<b>TT Cluster X Big Five Inventory</b>	<b>TxN</b>
15	Big Five Inventory	N	Big Five Inventory X Orosensory Responsiveness	NxN
16	Big Five Inventory	N	Big Five Inventory X Spicy Food Liking	NxD
17	Neophobia	I	Big Five Inventory X Demographics	NxN
18	Private Body Consciousness	N	Food Modification X Private Body Consciousness	DxN
19	<b>TT Cluster</b>	<b>T</b>	Demographics X Orosensory Responsiveness	NxN
20	Big Five Inventory	N	Oral Processing X Spicy Food Liking	DxD

Looking at components selected when fitting the common component model reveals where different data blocks have common latent factors (common components) in predicting food choice. These components are, by definition, better at predicting food choice than separate (distinct) components extracted from the individual data blocks.

The common component model revealed patterns different from those of the distinct component model. For example, although a component from Demographics was still selected first, this was a component which described both Demographics and Ethnicity (Table 5-5). In addition, components relating to all three thermal taste data blocks were selected. The first thermal taste-related component occurred earlier in the model (component 2: TT Cluster x Demographics), and five thermal taste-related components were selected in total (component 2: TT Cluster x Demographics; component 9: TT Cluster x Demographics, component 10: TT Temperature x Sweet liking; component 13: TT Taste x TT Temperature; component 14: TT Cluster x Big Five Inventory). Except for TT Temperature x Sweet Liking (component 10), thermal taste-related components were either common to other thermal taste-related data blocks (component 13) or to thermal taste and aspects that do not measure food behaviour directly or indirectly (components 2, 9, and 14). These patterns suggest that thermal taste is independent of measures of food behaviour in its effect on food choice.

#### 5.3.1.2.2 Overall variable importance

In the distinct component model, 67 variables (of a possible 148) exceeded the VIP threshold of 1 (Table N1, Appendix N). Overall, agreement with the statement “I find it hard to enjoy the flavour of food when it contains hot spices” was the most important predictor of food choice (VIP = 102.0 Table 5-6). The top ten most important variables included two from Spicy Food Liking, three from Oral Processing, one from TT Cluster, one from Orosensory Responsiveness, two from the Big Five Inventory and one from Demographics (Table 5-6). Further details of VIP values are discussed in section 5.3.2 under the relevant data blocks.

### **5.3.2 Focus on data blocks**

This section describes the results of the ROSA model, organised by data block. Results

Table 5-6 Ten variables from the distinct component ROSA model with the highest variable importance in projection scores.

Rank	Variable	Data block Origin	VIP
1	"I find it hard to appreciate the flavours of food when the food contains hot spices"	Spicy Food Liking	102.0
2	Smoosher	Oral Processing	75.07
3	TT Cluster 2	TT Cluster	19.59
4	Sweetness intensity 12% sucrose	Orosensory Responsiveness	17.82
5	"I am someone who is talkative"	Big Five Inventory	16.09
6	"I think spices make food taste better"	Spicy Food Liking	11.74
7	Chewer	Oral Processing	8.97
8	Age	Demographics	8.50
9	Sucker	Oral Processing	8.09
10	"I am someone who is depressed, blue"	Big Five Inventory	7.38

related to thermal taste are presented first, followed by the remainder of the blocks in the order in which the ROSA algorithm selected them while fitting the distinct component model. The selection order and frequency (Table 5-5) are revisited for each block, followed by a more in-depth presentation of the VIP values (first presented in section 5.3.1.2.2). Next, the components extracted from selected data blocks are explored by presenting loading weights (thermal taste blocks only) and X loadings. Loading weights are used to calculate components, giving insight into the relationships between extracted components and the variables in the data block from which that component was extracted. In contrast, X-loadings are a projection of *all* X- variables on the calculated components, highlighting relationships between variables used in component calculation and those in other data blocks. Finally, large regression coefficients are presented for all blocks, highlighting relationships where specific variables had a notable impact on the consumption prediction of specific items.

### 5.3.2.1 Thermal taste (TT Cluster)

TT Cluster was the only thermal taste-related data block selected while fitting the distinct component model, and it was selected in positions 6, 13, and 19. While fitting the common component model, components related to all three thermal taste-related data blocks were selected. Components common to TT Cluster and Demographics (positions 2 and 9) and Big Five Inventory (position 14) were selected. In addition,

components common to TT Temperature and Sweet Liking (position 10) and TT temperature and TT Taste (position 13) were selected.

In the distinct component model, two of the three variables in the TT Cluster data block met the VIP threshold of 1, with Cluster 2 membership being one of the best-performing predictors overall (VIP = 19.59 Table 5-6), with Cluster 1 membership barely meeting the criteria for inclusion (VIP = 1.07 Table N1 Appendix N).

Loading weights highlight prominent variables in calculating specific components. In ROSA, as loading weights of blocks not selected are set to zero (Smilde et al., 2022), loading weights also allow for interpretation of variable impact in isolation from relationships with variables from other blocks. Loading weights plots (Figure 5-3a) show that Cluster 2 and Cluster 1 memberships were used in calculating Component 6. In contrast, the difference between Cluster 3 (positive weight) and Cluster 1 (negative weight) was used when calculating Components 13 and 19 (Figure 5-3a) because those were the largest weights in the positive and negative directions.

In contrast to loading weights, loadings highlight the relationships between variables and components, and can specifically show how variables that are not contained in the data block used to calculate the component are related to that component. Figure 5-3b shows that two of the components extracted from the TT Cluster data block (Component 6 and 19) explained a large portion of the RapCoTT trial (section 2.2.1.2.2), with a section spanning from 25°C during cooling to 40°C during warming (“Temp 25C”, “Temp 20C”, “Temp 15C”, “Temp 10C”, “Temp 5C”, “Temp 10W”, “Temp 15W”, “Temp 20W”, “Temp 25W”, “Temp 30W”, “Temp 35W”) having high positive loadings for both components. In contrast, a section at the start of the cooling trial portion between 30°C and 15°C (“Temp 30C”, “Temp 25C,” “Temp 20C”) and one at

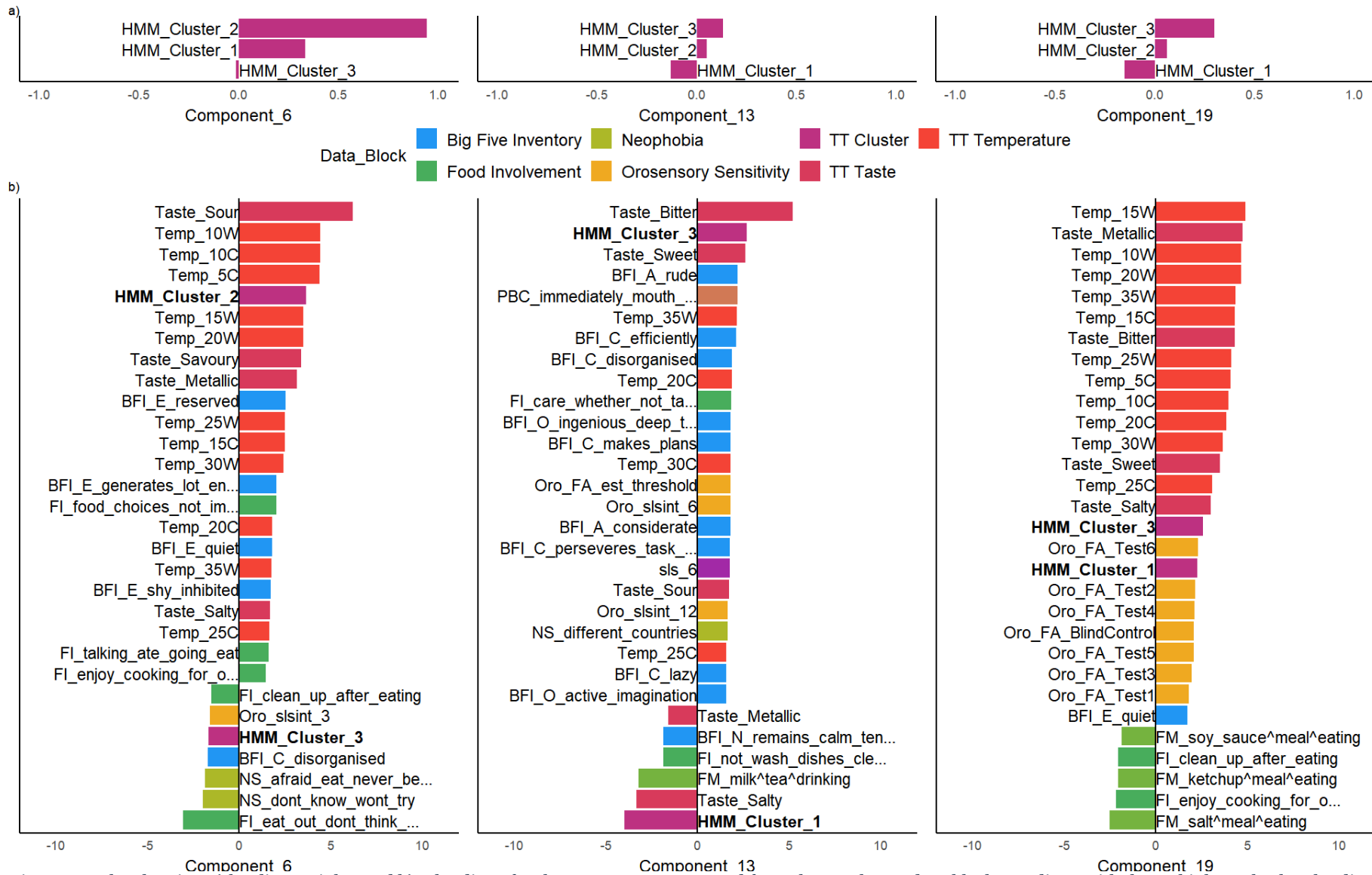


Figure 5-3 Plot showing a) loading weights and b) x-loadings for the components extracted from the TT Cluster data block. Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M

the end of the warming portion (“Temp 35W”) had a high positive loadings for Component 13.

All three components encoded the nature (identity) of thermal taste reported. Component 6 had the highest positive loading for Sour thermal taste and high positive loadings for Savoury and Metallic thermal taste. Component 13 had the highest positive loading for Bitter thermal taste, but also had positive loadings for Sweet thermal taste and negative loadings for Salty and Metallic thermal tastes. Component 19 had positive loadings for Metallic, Bitter, Sweet and Salty thermal taste. Therefore, overall, the components extracted from the TT Cluster data block encoded information from all three thermal taste data blocks.

Regression coefficients are helpful in identifying specific relationships between predictor variables and the consumption of items, as well as the nature of those relationships (positive or negative). Positive regression coefficients can be interpreted as an indicator that an increase in the relevant variable (or the presence of that variable in binary cases) is linked to an increase in consumption of that item. Similarly, negative regression coefficients indicate that an increase in the variable in question is linked to a decrease in consumption.

As TT Cluster was the only thermal taste-related data block selected during the fitting of the distinct component model, only variables from the TT Cluster data block had the potential to affect consumption prediction. TT Cluster affected the consumption prediction of 38 items (Table 5-7).

Cluster 1 membership had a notable positive effect on *Other Pasta* consumption. Cluster 2 membership had a notable positive effect on *Processed meat, Sausages, Hotdogs* and *Pizza/Hamburgers* consumption. Cluster 3 membership had a notable

Table 5-7 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the TT Cluster data block. The sign of the coefficient (+/-) relates the impact on prediction (positive/negative).

Variable	Positive contribution		Negative contribution	
	Item	coeff	Item	coeff
Cluster1	<i>Other Pasta</i>	0.577	Salad Greens	-0.434
	Other Vegetables	0.439	Eggs	-0.356
	International Takeaway	0.422	Cake/Pastries	-0.322
	Nuts	0.398	Berries	-0.295
	Citrus fruit	0.334	Yogurt	-0.269
	Hot chips/French fries	0.319	Poultry	-0.268
	Other Root Vegetables	0.290	Legumes	-0.258
	Stone fruit	0.288		
	Fried fish	0.272		
	Pancakes	0.262		
Cluster2	<i>Processed meat</i>	0.858	Other Greens	-0.498
	<i>Sausages, Hotdogs</i>	0.757	Legumes	-0.498
	<i>Pizza/Hamburgers</i>	0.550	High Fibre Cereals	-0.488
	Tomatoes	0.460	Yogurt	-0.391
	International Takeaway	0.434	Nuts	-0.383
	Oils	0.412	Cake/Pastries	-0.380
	White rice / Couscous	0.394	Chocolate/Chocolate bars	-0.361
	Other Pasta	0.389	Berries	-0.353
	Hot chips/French fries	0.379	Potatoes, Kumara, Pumpkin	-0.311
	Low fat cheese	0.342	Wholemeal bread	-0.295
	Sweets, Lollies	0.271	Brown Rice/Wholemeal Pasta	-0.287
	Dried fruit	0.269	Crackers, Crispbread	-0.274
	White Bread	0.256	Stone fruit	-0.268
			Other Root Vegetables	-0.266
Cluster3	<i>Sweets, Lollies</i>	0.576	<i>Apples/Pears</i>	-0.526
	<i>Crackers, Crispbread</i>	0.528	<i>Legumes</i>	-0.501
	Meat pie	0.457	Berries	-0.404
	Hot chips/French fries	0.411	Yogurt	-0.367
	Other Vegetables	0.373	White Bread	-0.359
	Nuts	0.315	Brown Rice/Wholemeal Pasta	-0.350
	Ice cream	0.280	High Fibre Cereals	-0.274
	Cheese	0.276	Oils	-0.272

positive effect on *Sweets, Lollies* and *Crackers, Crispbread*, and a notable negative effect on *Apples/Pears* and *Legumes* consumption prediction. The large number of items affected by TT Cluster, and the presence of notable residuals of affected items, justify the additional exploration of this aspect of food choice behaviour.

### 5.3.2.2 Demographics

The Demographics data block, encoding Age and Sex/Gender, was the first data block selected during the fitting of the distinct component model. Note that the other data block related to demographics, Ethnicity, was not selected. In the common component model, components common to Demographics and Ethnicity (component 1), TT

Cluster (components 2 and 9), Food Involvement (component 5), Big Five Inventory (component 17) and Orosensory sensitivity (component 19) were selected.

In the distinct component model, one of the two variables in the Demographics data block, namely Age, met the VIP threshold of 1 (Table N-1, Appendix N) and the extracted component was primarily calculated using Age, as demonstrated by the large absolute loading value (Figure 5-4). When considering the relationship between thermal taste and Component 1, one taste and three consecutive temperature sections had high absolute loadings. Reporting Savoury thermal taste, as well thermal taste between 10°C and 20°C in the warming trial portion (“Temp 10W”, “Temp 15W”, “Temp 20W”) were negatively associated with this component, but positively associated with Age (Figure 5-4).

Demographic variables significantly impacted consumption prediction for five items. The effect for two items was positive (*Fats* and *Cheese*) whereas it negatively affected consumption prediction of *Banana*, *Fatty Fish* and *Eggs* (Table 5-8). For all five items, consumption prediction was affected by the variable Age.

Table 5-8 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Demographics data block. The sign of the coefficient (+/-) relates the impact on prediction (positive/negative)

Variable	Positive Contribution		Positive Contribution	
	Item	coeff	Item	coeff
Age	Fats	0.437	Banana	-0.299
	Cheese	0.373	Fatty Fish	-0.265
			Eggs	-0.250

### 5.3.2.3 Orosensory response

When fitting the distinct component model, the data block describing Orosensory response was selected twice; initially in the second position, and again in the 14<sup>th</sup> position. Components common to Orosensory response and the Big Five Inventory (component 15) and Demographics (component 19) were selected when fitting the

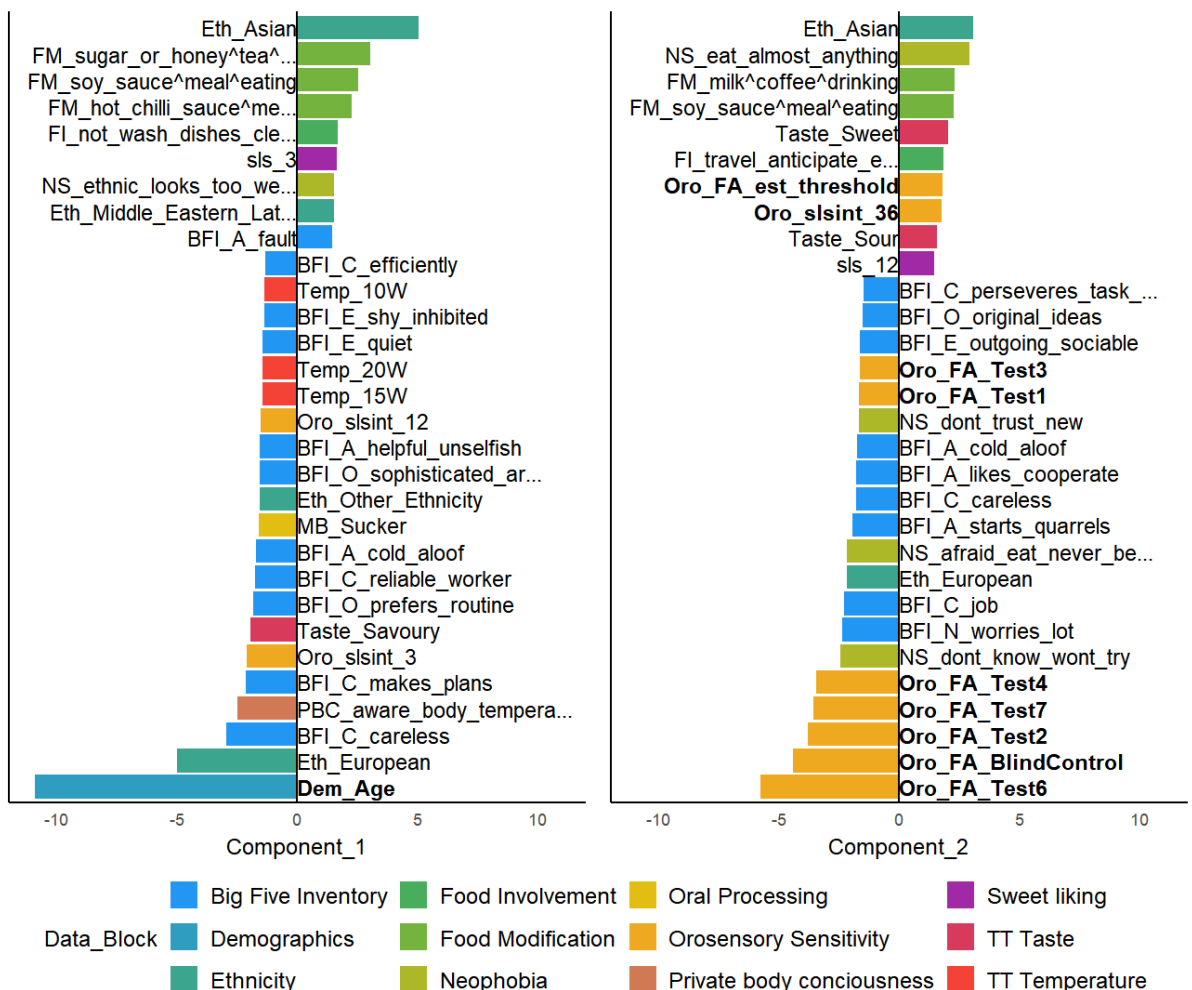


Figure 5-4 Plot showing x-loadings for Component 2 (Demographics) and Component 2 (Orosensory response). Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M

common component model.

In the distinct component model, 12 of the 15 variables in the data block met the VIP threshold of 1 (Table N-1, Appendix N), and one variable (sweetness intensity of the 12% sucrose solution) was included in the top 10 predictors (Table 5-6).

When considering the loadings of Component 2, only two variables from thermal taste were included in the top thirty correlated variables. Reporting Sour and Sweet thermal taste both had positive loadings on Component 2 (Figure 5-4), which was negatively correlated with most of the orosensory response variables, except for the estimated fatty acid threshold and the hedonic response to the intensity of the 36% sucrose

solution (Figure 5-4). Responses to most of the fatty acid samples and two sucrose concentrations (6% and 36%) loaded negatively to Component 14 (Figure O-1, Appendix O). Five thermal taste-related variables also loaded negatively on Component 14, including one taste (Savoury) and four consecutive temperature periods (Figure O-1, Appendix O). Reporting thermal taste between 30°C and 10°C during the cooling trial portion (“Temp 30C”, “Temp 25C”, “Temp 20C” and “Temp 15C”) was negatively associated with Component 14

Fatty acid responsiveness was the only aspect of orosensory response that impacted consumption prediction of specific items, influencing 11 items. Ten of the 11 items were influenced by the response to the level 5 Oleic acid sample - for four (*White rice / Couscous, Red meat, Fried fish and Sausages, Hotdogs*) this had a positive effect on consumption prediction, whereas for six items (*Low fat cheese, Dried fruit, Wholemeal bread, Cheese, Tomatoes, and Banana*) this had a negative effect on consumption prediction (Table 5-9). The remaining item, *Pancakes* was influenced by the response to the level 7 Oleic acid sample.

Table 5-9 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Orosensory sensitivity data block. The sign of the coefficient (+/-) relates the impact on prediction (positive/negative).

Variable	Positive Contribution		Negative Contribution	
	Item	coeff	Item	coeff
FA_Test5	White rice / Couscous	0.321	Low fat cheese	-0.345
	Red meat	0.294	Dried fruit	-0.337
	Fried fish	0.266	Wholemeal bread	-0.281
	Sausages, Hotdogs	0.253	Cheese	-0.274
FA_Test7	Pancakes	0.289	Banana	-0.269
			Tomatoes	-0.258

#### 5.3.2.4 Spicy Food Liking

Spicy Food Liking related components were selected twice while fitting the distinct component model; in the third and fourth positions. When fitting the common component model, components from this data block were selected five times in total,

first in isolation (Component 7), three times in combination with Oral Processing (Components 8, 12 and 20) and in combination with Big Five Inventory (component 16).

In the distinct component model, all six variables met the VIP threshold of 1 (Table N-1, Appendix N), and two variables (“I find it hard to appreciate hot food” and “Spices make food taste better”) were included in the top ten predictors (Table 5-6).

Component 3 mainly described the difference between agreeing that foods taste bland without spices (positive loadings), spicy food consumption frequency and the liking of the taste of spices (negative loadings). Here, reporting sweet thermal taste and thermal taste between 10°C and 15°C during the warming trial portion (“Temp 10W”) had notable negative loadings (Figure 5-5). In contrast, Component 4 described the difference between spice consumption level, agreeing that foods taste bland without spices (negative loadings), and finding it difficult to appreciate hot food. Most of the cooling trial portion loaded positively to this component – starting from 30°C to the start of the warming trial portion (“Temp 30C”, “Temp 25C”, “Temp 20C”, “Temp15C”, “Temp10C”, “Temp 5C”). In addition, reporting Bitter thermal taste and reporting thermal taste at the end of the warming trial portion (between 30°C and 40°C; “Temp 30W” and “Temp 35W”) loaded positively to this component.

Consumption prediction of ten items in total was affected by items on the spicy food liking questionnaire (Table 5-10). Here, the hedonic response to the burn of spices affected most items (seven). Notably, consumption prediction for three items (*Processed Meat* and *Pizza/Hamburgers* and *Jam/Syrup*) was influenced by three different items on the spicy food liking questionnaire. The hedonic response to the burn of spices affected consumption prediction positively for both items, whereas

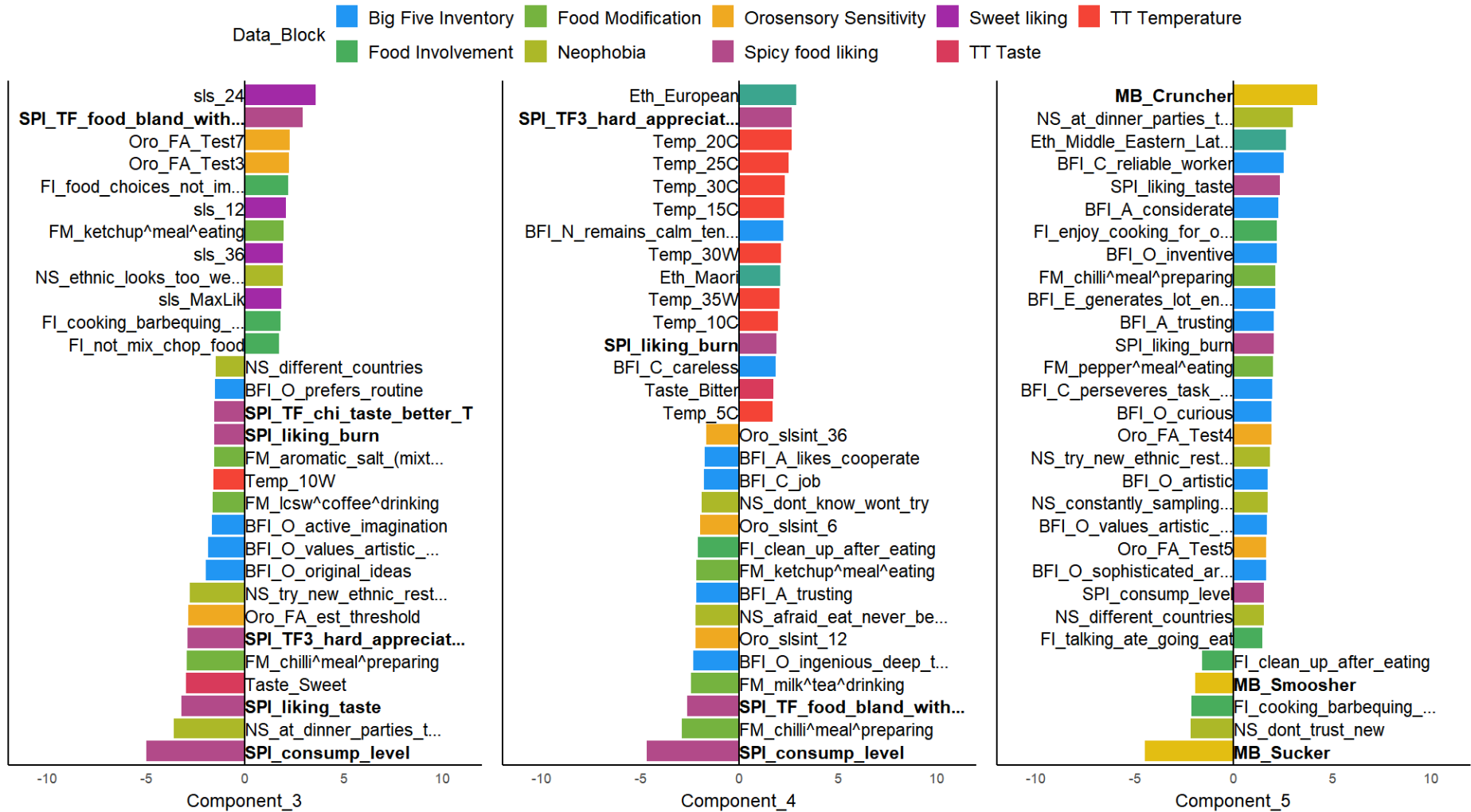


Figure 5-5 Plot showing x-loadings for Component 3 (Spicy Food Liking), Component 4 (Spicy Food Liking) and Component 5 (Oral processing). Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M

Table 5-10 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Spicy Food Liking, Oral Processing, Food Modification and Food involvement data blocks. The sign of the coefficient (+/-) relates the impact on prediction (positive/negative).

Data Block	Variable	Positive		Negative	
		Item	coeff	Item	coeff
Spicy Food Liking	Hedonic response burn of spices	Processed meat	0.405	Onion, Leeks	-0.391
		Jam/Syrup	0.389	Legumes	-0.324
		Pizza/Hamburgers	0.336		
		Sausages, Hotdogs	0.331		
		Apples/Pears	0.255		
	I find it hard to appreciate the flavours of food when it contains hot spices	Oils	0.356	Cake/Pastries	-0.416
		Apples/Pears	0.256	Pancakes	-0.261
	I think spices make food taste better			Processed meat	-0.393
				Jam/Syrup	-0.326
				Pizza/Hamburgers	-0.297
	Spice consumption frequency	Onions, leeks	0.267	Processed meat	-0.400
				Jam/Syrup	-0.344
Pizza/Hamburgers				-0.309	
Without hot spices, I find that food tastes bland	Cake/Pastries	0.416	Oils	-0.356	
	Pancakes	0.261	Apples/Pears	-0.256	
Oral Processing	Smoocher	Cream based dairy	0.406	Other Greens	-0.359
		White bread	0.335	Other Root Vegetables	-0.306
		Cake/Pastries	0.311	Yogurt	-0.257
		Chips shelf stable	0.293		
		Meat pies	0.293		
	Sucker	Jam/Syrup	0.400	Salad Greens	-0.359
		Processed meat	0.335	Tomatoes	-0.357
		Meat pie	0.293	Other Cold Breakfast Cereal	-0.284
		Sausages, Hotdogs	0.277	Other Vegetables	-0.263
		Cake/Pastries	0.270		
		Hot chips/French fries	0.256		
Food Modification	Adding low-calorie/ non-nutritive sweetener to tea	Other Greens	0.257	Milk	-0.272
				Tomatoes	-0.253
Food Involvement	I enjoy cooking for myself and others			Low fat cheese	-0.259

spice consumption frequency and agreement with the item “I think spices make food taste better” negatively affected consumption prediction, suggesting two different underlying phenomena, likely related to the Component 3 and Component 4, respectively.

#### 5.3.2.5 Oral Processing

When fitting the distinct component model, components from the Oral Processing data block were selected twice, in the fifth and ninth positions. In the common component model, the three selected components from Oral Processing were always common to Spicy Food Liking (Components 8, 12 and 20).

In the distinct component model, three of the four variables from this block met the VIP threshold of 1, and all three of these variables (“Chewer”, “Smoosher” and “Sucker”) were included in the top 10 most important variables (Table 5-6).

Component 5 mainly encoded the difference between the “Cruncher” (positive loading) and “Sucker” (negative loading) oral processing groups (Figure 5-5). In contrast, Component 9 encoded the difference between the “Smoosher” (positive) and “Sucker” (negative) groups. Reporting Sour thermal taste was the only thermal taste-related variable with a high loading for Oral Processing; it loaded positively to Component 9 (Figure 5-6). Overall, this indicates that oral processing behaviour is independent of thermal taste in determining food choice.

Oral Processing behaviour grouping positively affected the consumption prediction of 16 items (Table 5-10), with only “Smoosher” and “Sucker” memberships influencing the prediction of specific items. Consumption prediction of ten items was positively affected, with both subgroups (“Smoosher” and “Sucker”) having an effect on *Meat Pie* and *Cakes/Pastries* consumption prediction (Table 5-10). In addition, the

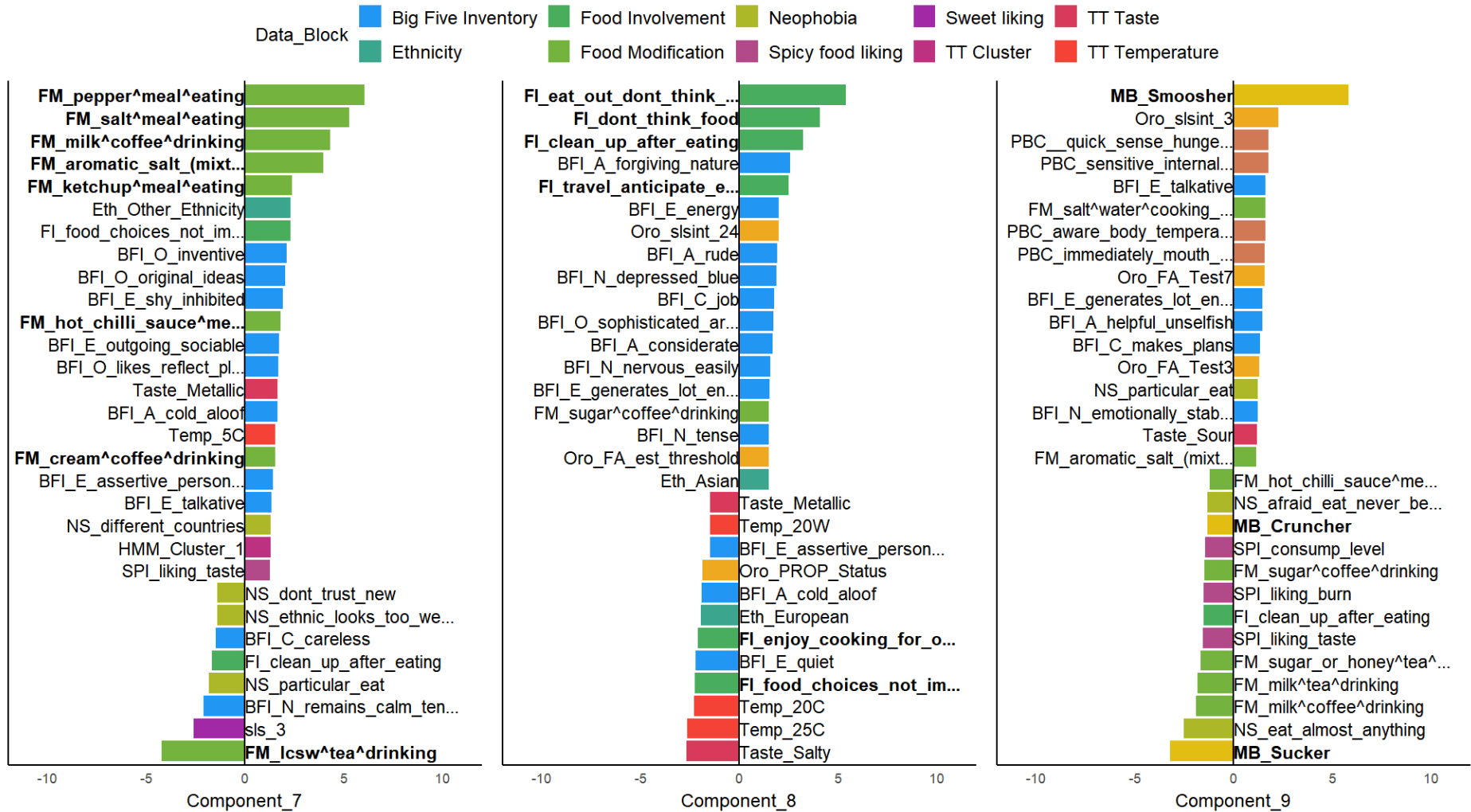


Figure 5-6 Plot showing x-loadings for Component 7 (food modification), Component 8 (food involvement) and Component 9 (Oral Processing). Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M

consumption prediction of seven items was negatively affected by Oral processing behaviour subgroup.

#### 5.3.2.6 Food Modification

When fitting the distinct component model, only one component from the Food Modification data block was selected, in the seventh position. In the common component model, components common to Food Neophobia (Component 3), Sweet liking (Component 11) and Private Body Consciousness (Component 18) were selected.

Seven of the 15 variables in this data block met the VIP threshold of 1 in the distinct component model (Table N-1, Appendix N), but no variables were included in the top ten predictors. The addition of pepper (positive), salt (positive), aromatic salt (positive) or ketchup (positive) to food, milk to coffee (positive) and non-nutritive sweetener to tea (negative) were the variables with the highest loadings to Component 7 (Figure 5-6). One variable from each thermal taste-related data block (TT Cluster 1, Metallic and “Temp 5C”) had a positive loading on Component 7 (Figure 5-6).

One variable from the Food Modification data block, namely “Adding low-calorie/non-nutritive sweetener to tea”, affected consumption prediction of three items; *Other Greens* was affected positively, whereas *Milk* and *Tomatoes* were affected negatively (Table 5-10).

#### 5.3.2.7 Food Involvement

In the distinct component model, Food Involvement was selected only once, in the eighth position. In the common component model, components from Food Involvement were selected twice, combined with Sweet Liking (Component 4) and Demographics (Component 5).

In the distinct component model, five of the twelve variables met the VIP threshold of 1 (Table N-1, Appendix N), and no variables were included in the top ten variables. Variables describing whether participants did not generally think about food (“I don’t think much about food each day”) or did not talk about food when eating out (“When I eat out, I don’t think or talk much about how the food tastes”) had positive loadings, whereas enjoyment of cooking, but also not seeing food choices as important had negative loadings (Figure 5-6). Reporting Salty and Metallic thermal taste had negative loadings to Component 8. In addition, reporting thermal taste between 25°C and 15°C on the cooling trial portion (“Temp 25C”, “Temp 20C”) or between 20°C and 25°C (“Temp 20W”) on the warming trial portion loaded negatively to Component 8 (Figure 5-6). Food involvement only affected the consumption prediction of a single item; the enjoyment of cooking negatively affected the prediction of *Low Fat cheese* consumption (Table 5-10).

#### 5.3.2.8 Sweet Liking

When fitting the distinct component model, three components were selected from the Sweet Liking data block in the tenth, eleventh and twelfth positions. In the common component model, Sweet Liking was selected in combination with Food Involvement (Component 4), TT Temperature (Component 10) and Food Modification (Component 11). In the distinct component model, five of the nine variables in the data block met the VIP threshold of 1 (Table N-1, Appendix N) but no variables from this block were included in the top ten variables.

The most liked sucrose concentration, and the liking of 36% sucrose solution loaded positively to Component 10, whereas liking of the 24% sucrose solution loaded negatively to this component 1 (Figure 5-7). The only thermal taste-related variable strongly loading Component 10 was the reporting of bitter thermal taste. Hedonic

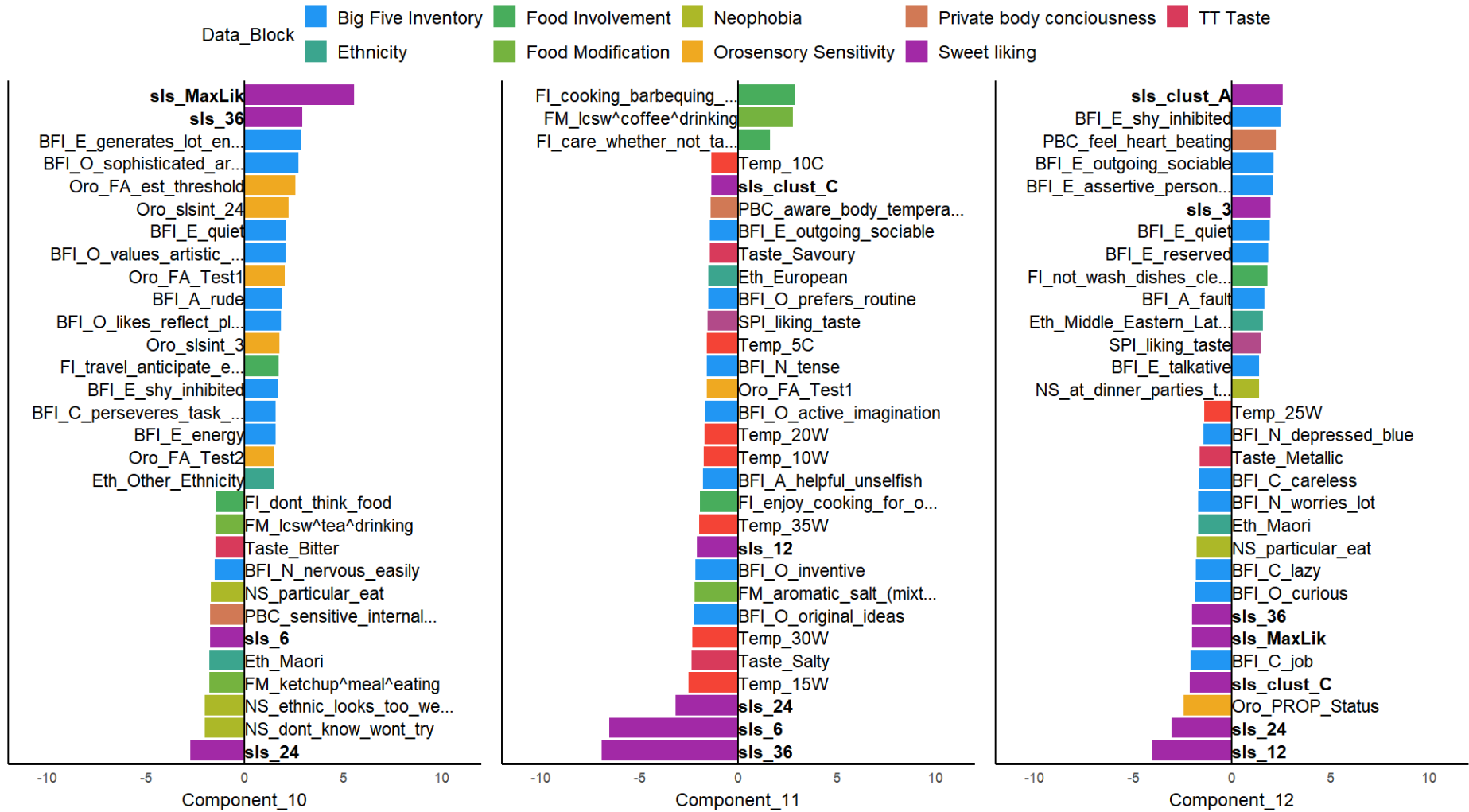


Figure 5-7 Plot showing x-loadings for Components 10, 11 and 12, extracted from the Sweet Liking data block. Loadings with the 30 highest absolute loadings are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M

response to three sucrose concentrations (36%, 24% and 6%) loaded negatively to Component 11 (Figure 5-7). Reporting Savoury and Salty thermal taste loaded negatively to Component 11, as did a large portion of the thermal exposure, stretching from the end of the cooling trial portion (10°C), throughout the warming trial portion (35°C), with only the period between 25°C and 30°C not having a notably negative loading (“Temp 10C”, “Temp 5C”, “Temp 10W”, “Temp 15W”, “Temp 20W”, “Temp 30W”, “Temp 35W”). Hedonic response to 12% and 24% sucrose solutions loaded negatively to Component 12 (Figure 5-7), as did reporting Metallic thermal taste and thermal taste between 25°C and 30°C during the warming trial portion (“Temp 25W”).

Different aspects of the hedonic response to sucrose affected the consumption of most items (38 in total). Here, High Sweet Liking (HSL) and Medium Sweet Liking (MSL) had the greatest effect, with the two variables affecting consumption prediction for 29 and 20 items, respectively. HSL had a notably large positive effect on *Other fruit* (0.842), *Apples/ Pears* (0.783), *Chips shelf stable* (0.720), *Sausages, Hotdogs* (0.696), *Meat pie* (0.633), *Low fat cheese* (0.522), *Wholemeal bread* (0.516), and *Cheese* (0.503) consumption prediction. Unclassified sweet liker status had a notably large effect on consumption prediction for *Other fruit* (0.527). In addition, HSL had a notably large negative effect on *Red meat* (-0.929), *Other Fish and Seafood* (-0.695) and *White Bread* (-0.663) consumption prediction (Table 5-11).

#### 5.3.2.9 Big Five Inventory

When fitting the distinct component model, components extracted from the Big Five Inventory data block were selected three times, although only occurring later in the model (Components 15, 16 and 20). In the common component model, components from Big Five Inventory were selected in combination with Food Modification (Component 6), Orosensory sensitivity (Component 14), Spicy Food Liking

Table 5-11 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Sweet Liking data block. The sign of the coefficient (+/-) relates the impact on prediction (positive/negative).

Variable	Positive		Negative	
	Item	coeff	Item	coeff
High Sweet Liking	<i>Other fruit</i>	0.842	<i>Red meat</i>	-0.929
	<i>Apples/Pears</i>	0.783	<i>Other Fish and Seafood</i>	-0.733
	<i>Chips shelf stable</i>	0.720	<i>White Bread</i>	-0.663
	<i>Sausages, Hotdogs</i>	0.696	Salad Greens	-0.481
	<i>Meat pie</i>	0.633	Berries	-0.447
	<i>Low fat cheese</i>	0.522	International Takeaway	-0.442
	<i>Wholemeal bread</i>	0.516	Poultry	-0.428
	<i>Cheese</i>	0.503	Fatty Fish	-0.413
	Hot chips/French fries	0.424	High Fibre Cereals	-0.396
	Biscuits	0.418	Onion Leeks	-0.370
	Citrus fruit	0.393	Chocolate/Chocolate bars	-0.350
	Pancakes	0.342	Crackers, Crispbread	-0.310
	Dried fruit	0.333	Eggs	-0.257
	Ice cream	0.306		
	Brown Rice/Wholemeal Pasta	0.305		
	Processed meat	0.258		
Medium Sweet Liking	Fried fish	0.480	Salad Greens	-0.477
	Meat pie	0.435	Crackers, Crispbread	-0.467
	Pancakes	0.433	Other Greens	-0.399
	Apples/Pears	0.376	Other Cold Breakfast Cereal	-0.391
	Hot chips/French fries	0.356	Yogurt	-0.332
	Banana	0.352	Biscuits	-0.289
	International Takeaway	0.343	Fatty fish	-0.287
	Wholemeal bread	0.333		
	Ice cream	0.311		
	Fats	0.278		
	Jam/Syrup	0.276		
	Processed meat	0.270		
Chips shelf stable	0.266			
Unclassified Sweet Liking	<i>Other fruit</i>	0.527	High Fibre Cereals	-0.346
	Meat pie	0.402	Legumes	-0.298
	Dried fruit	0.361	Red meat	-0.280
	Other Pasta	0.252	White Bread	-0.279
			Banana	-0.274
Sucrose liking 36	International Takeaway	0.395	Fried fish	-0.265
Sucrose liking 12			Cheese	-0.254
			International Takeaway	-0.252

(Component 15) and Demographics (Component 17).

In the distinct component model, eighteen out of the 44 variables met the VIP threshold of 1 (Table N-1, Appendix N) and two variables, “I am someone who is talkative” and “I am someone who is depressed, blue” were included in the top ten predictors (Table 5-6).

Component 15 mainly had variables from the Neuroticism, Openness to Experience and Agreeableness with large absolute loadings (Figure O-1 Appendix O). In addition,

reporting Metallic thermal taste and reporting thermal taste in the period between 30°C in the cooling trial portion and 15°C in the warming trial portion (“Temp 30C”, “Temp 25C”, “Temp 20C”, “Temp 10C”, “Temp 5C”, “Temp 10W”) loaded negatively to this component (Figure O-1 Appendix O).

Component 16 mainly had variables from the Neuroticism, Openness to Experience and Conscientiousness with large absolute loadings (Figure O-1 Appendix O). Reporting Salty or Bitter thermal taste had positive loadings to this component. In addition, three sections of the RapCoTT trial had positive loadings to this component; between 20°C and 10°C during the cooling trial portion (“Temp 20C”, “Temp 15C”), between 5°C and 10°C during warming (“Temp 5C) and between 15°C and 30°C during the warming trial portion (“Temp 15W”, “Temp 20W”, “Temp 25W”). Component 20 included all five dimensions of the BFI in its top 30 variables and did not include any thermal taste-related variables in this set (Figure O-2, Appendix O).

Three items from the BFI affected consumption prediction of five items (Table 5-12). The reversed item related to being shy and inhibited positively affected consumption prediction of *Apples/Pears*, *Citrus Fruit*, and *Other Fruit* – indicating that higher extraversion affects consumption prediction. In addition, being original positively affected *Citrus Fruits* consumption prediction and negatively affected *Cream-based dairy* consumption prediction.

Table 5-12 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Big Five Inventory data block. The sign of the coefficient (+/-) relates the impact on prediction (positive/negative).

Variable	Positive Contribution		Negative Contribution	
	Item	coeff	Item	coeff
Is sometimes shy, inhibited (E)	Apples/ Pears	0.339	Pizza/ Hamburgers	-0.290
	Other fruit	0.299		
	Citrus fruit	0.287		
Is depressed, blue (N)	Apples/ Pears	0.267		
Is original, comes up with new ideas (O)	Citrus fruit	0.259	Cream based dairy	-0.270

#### 5.3.2.10 Neophobia

When fitting the distinct component model, one component from the Food Neophobia Scale was selected, in the 17<sup>th</sup> position. In the common component model, a common component from Food Neophobia and Food modification was selected in the 3<sup>rd</sup> position. In the distinct component model, five of the ten variables met the VIP threshold of 1 (Table N-1, Appendix N), and no variables were included in the top ten predictors.

Items related to food adventurousness had high positive loadings to Component 17, and items related to the avoidance of unfamiliar foods were negatively loaded onto this component (Figure O-2 Appendix O). In addition, reporting Sweet thermal taste or reporting thermal taste between 25°C and 35°C during the warming trial portion (“Temp 25W”, “Temp 30W”) loaded negatively to this component.

Neophobia only affected consumption prediction of four items, although for two of these items (*Cheese* and *Sweets, Lollies*), several items from the FNS affected prediction in different directions. Variables which positively affected Cheese consumption prediction (“At dinner parties, I will try a new food” and “I like foods from different countries”) affected *Sweets, Lollies* consumption prediction negatively. Similarly, variables which affected *Cheese* consumption prediction negatively (“I like to try new ethnic restaurants” and “Ethnic food looks too weird to eat”) affected *Sweets, Lollies* consumption prediction positively. Trying new ethnic restaurants notably positively affected *Sweets, Lollies* consumption (Table 5-13).

#### 5.3.2.11 Private Body Consciousness

In the distinct component model, one component from the Private Body Consciousness data block was selected, in the 18<sup>th</sup> position. In the common component

Table 5-13 Food item and regression coefficient (coeff) for variables where absolute coefficient exceeds 0.25 for variables from the Food Neophobia data block. The sign of the coefficient (+/-) relates the impact on prediction (positive/negative).

Variable	Positive Contribution		Negative Contribution	
	Item	value	Item	value
I like to try new ethnic restaurants	Sweets, Lollies	0.462	Cheese	-0.377
			Wholemeal bread	-0.262
At dinner parties, I will try a new food	Cheese	0.319	Sweets, Lollies	-0.391
I like foods from different countries	Cheese	0.263	Sweets, Lollies	-0.322
I don't trust new foods			Sweets, Lollies	-0.265
Ethnic food looks too weird to eat	Sweets, Lollies	0.264	Cheese	-0.254

model, Private Body Consciousness was also selected in the 18<sup>th</sup> position, in combination with Food Modification. When fitting the distinct component model, two of the five variables in this data block met the VIP threshold of 1, and neither were included in the top ten predictors.

Component 18 mainly described the difference between awareness of one's mouth being dry (positive), awareness of temperature (positive) and, sensitivity to hunger (negative) and bodily tensions (negative) (Figure O-2 Appendix O). Reporting thermal taste between 10°C and 20°C during the warming trial portion ("Temp 10W", "Temp 15W") loaded negatively to this component. No variables from this block significantly affected the consumption prediction of any items.

### 5.3.3 Impact of aspects of TT on food consumption prediction

As the regression coefficients of different TT clusters (as established in Chapter 3) showed the most potential for further analysis, these were chosen. Results from the NBClust procedure recommended either three or 22 clusters, as six indexes indicated each respective clustering amount to be the correct one (full results: 2 clusters: 4 indexes; 3 clusters: 6 indexes; 5 clusters: 2 indexes; 6 clusters: 1 index; 12 clusters: 1 index; 15 clusters: 1 index; 20 clusters: 1 index; 21 clusters: 1 index; 22 clusters: 6 indexes). Therefore, the products were arranged in three clusters with 22 subclusters (Figure 5-8). This clustering structure informed the outline of this section – three

subsections (Cluster A, B, C), with each subcluster (A1, A2 etc.) discussed individually.

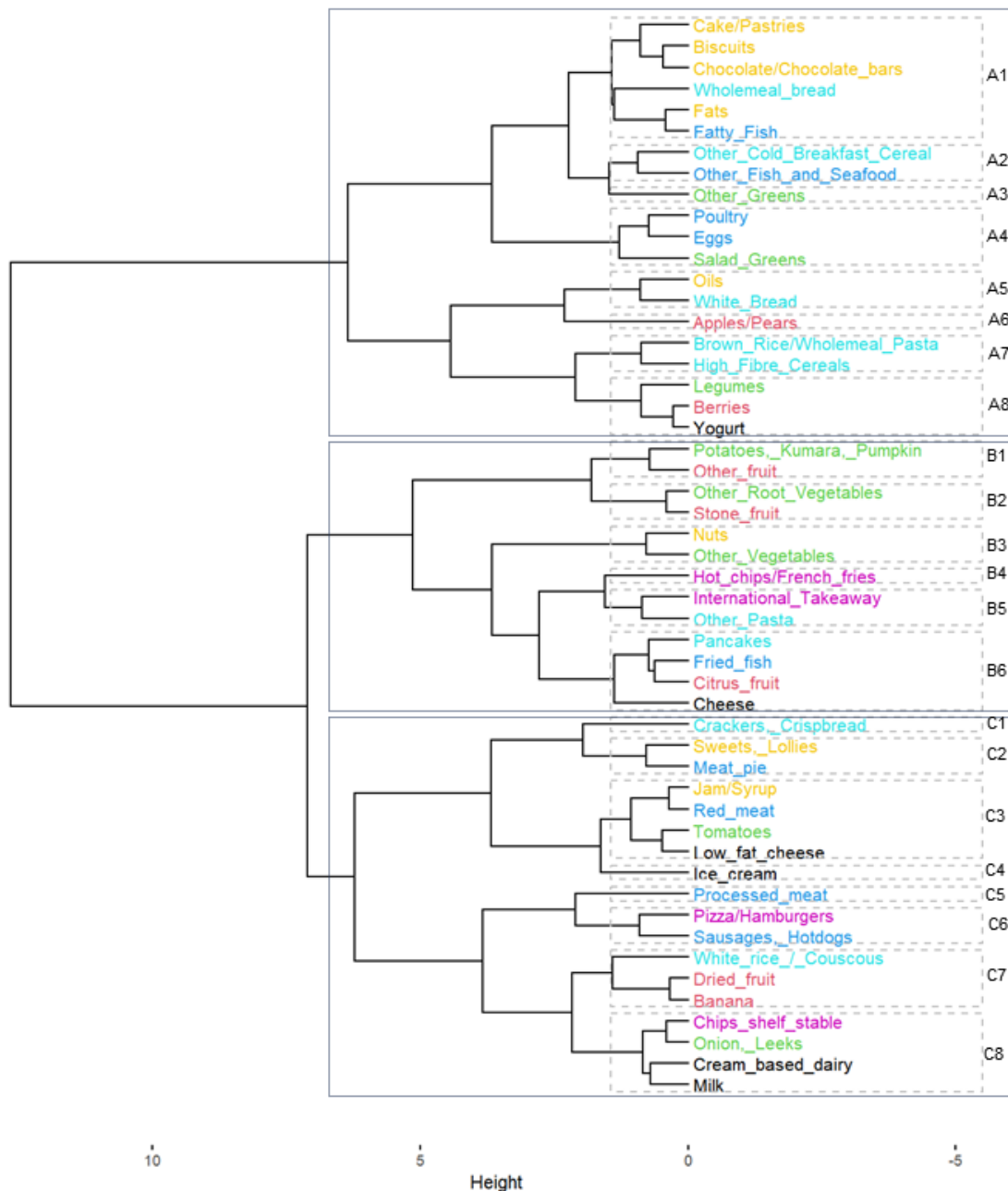


Figure 5-8 Clustering of different food items based on the impact of thermal taste on the consumption. Items are colour-coded according to their categories as described in section 5.2.1.3

The clustering process did not group items according to their predetermined categories (e.g. Cereals), although some patterns were apparent (Figure 5-8). Subcluster A1 contained predominantly items from the “Sweets/baked goods misc.” category, and Subcluster A7 contained only “Cereals” items. Interestingly, two subclusters in Cluster B (B1 and B2) each contained one “Fruit” and one “Vegetables”

item. In addition, none of the “Fast Foods” items formed part of Cluster A.

#### 5.3.3.1 Cluster A

Cluster A contained eight subclusters with two subclusters containing a single item (A3: *Other Greens*; A6: *Apples/Pears*), three containing two items (A2: *Other Breakfast Cereal* and *Other Fish and Seafood*; A5: *Oils* and *White Bread*; A7: *Brown Rice/Wholemeal Pasta* and *High Fibre Cereals*), two containing three items (A4: *Eggs, Poultry* and *Salad Greens*; A8: *Berries, Legumes* and *Yogurt*) and one containing six items (A1: *Berries, Cakes/Pastries, Chocolate/Chocolate bars, Fats, Fatty Fish* and *Wholemeal bread*). Overall, this cluster was negatively influenced by TT cluster membership and, in some cases, by extension, TT Status. Specifically, items in Cluster A tended to have negative regression coefficients with the TT clusters, as described below. This pattern indicates that participants in the relevant TT Cluster were less likely to consume these items. Consumption prediction can be said to be affected by TT status if the effect of all the TT Clusters is in the same direction and of similar magnitude. (Note that that the effects of TT Cluster on consumption *prediction* do not imply a direct causative effect between TT Cluster and consumption.)

Consumption prediction of items in subcluster A1 (*Berries, Cakes/Pastries, Chocolate/Chocolate bars, Fats, Fatty Fish* and *Wholemeal bread*) was negatively affected by TT Cluster 1 and 2 memberships. In contrast, TT Cluster membership affected consumption prediction of items in subcluster A2 (*Other Breakfast Cereal* and *Other Fish and Seafood*) less. Consumption prediction of the item in subcluster A3 (*Other Greens*) was strongly negatively affected by TT Cluster 2 membership and those in subcluster A4 (*Eggs, Poultry* and *Salad Greens*) were negatively affected by TT Cluster 1 membership and positively affected by TT Cluster 2 membership in some cases (Figure 5-9).

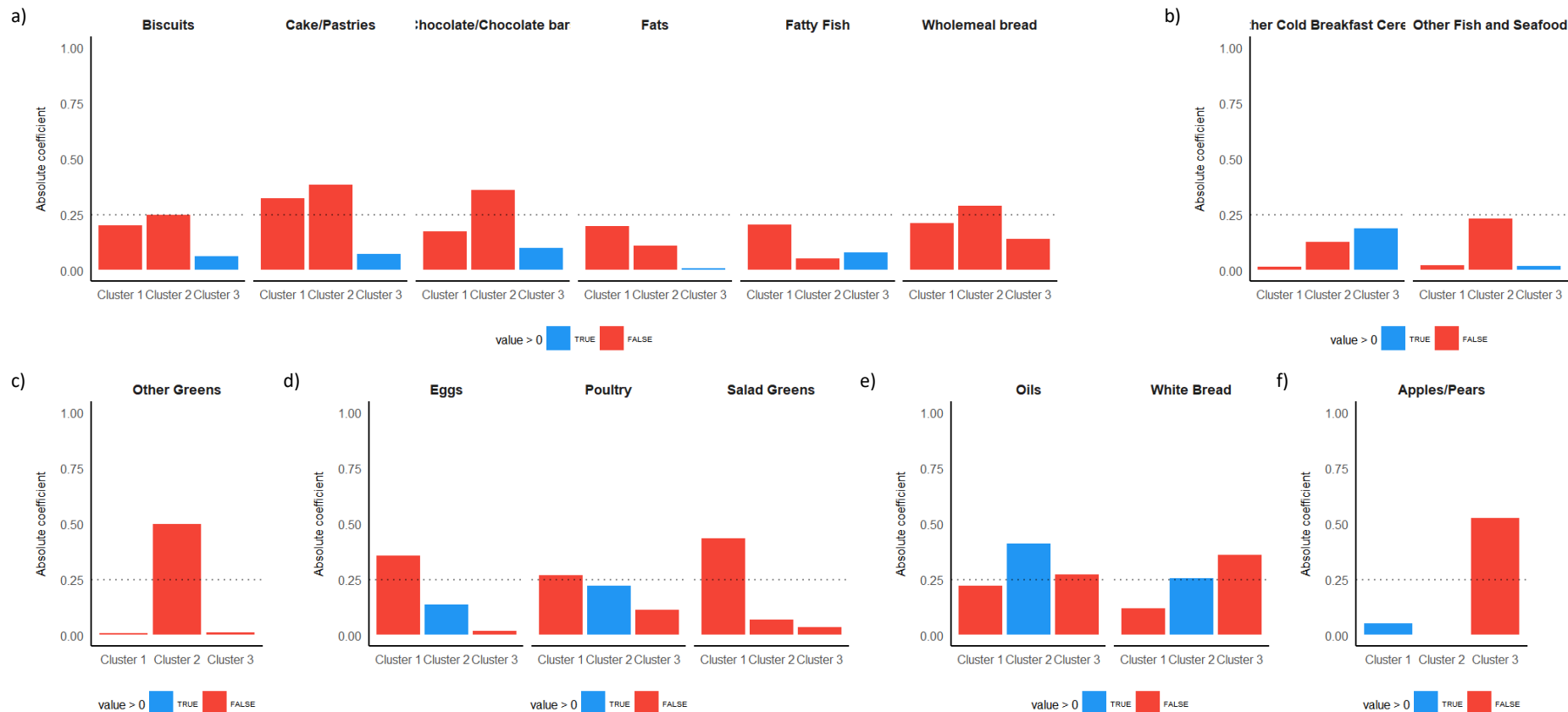


Figure 5-9 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster A1, b) subcluster A2, c) subcluster A3, d) subcluster A4, e) subcluster A5 and f) subcluster A5. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2.

Consumption prediction of item in the remaining four subclusters (A5, A6, A7 and A8) tended to be negatively affected by TT Cluster 3 membership. Consumption prediction of items in subcluster A5 (*Oils and White Bread*) were negatively affected by TT Cluster 1 and 3 membership and positively affected by TT Cluster 2 membership (Figure 5-9). In contrast, consumption prediction of the item in subcluster A6 (*Apples/Pears*) was strongly negatively affected by TT Cluster 3 membership, with this being the only cluster affecting consumption prediction. The prediction of consumption of items in subcluster A7 (*Brown Rice/Wholemeal Pasta and High Fibre Cereals*) was negatively affected by both TT Cluster 2 and 3 memberships. Consumption prediction of items in subcluster A8 (*Berries, Legumes and Yogurt*) was negatively affected by TT Status, as all three TT Clusters affected consumption prediction (Figure 5-10).

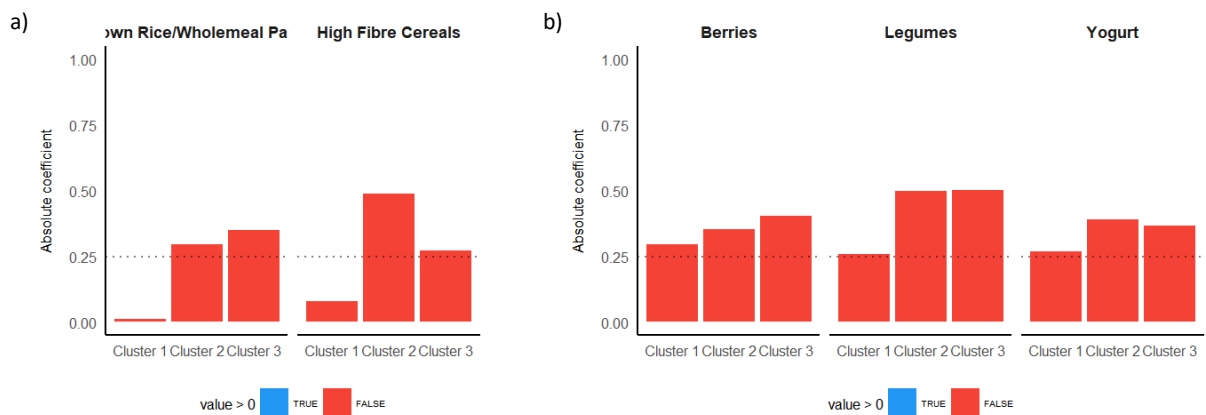


Figure 5-10 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster A7 and b) subcluster A8. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2.

### 5.3.3.2 Cluster B

Cluster B contained six subclusters, four of which contained two items (B1: *Other fruit and Potatoes/Kumara/Pumpkin*; B2 *Other Root Vegetables and Stone Fruit*; B3: *Nuts and Other Vegetables*; B5: *International Takeaways and Other Pasta*), one which contained a single item (B4: *Hot Chips/French Fries*) and one containing four items (B6: *Cheese, Citrus Fruits, Fried Fish and Pancakes*). Overall, consumption prediction

of items in Cluster B was positively affected by Cluster 1 membership, although the effect of other TT Cluster memberships differed between clusters. In other words, participants in Cluster 1 were more likely to consume items from Cluster B.

Consumption prediction of items in subcluster B2 (*Other Root Vegetables* and *Stone Fruit*) was negatively affected by both other TT Cluster memberships, whereas subcluster B3 (*Nuts* and *Other Vegetables*) consumption was positively affected by TT Cluster 3 membership. The consumption prediction of items in subcluster B1 (*Other fruit* and *Potatoes/Kumara/Pumpkin*) was unaffected by TT Cluster C membership (Figure 5-11).

Consumption prediction of items in the remaining subclusters (B4, B5 and B6) was positively affected by all three TT Cluster memberships and can, therefore, be said to have been positively affected by TT status overall. However, the relative effect of TT Clusters differed between subclusters. Overall, this effect was least pronounced for items in subcluster B6 (*Cheese, Citrus Fruits, Fried Fish* and *Pancakes*), with the highest regression coefficient for each item barely exceeding 0.25 (Figure 5-11). In contrast, consumption prediction of the item in subcluster B4 (*Hot Chips/French Fries*) (Figure 5-11) was strongly positively affected by all three TT cluster memberships. Finally, the prediction of consumption of items in subcluster B5 (*International Takeaways* and *Other Pasta*) was more strongly affected by TT Cluster 1 and 2 memberships.

#### 5.3.3.3 Cluster C

Cluster C contained eight subclusters, with three containing a single item (C1: *Crackers/Crispbread*; C4: *Ice cream*; C5: *Processed Meat*), two containing two items (C2: *Meat pie* and *Sweets, Lollies*; C6: *Pizza/Hamburgers* and *Sausages/Hotdogs*)

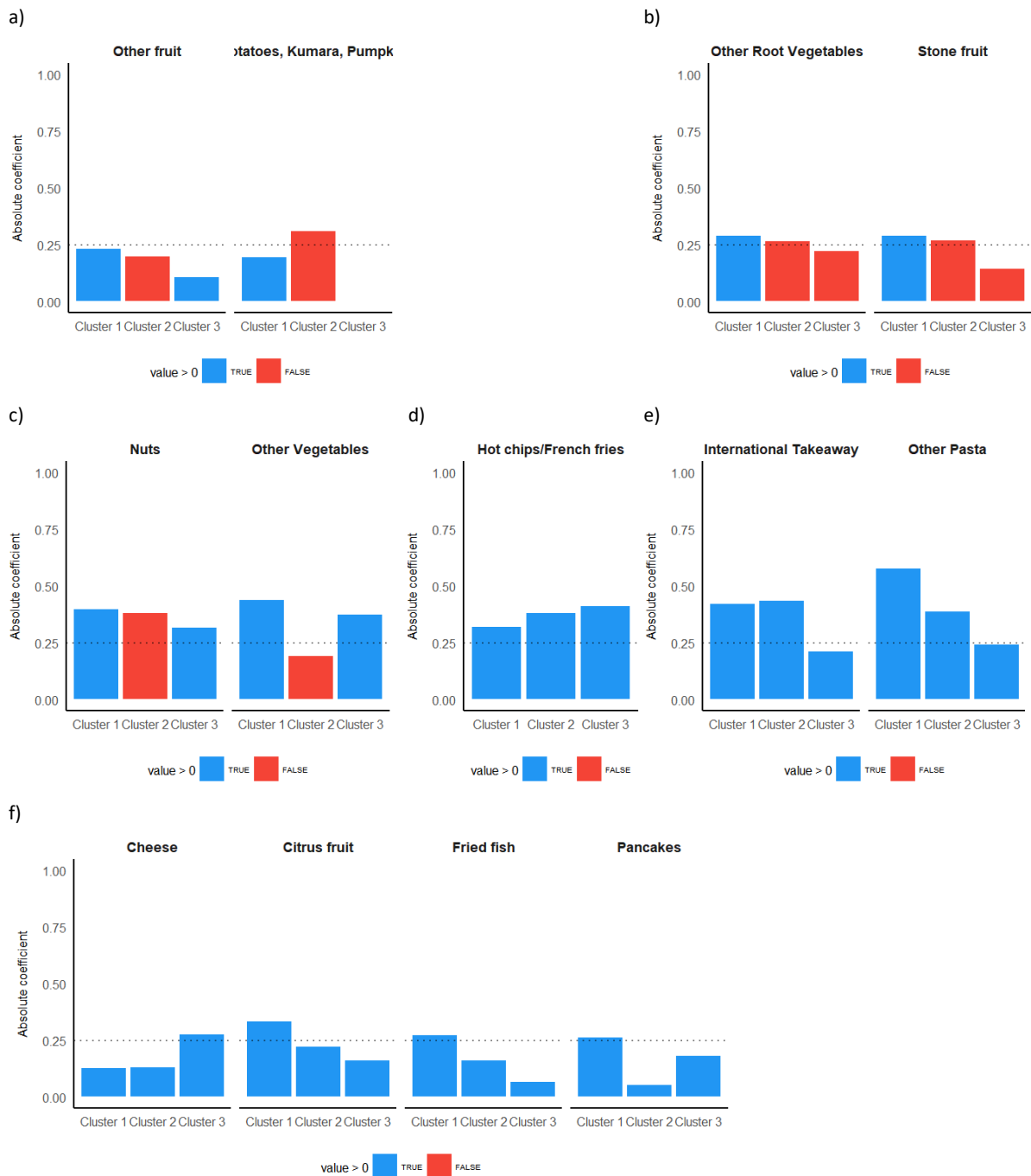


Figure 5-11 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster B1, b) subcluster B2, c) subcluster B3, d) subcluster B4, e) subcluster B5 and f) subcluster B6. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2.

one containing three items (C7: Banana, Dried Fruit, White rice/Couscous) and two containing four items (C3: Jam/syrup, Low fat cheese, Red meat and Tomatoes; C8: Chips shelf stable, Cream based dairy, Milk, and Onions/leeks).

Overall, the prediction pattern for Cluster C was characterised by a less pronounced

effect of TT Cluster 1 membership, although predictive patterns differed for the rest of the TT Clusters. In other words, for the items in Cluster C, there was no clear consumption pattern for participants in TT Cluster 1. Participants in other TT Clusters tended to consume items in Cluster C more or less frequently, depending on the item's subcluster.

The first two subclusters (C1 and C2) were characterised by their consumption prediction being strongly influenced by TT Cluster 3 membership. However, subcluster C1 (*Crackers/Crispbread*) consumption prediction was negatively influenced by TT Cluster 2 membership and subcluster C2 (*Meat pie* and *Sweets, Lollies*) consumption prediction was positively influenced by TT Cluster 2 membership (Figure 5-12).

Consumption prediction for items in subcluster C3 (*Jam/syrup, Low fat cheese, Red meat* and *Tomatoes*) was mildly negatively affected by TT Cluster 1 membership and positively affected by TT Cluster 2 and 3 membership (Figure 5-12). Similarly, consumption prediction of the item in subcluster C4 (*Ice cream*) was positively affected by TT Cluster 3 membership and negatively affected by TT Cluster 1 and 2 memberships, although the effect for TT Cluster 2 was negligible.

Consumption prediction of items in subclusters C5 (*Processed Meat*) and C6 (*Pizza/Hamburgers* and *Sausages/Hotdogs*) was strongly positively affected by TT Cluster 2 membership, with TT Cluster 3 having a weak positive effect on prediction for the item in subcluster C5 and a weak positive effect on the items in subcluster C6 (Figure 5-12). Although TT Cluster 2 membership also had the most considerable impact on consumption prediction of the items in subcluster C7 (*Banana, Dried Fruit, White rice/Couscous*), this effect was less pronounced and accompanied by a positive

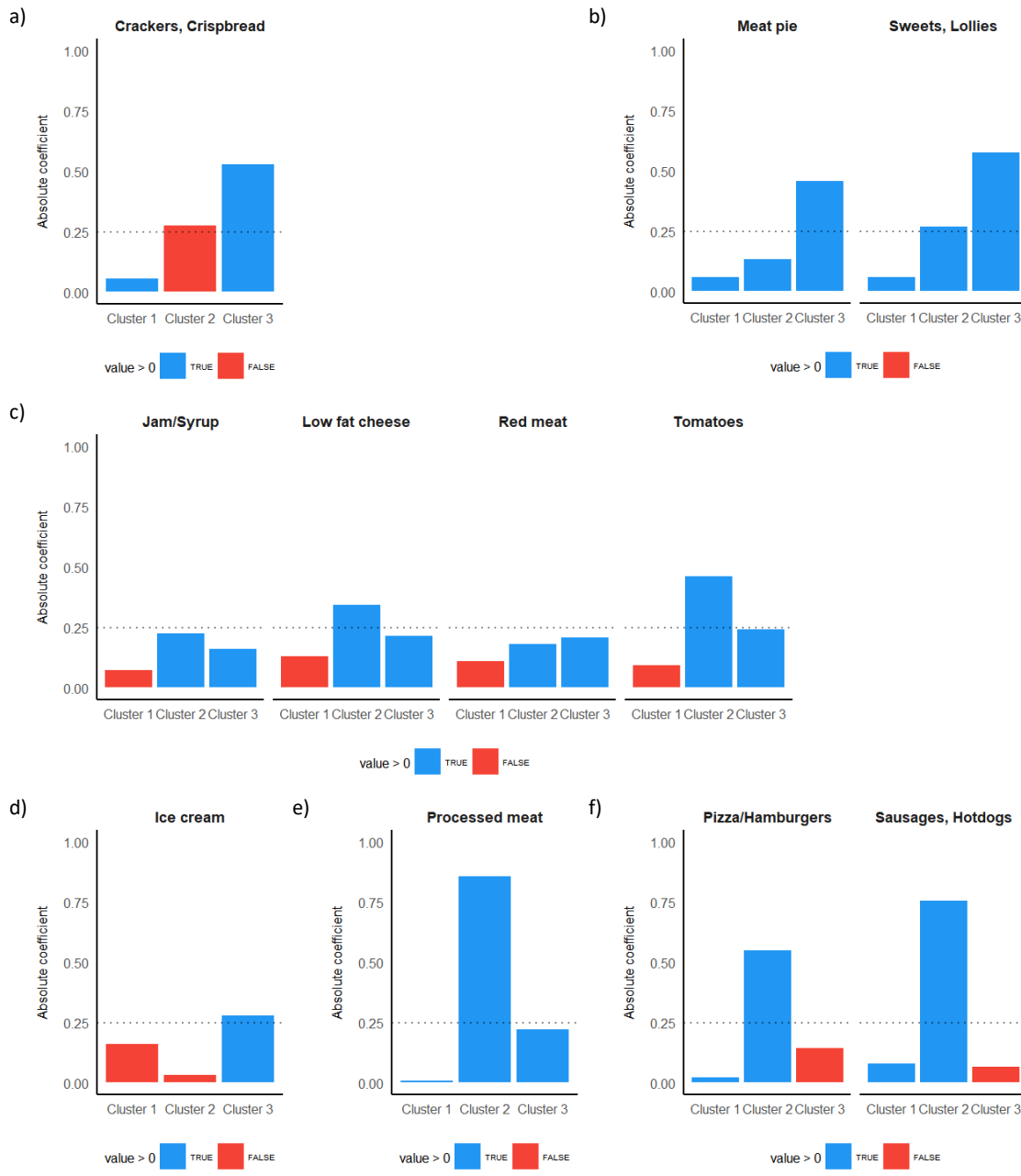


Figure 5-12 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster C1, b) subcluster C2 and c) subcluster C3. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2.

effect of Cluster 1 membership (Figure 5-13). Consumption prediction of the final subcluster (C8: Chips shelf stable, Cream based dairy, Milk, and Onions/leeks) was not strongly affected by any TT cluster membership and can, therefore, be said to be generally unaffected by TT status (Figure 5-13).

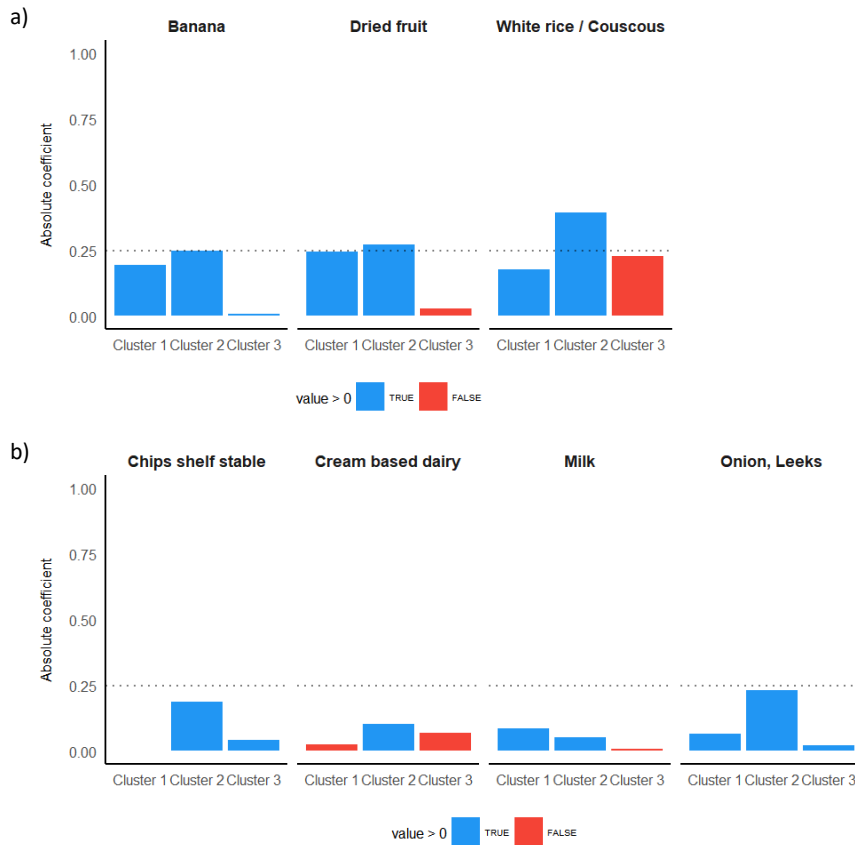


Figure 5-13 Absolute regression coefficients from the ROSA model predicting food choice based on TT Cluster a) subcluster C7 and b) subcluster C8. Blue shaded blocks indicate positive regression coefficients and red shaded blocks indicate negative regression coefficients. The dotted vertical line at 0.25 indicates the threshold for inclusion in results in section 5.3.2.

## 5.4 Discussion

Overall, this study revealed three patterns of influence of aspects of thermal taste, demographics, food behaviour, personality and orosensory responsiveness on food choice. The first pattern was where a specific data block (for example, Food Involvement) had a large or proportionally large effect on the explained variance, suggesting the overall importance of this data block in determining food choice overall. The second pattern was where a specific data block had a notable impact on the explained variance of a specific item (for example, Demographics on *Fats* consumption), highlighting a robust determining effect of that data block on the

consumption behaviour of a specific item. The third pattern was where specific variables greatly influenced the consumption prediction of many items (for example, TT Cluster), suggesting a generalised but latent effect of those specific variables on food choice. Note that thermal taste featured in the third category, with specific TT Clusters featuring in predictions, suggesting a generalised effect of the thermal taste experience on food choice.

#### **5.4.1 Relative predictive effect of selected factors on food choice**

As thermal taste is the focus of this thesis, this section will first focus on the role of thermal taste in the predictive model. The remainder of the data blocks are discussed based on their ranking in terms of selection order, and all data blocks with components selected before thermal taste are included. Further, any additional data blocks with a prominent predictive position are included in the “Other important data blocks” section.

##### *5.4.1.1 Thermal taste*

While fitting the distinct component model, components from only one thermal taste-related data block, TT Cluster, were selected. Three separate and unrelated components were selected in the 20-component model. The TT Cluster data block encodes the most information about thermal taste of the three data blocks. All three blocks encode TT Status, as having reported no tastes (TT Taste) or reporting thermal taste at no temperature (TT Temperature) would encode TnT, but the patterns in the TT Cluster block are more direct. TT Cluster also represents both TT Taste and TT Temperature, as this information was used in the clustering process in Chapter 3. This relationship was apparent from the high absolute loadings of variables from the TT Taste and TT Temperature data blocks on the components extracted from TT Cluster in the distinct component model. Finally, TT Cluster also represents an additional

layer of latent information about thermal taste. Therefore, it is unsurprising that this data block was the thermal taste-related data block that was most informative in predicting food consumption.

Five thermal-taste-related components were included in the common component model, which included components from all three data blocks (TT Cluster, TT Temperature and TT Taste), although none were related to a single TT data block.

The most prominent common data space intersecting with thermal taste, was the intersection of TT Cluster and Demographics, the source of the second and ninth selected components. As the TT Cluster data block also encodes TT Status, the common component likely represents the higher age of TT in the current cohort (Chapter 4). However, if this were the only relevant aspect, the model would have selected a component directly related to Age (originating from the Demographics data block). Overall, this points to latent factors common to thermal taste and Demographics (Age and Sex/Gender), which were not apparent in the distinct component model. This aspect warrants further exploration in future work using common component models.

Similarly, the component extracted from the TT Cluster x Big Five Inventory blocks likely encodes some of the personality differences found between TT Cluster and TT Status groups in Chapter 4 but also describes an interaction between the two factors in determining food choice. However, this interaction was not prevalent in the distinct component model, as none of the thermal taste clusters had high absolute loadings for any of the components extracted from the BFI data blocks.

The remaining two thermal taste-related common components selected included TT temperature as one of the two data blocks in the space from which the component was

extracted. The first component describes the joint space between TT temperature and TT Taste. Although the clustering process described in Chapter 3 attempted to combine temperature and taste aspects of thermal taste, the selection of this component suggests that the clustering process does not fully describe the relationship – at least concerning food choice. (If the relationship was fully described, the algorithm would have selected a component from TT Cluster instead.) The second common component describes the joint space between TT temperature and Sweet Liking, suggesting an interaction between the mechanisms of thermal taste and Sweet liking in determining food choice, rather than the resulting taste.

#### *5.4.1.2 Initially selected data blocks*

Component selection order was related to the overall importance of each data block in the model, as ROSA modelling aimed to provide a holistic explanation of food choice. The factors discussed below were selected based on their selection sequence, the number of times a component from a specific data block was selected, and the number of items for which the specific data block was important.

##### *5.4.1.2.1 Demographics*

The Demographics data block was selected first when fitting the distinct component model. From this data block, “Age” was an important variable overall and was also involved in predicting five items, with *Fats* and *Cheese* consumption being positively correlated with increased “Age”, and *Fatty fish*, *Eggs* and *Banana* being negatively correlated. These results suggest that in the cohort included in the study, older participants tended to eat more *Fats* and *Cheese*, and less *Bananas*, *Fatty Fish* and *Eggs*. Previous research (Cao et al., 2019; García-Esquinas et al., 2019; Granic et al., 2016; Panurywanti et al., 2021) has linked four of these items to health outcomes in older adults.

Consumption of the two items negatively associated with increased age has been linked to positive health outcomes for older adults. Banana consumption at bedtime has resulted in better sleep quality in older adults, linked to the melatonin (neurotransmitter important for sleep) and tryptophan (neurotransmitter precursor) in bananas (Panurywanti et al., 2021). Similarly, regular fatty fish consumption is linked to overall healthy ageing outcomes, including physical and mental health metrics (García-Esquinas et al., 2019).

The two items consumed at increased frequency with increasing age (*Fats and Cheese*) have potential adverse health outcomes in older adults. Both these items contain high levels of saturated fat, intake of which is linked to a higher risk of cognitive decline in older adults (Cao et al., 2019; Granic et al., 2016). These consumption patterns align with recent findings in NZ; recent nutritional surveys have found that NZ older adults are over-indexed on their overall fat intake, leading to a micronutrient-poor diet (Tay et al., 2021), linked to age and ethnicity (Lovell et al., 2023).

These consumption habits may be a result of ageing-related changes in taste (Jilani et al., 2022; Mojet et al., 2001; Mojet et al., 2003; Sato et al., 2022; Yang et al., 2022), chemesthetic (Yang et al., 2022) and olfactory function (Honnens de Lichtenberg Broge et al., 2021). These changes lead to lower hedonic responses to some foods (Honnens de Lichtenberg Broge et al., 2021), and increased fat addition may be an attempt to increase the calorific value of food or to enhance the taste of foods which have lost their olfactory properties, as older adults specifically have impaired ability to smell savoury foods (Honnens de Lichtenberg Broge et al., 2021).

#### 5.4.1.2.2 Orosensory Responsiveness

A component describing the Orosensory Responsiveness data block was selected

second, and a subsequent component was selected in the 14<sup>th</sup> position. A variable related to sweetness intensity was included in the top predictors, whereas variables related to fatty acid responsiveness were useful in predicting the consumption of several items. Notably, PROP taster status was not one of the variables with prediction potential, and was unrelated to either of the extracted predictive components, as evidenced by the variable's low absolute loading on both components.

The curious absence of high regression coefficients with Sweetness intensity may be because of its relationship with Sweet Liker Status (SLS). Sweet liking groups differ in their sweetness intensity ratings (Kavaliauskaite et al., 2023), and it may be that this specific variable is indirectly encoding SLS Cluster membership, which had a notable effect on the prediction of several items.

The lack of predictive potential of PROP sensitivity is notable. Numerous studies have explored the relationship between PROP taster status and food behaviour (Chamoun et al., 2018), but many with inconclusive results (for example Deshaware & Singhal, 2017), implying that, in the larger context of various factors influencing food behaviour, PROP taster status is less important than previously thought.

#### 5.4.1.2.3 Spicy Food Liking

In the distinct component model, Spicy Food Liking was selected third, and two subsequent components were extracted, resulting in two consecutive components (Components 3, and 4). The position and the prevalence of this data block suggest the importance of spicy food liking in determining food choice, which is further evidenced by the fact that two of the top ten variables originate from this data block, including the most important variable. In addition, the consecutive nature of the extracted components indicates the presence of several independent spicy food liking-related

factors in determining food choice. The presence of independent factors is further evidenced by the seemingly counterintuitive relationship seen in the regression coefficients of some items; for specific items (*Processed meat, Jam/Syrup* and *Pizza/Hamburgers*), enjoying the burn of spices was a positive consumption predictor, whereas spice consumption frequency and agreeing that spices make food taste better were negative consumption predictors.

One possible latent factor is cultural groupings, which are more nuanced than those contained in the Ethnicity data block recorded in the overall study. However, it remains curious that many food items that have their prediction significantly influenced by spicy food liking are not foods one would typically associate with spice addition, such as *Cakes/Pastries, Jam/Syrup, Pancakes, and Apples/Pears*. These findings may be related to previous findings that (in a Chinese population) spicy food liking is associated with a higher calorie consumption (Choi & Chan, 2015) and BMI (K. Yang et al., 2018); these findings have not been repeated in a predominantly Western population.

A second possible latent factor is sensation-seeking-related food adventurousness, which was not directly measured. Sensation-seeking has been linked to spicy food liking (Byrnes & Hayes, 2013, 2016; Scott et al., 2019), although there is a gender difference in the overall effect (Byrnes & Hayes, 2015; Spinelli et al., 2018). Sensation-seeking has been shown to influence beverage liking (Higgins et al., 2020) and dietary adherence (Gibson et al., 2016). It is also correlated with food adventurousness (Tornwall et al., 2014), not measured in this study. Therefore, one of the components related to the Spicy Food Liking block may be describing food adventurousness.

#### 5.4.1.2.4 Oral Processing behaviour

The impact of oral processing behaviour on food choice was independent of thermal taste, as no thermal taste related variables had high absolute loadings on the components extracted from the Oral Processing data block. This pattern appears to contrast previous findings (Bajec & Pickering, 2010) relating TT Status to texture preference. However, this apparent contradiction can be explained by the fact that oral processing subgroups do not translate directly to texture preference (Kim & Vickers, 2020), and therefore does not rule out texture liking differences between TT groups.

Predictive components related to oral processing behaviour were selected twice in the first ten components (fifth and ninth), and three variables from the Oral Processing block (“Smoosher”, “Chewer”, and “Sucker”) were included in the top ten variables. Notably, “Smoosher” status was the second most important variable. However, membership to one of the oral processing behavioural groups was only a prominent consumption predictor for a third of items. Together, these results indicate that oral processing behaviour impacts food behaviour overall without impacting the consumption of a specific item or group of items.

Although Jeltema et al. (2016) demonstrated some food choice differences between oral processing groups, the potential for manipulation of foodstuffs to suit the desired behaviour is a key driver rather than the texture itself (Jeltema et al., 2015; Jeltema et al., 2020; Jeltema et al., 2014). As the same foods can suit different oral processing patterns (for example, hard candies can be chewed or sucked (Jeltema et al., 2016)), or their texture can be manipulated via preparation methods (e.g. boiled /mashed potatoes or toast/ fresh bread) differences in oral processing behaviours likely result in complex differences in food consumption patterns, rather than differences in the consumption frequency of specific items, consistent with the findings of the present

study.

#### 5.4.1.3 Other data blocks of importance

##### 5.4.1.3.1 Sweet liking

Although the Sweet Liking data block was only selected once in the first ten components (Component 10), it was selected three consecutive times when fitting the distinct component model (Components 10, 11 and 12). None of the variables from Sweet Liking was contained in the top ten predictors, but variables from Sweet Liking did impact the prediction of the majority (39) of specific items, evidenced by the large absolute residuals. Although the absence of Sweet Liking in the top (VIP) predictors suggests that the effect is more specific than general, the large number of items affected gives further evidence to the hypothesis put forward in section 5.4.1.2.2 that the overall effect of the intensity rating of the 12% sucrose solution is a proxy for sweet liking.

High Sweet Liking had the most pronounced effect on food consumption prediction, as most items were significantly influenced by HSL, and the highest regression coefficients occurred. Notably, many of the items influenced by HSL were not sweet. For example, *Sweets*, *Lollies* and *Cakes/Pastries* consumption were not predicted by sweet-liking groups, in contrast with previous findings (Holt et al., 2000; Lim et al., 2020; Pangborn & Giovanni, 1984). In addition, HSL was a negative predictor of *Chocolate/Chocolate bars* consumption, and a positive predictor of the consumption of several fruit items. This counterintuitive relationship is likely an example of intuitive eating stemming from better interoception (Iatridi et al., 2021), as HSL may be choosing fruit as a healthier sweet substitute to more typical sweet treats due to unpleasant post-ingestive consequences of sugar consumption resulting from blood

sugar fluctuations (peak and subsequent crash). In addition, HSL may choose the healthier sweet substitute due to their better resistance to emotional eating behaviours (Iatridi et al., 2021).

#### **5.4.2 Impact of thermal taster status (Cluster) on food choice**

Generally, the effect of TT Cluster on food choice could be grouped into three categories: items for which TT Cluster membership (and by extension, TT Status) negatively affected consumption, items for which TT Cluster 1 membership had a positive effect on consumption, and items for which either Cluster 2 or Cluster 3 membership had a positive effect on consumption. Notably, the subgroup for which TT Status had the clearest effect (negative regression coefficients of similar magnitudes) contained three items which could be considered “mushy” (*Berries*, *Legumes* and *Yogurt*), in line with previous findings (Bajec & Pickering, 2010) that TT tend to dislike mushy foods. In contrast, the item with the clearest positive effect of TT Status (*Hot Chips/French fries*) is an item which is specifically consumed hot, with a room-temperature alternative available (*Shelf-stable chips*). Although this could be an effect of temperature driving consumption, it is also possible that it is a latent effect of texture (mushy vs crunchy). However, in the latter case, it would contradict other findings in the present study and the literature (Bajec & Pickering, 2010).

Overall, TT Cluster membership exhibited a complex set of patterns across food consumption behaviour, indicating that the combination of the time at which thermal taste is perceived and its underlying mechanisms are drivers of food choice. Therefore, future studies on TT and food choice should consider TT as subgroups rather than a single group.

### **5.4.3 Model fitting**

The cross-validation showed that the model was overfitted in most instances, with the patterns of overfit differing between items. Overall, this indicates that the model differs substantially between different items, suggesting that the model was able to optimise its fit for some food items (responses) but not all. As the present study was the first to use ROSA with multiple response variables, it is possible that strategies not yet explored in the literature are required for successful multi-response ROSA fitting.

One such strategy is the strategy of dual cross-validation, as suggested by Smilde et al. (2022). Typically, cross-validation in ROSA happens in two steps. First, the component choice is cross-validated to check the validity of the component choice in different samples. Next, the prediction is cross-validated to check the overall predictive ability in different scenarios. Note that the results of prediction cross-validation do not influence any of the component choices. In dual cross-validation, the two cross-validation steps happen simultaneously, and results of the predictive cross-validation influences the component choice.

A second possible strategy is using a more complex approach to common components. The present study only allowed for the extraction of components in the spaces of two data blocks and was cautious in its interpretation of the common component model. However, it is possible that the use of common component model that allows for the extraction of components from any higher-order blocks could increase the overall fit of the model without overfitting.

### **5.5 Limitations**

A limitation of this study was that participants used a self-report food questionnaire validated mainly for determining nutritional status. Food diaries are an alternative but

introduce the risk of participant dropout. In addition, participants were not screened to exclude those following limited diets, for medical or other reasons.

A further limitation is that data was collected during 2020, at the start of the COVID-19 pandemic. The first round of data was collected at the end of a reasonably short lockdown. However, as it was expected people might have changed their eating behaviour, they were asked to comment on this and say how different it was. It was not found to be an issue. By the time the second round of data was collected, there had been another lockdown, although this was a lockdown with relatively low restrictions. Although numerous studies indicate that eating behaviours changed during the pandemic, it is important to note that there were relatively few lockdowns/restrictions in New Zealand in the early part of the pandemic.

## **5.6 Conclusion**

This chapter used a ROSA modelling approach to attempt to untangle the complexity of food choice overall and contextualise thermal taste with other factors potentially affecting food choice. Here, TT Cluster was the most important thermal taste-related data block in predicting food choice, likely because it encodes all aspects of thermal taste (temperature, taste and status). In addition, TT Cluster was found to be linked to the specific consumption patterns of most items. When interrogating patterns of TT Cluster on the consumption of specific items, items from the same food groups tended not to cluster together. Some patterns around “mushy” textures were found, in agreement with previous research.

Overall, this chapter established that although thermal taste is related to differences in food choice, it does so in a manner that interacts with other aspects related to food choice.

## **Chapter 6 Discussion and future work**

### **6.1 Introduction**

From the review of literature surrounding thermal taste, including existing TT phenotyping methods, orosensory differences between TT and TnT, and studies investigating food choice and TnT several key themes emerged:

- Studies differ in their definition of thermal taste
- Thermal taste phenotyping is generally inefficient
- Traditionally phenotyped TT are more orosensitive than TnT
- Research into thermal taste and food choice is limited to small sample sizes
- There is some evidence of behavioural differences between TT and TnT
- Mechanistic information about thermal taste is limited
- Some additional measures of food/orosensory behaviour might have a mechanistic link with thermal taste

Several research gaps transpired from these themes: a need for a better phenotyping method, which was both more effective and efficient; a lack of information about thermal taste as a phenomenon, partially because historically, data was only collected at a single time point, after thermal exposure; and a lack of understanding about how different factors potentially related to thermal taste interact.

These gaps led to the central idea of this thesis, which was that an improvement in TT phenotyping, which is descriptive rather than primarily classificatory, would vastly expand the possibility of what can be known about thermal taste. It led to the four main objectives of developing an improved phenotyping method for thermal taste; grouping and describing thermal taste responses; determining relationships between

thermal taste and demographic, food behaviour, personality, and orosensitivity; and establishing the relative importance of thermal taste in determining food choice.

A secondary idea explored throughout the thesis was that advanced data collection tools require advanced modelling. Theoretical themes around modelling were explored extensively in the relevant chapter introductions and will not be revisited in the summary. However, relevant sections will address modelling challenges, limitations and future work.

## **6.2 Summary**

### **6.2.1 Chapter 2: Fast and Furious**

Chapter 2, “Fast and furious”, aimed to develop a novel phenotyping method for thermal taste. This method followed the proposal of Skinner et al. (2018) to combine the two temperature trials traditionally used for phenotyping (Bajec & Pickering, 2008). In addition, it incorporated the suggestion of Skinner (2017) to include TCATA collection concurrently with temperature exposure. Existing sensory data collection software was unsuitable for this purpose, as it was impossible to align temperature exposure with data collection. A custom system was built with the help of two engineers, and the overall method (system and data collection) was named RapCoTT.

The second aim of Chapter 2 was to validate the RapCoTT method in terms of effectiveness and efficiency (furious). This aim was achieved through a method validation study in which 132 participants completed both protocols. In addition, two different classification schemes were used with the traditional protocol. Overall, the RapCoTT method was more effective than the traditional method, as it classified 93% of participants, compared to between 68% (Scheme A) and 80% (Scheme B) by the traditional approach. In addition, RapCoTT was found to agree with previous schemes

and showed no directional bias compared to the other schemes. Finally, RapCoTT achieved TT phenotyping significantly faster than traditional approaches, roughly halving the time taken to phenotype participants.

Overall, the RapCoTT method enabled the effective phenotyping of participants and allowed for better data to be collected, enabling the establishment of clusters and the rest of the PhD.

### **6.2.2 Chapter 3: Divide and conquer**

Chapter 3, “Divide & Conquer”, aimed to cluster TT based on their thermal taste responses and to use the clustered responses to extract patterns of thermal taste response. It was expected that clusters of TT would exist and that these clusters would be similar in the time (temperature) at which thermal taste was reported (Skinner et al., 2018) and not only the nature of thermal taste reported. In addition, clear relationships between patterns of thermal taste response were expected, linking different thermal tastes to specific mechanisms, for example, confirming previous work linking TRPM5 to sweet thermal taste (Skinner et al., 2018 ; Talavera et al., 2005). Specific temperature-taste relationships were also expected to be stronger than temperature ramp-taste relationships.

Clustering was addressed by using an exploratory clustering approach, followed by a mixture hidden Markov model (MHMM) of state sequences. The modelling process split responses into three groups with either four (Cluster 1 and 2) or five (Cluster 3) hidden states. All clusters of responses contained a state where Metallic was most frequently reported. However, specific combinations of tastes also characterised response clusters. Responses in Cluster 1 tended to contain reports of Metallic and Salty and Metallic and Salty thermal tastes reported concurrently. Responses in

Cluster 2 tended to contain reports of Savoury/Umami thermal taste and Sour thermal taste. Finally, responses in Cluster 3 tended to contain reports of Sweet and Bitter thermal taste and concurrent reports of Bitter and Metallic thermal taste. The presence of three clusters was consistent with Skinner et al. (2018), who found four clusters of distinct thermal taste responses, although the clusters described in Skinner et al. (2018) were characterised by the time at which they experienced thermal taste most intensely, and not the nature of thermal taste reported.

When establishing specific participant clusters, 64% of thermal tasters were consistently assigned to the same cluster. The inconsistently classified participants aligned with previous findings as Skinner et al. (2018) found a similarly-sized group of participants with inconsistent responses for thermal taste. Inconsistently classified participants tended to have a higher affinity to one of the two clusters they were assigned, resulting in their final cluster assignment.

Three sets of data analyses were used to establish links between temperatures, temperature ramps and reported thermal tastes. Overall, relationships between temperature ramps and tastes failed to replicate previous findings (Skinner et al., 2018; Thibodeau et al., 2020), suggesting a methodological influence in previous studies.

Functional data analysis (fda) and an adapted approach to traditional TCATA data analysis aided the establishment of links between specific temperatures and tastes. Here, fda was more useful in establishing these links. Existing hypothesised links were confirmed, but novel links were also established. The fda analysis confirmed previous findings (Skinner et al., 2018 ; Talavera et al., 2005), linking the TRPM5 channel to Sweet thermal taste. However, evidence for additional mechanisms for Sweet thermal

taste was also found, including TRPA1 and TRPM5 during cooling, potentially linked to TRPM4. Umami/Savoury thermal taste was likely produced through similar mechanisms (TRPA1 and TRPM5) than Sweet thermal taste. These findings aligned with literature linking TRPA1 with Sweet and Savoury/Umami taste (Matsumoto et al., 2023), TRPM5 with Umami/Savoury (Zhang et al., 2003 ) and TRPM4 with Sweet (Banik et al., 2018; Dutta Banik & Medler, 2023).

There was also evidence of the production of Bitter and Sour thermal taste through a TRPA1-dependent mechanism, but likely through perceptual association rather than true taste mechanisms. TRPA1 is associated with the nociceptive response to CO<sub>2</sub> (Wang et al., 2010) and is a sensor for weak acids (de la Roche et al., 2013; Wang et al., 2011), and participants likely misreported the sensation of TRPA1 activation as Sour thermal taste. Similarly, although there is no direct link between Bitter taste and TRPA1, the channel responds to astringent polyphenols (Kurogi et al., 2015; Takahashi et al., 2021) and caffeine (Nagatomo & Kubol, 2008). Therefore, participants may have misreported bitter-associated sensations (astringency and orosensation resulting from caffeine) as Bitter thermal taste. Evidence for an additional thermal taste mechanism for Sour thermal taste was found for TRPV4 (Matsumoto & Kato, 2019; Matsumoto et al., 2019).

Metallic thermal taste exhibited complex thermal patterns. However, there was evidence that Metallic thermal taste was produced directly through thermal activation of channels previously linked to Metallic taste, namely TRPM5 (Riera et al., 2009), TRPA1 (Ecarma & Nolden, 2021) and TRPV1 (Riera et al., 2008, 2009). These clear links between Metallic thermal taste and specific mechanisms contradict previous criticisms (Nachtigal & Green, 2020; Yang et al., 2014) of Metallic thermal taste as a spurious sensation resulting from a mental association with the metal thermode.

Salty thermal taste also exhibited complex patterns, but links were established to three different mechanisms. First, evidence was found to confirm the involvement of the ENaC salt taste channel (Askwith et al., 2001), previously linked (Talavera et al., 2007) to Salty thermal taste. In addition, a second peak for Salty thermal taste could be connected to the TRPV1 and TMC4 taste channels, both previously associated with Salt taste (Kasahara, Narukawa, Ishimaru, et al., 2021; Kasahara, Narukawa, Kanda, et al., 2021; Kasahara et al., 2022; Kim et al., 2014; Lyall et al., 2004).

Overall, the work in this chapter laid the groundwork for the remaining chapters, where the TT clusters were further investigated. However, the links discovered between specific thermal tastes and mechanisms demonstrate the power of the RapCoTT method and constitute significant novel findings.

### **6.2.3 Chapter 4: Seeing the forest and the trees**

Chapter 4, “Seeing the forest and the trees” aimed to establish the relative relationships between TT Status and TT Subgroup (Cluster) to demographic, food behavioural, personality, and orosensory responsiveness measures. This overall aim resulted in two separate objectives: to profile a group of participants on the selected measures and use machine learning to establish each factor’s relative effect or importance and its constituents in determining TT status and TT cluster.

Although specific hypotheses informed the inclusion of the different factors or metrics, two main hypotheses informed the analysis in this chapter. The first was that differences exist between TT subgroups, skewing previous findings (null or otherwise) between TT and TnT. In particular, contradictory findings in food neophobia between Yang (2015) and Bajec and Pickering (2010) suggest the presence of different thermal taster subgroups. The second was that individual measures may be more effective than

global measures in predicting TT status or TT cluster, mirroring the findings of Yang (2015).

Although this work was exploratory, several factors not previously explored alongside TT were included based on the expectation that there would be differences between TT groups. Oral Processing behaviour (Jeltema et al., 2015) spicy food liking (Lawless et al., 1985) and fatty acid sensitivity (Liu et al., 2011; Motter & Ahern, 2012) were included as these were expected to have common mechanistic origins to thermal taste, linked to temperature-sensitive taste channels. Food modification was included due to evidence that participants with higher orosensitivity differ in their food modification behaviour (Puputti, Hoppu, et al., 2019). In addition, private body consciousness (Miller et al., 1981) was included as participants with higher interoceptive awareness (higher PBC) were expected to be more likely to report thermal taste.

Factors previously studied in TT were included to confirm previous differences, confirm previous null findings, or explore aspects where previous findings were contradictory. Food Involvement and Big Five inventory were included as preliminary findings (Yang, 2015) indicate that TT differ from TnT in these aspects. In addition, one preliminary finding (Thibodeau et al., 2019) points to differences in age between TT and TnT. It was also expected that null findings regarding PROP taster status (Bajec & Pickering, 2008; Bering et al., 2013; Thibodeau et al., 2019; Yang et al., 2014; Yang et al., 2020), gender/sex (Bajec & Pickering, 2008; Pickering & Kvas, 2016; Pickering et al., 2016; Thibodeau et al., 2019; Yang, 2015; Yang et al., 2020), Sweet Liker Status (Yang et al., 2020) would be repeated. In addition, the modelling approach and inclusion of TT Cluster groups were expected to shed light on previous contradictory findings regarding Food Neophobia (Bajec & Pickering, 2010; Yang, 2015) and ethnicity (Thibodeau et al., 2019; Yang et al., 2020).

Two data analysis approaches were used. Traditional statistics (Pearson's chi-squared test, Kruskal-Wallis test) to compare relationships between phenotype categories and numeric scores of aggregated measures; and a machine learning approach to determine the relative predictive effect of any raw or aggregated data collected. Here, an ensemble of random forest models was used to eliminate the worst performing variables to predict TT Status or TT Cluster. The remaining variables were assessed using interpretable machine-learning techniques.

When considering only traditional statistics, there were very few differences between TT and TnT. TT tended to be older than TnT. The only other significant difference between TT and TnT was in aspects of the five-factor model of personality; TT tended to be more extraverted and more open new experiences than TnT, the latter finding aligning with Yang (2015).

Traditional statistics revealed more differences between TT subgroups; here, there were differences in overall food involvement (including the Set & Disposal subscale) and food neophobia (including the Avoid dimension). In addition, differences existed in four of the five factors of the Big Five Inventory (Agreeableness, Conscientiousness, Extraversion, Openness) and in Sweet Liking. These findings support the idea that in previous studies, discrepancies may have resulted from the presence of different subgroups. In particular, differences in Neophobia between TT subgroups likely explain the discrepancies in findings regarding Neophobia between Yang (2015) and Bajec and Pickering (2010).

Using interpretable machine learning after an ensemble random forest approach resulted in more nuanced insights. Here, subtle differences were found between TT and TnT in many factors, including oral processing behaviour, food neophobia,

specific items relating to food involvement and private body consciousness. In addition, differences in food modification behaviour and specific items related to the enjoyment of spices were predictive of TT status. Finally, differences in liking of some sucrose concentrations and sensitivity to PROP and two levels of oleic acid were predictors of TT Status. Extraversion was also predictive.

Interpretable machine learning revealed different predictive patterns for TT Clusters. Here, the Set & Disposal subscale of Food Involvement was the best predictor of TT subgroup. Similarly to TT Status, aspects of personality and Food Neophobia were also predictive of TT Cluster. However, a general trend was that individual items from subscales were more predictive than the aggregates.

Overall, the results from this chapter paint a picture that thermal taste has a behavioural aspect, as behaviour was more different between TT and TnT. In addition, there were also behavioural differences between TT subgroups. Differences in thermal taste experiences may result in different behaviours, such as differences in food neophobia or modification. However, it is also possible that some behaviours (such as spicy food consumption) indirectly result in different thermal taste experiences. It is also possible that some behavioural aspects, like personality and thermal taste share a common, but independent underlying cause. These findings were unexpected, as it was expected that previous personality differences between TT and TnT (Yang, 2015) were a result of the specific phenotyping method used. This behavioural aspect, potentially underlying individual reporting or perception mechanisms, warrants further investigation.

#### **6.2.4 Chapter 5: Path to enlightenment**

Chapter 5 explored the relative influence of thermal taste on food choice. Here, the

overall hypothesis was that thermal taste would interact with different factors in a complex manner in determining food choice. In addition, it was expected that the details of thermal taste (temperature at which thermal taste is reported, thermal taste cluster or the identity of thermal taste) would be more influential than thermal taste itself. A second aim was to explore the different aspects of thermal taste on food choice by clustering food items based on the effect of the most important thermal taste predictor set. Here, it was expected that foods would not group according to typical categories (i.e. sweets, starches etc.), as the effect of thermal taste on food choice is likely to be complex. However, some item groupings could have been based on texture, in accordance with previous findings connecting differences in food choice between TT and TnT with texture (Bajec & Pickering, 2010).

Data obtained in Chapters 2, 3 and 4 was combined with food frequency data, collected using a shortened Food Frequency Questionnaire (Sam et al., 2020). Food frequency was modelled as an outcome, using ROSA.

Overall, the ROSA algorithm identified TT Cluster to be the most important aspect of thermal taste in predicting food choice. This result was unsurprising as this data block encoded the time at which taste was reported, and the specific thermal taste reported. In addition, of all possible data blocks, TT Cluster notably affected consumption prediction for the most individual food items. However, when considering their predictive impact on the set of food items as a whole, Demographics, Orosensory sensitivity and Spicy food liking were more important predictors than thermal taste overall. However, specific temperatures at which thermal taste was reported were related to the components (large loadings for this specific component) extracted from the spicy food liking data block, indicating that thermal taste interacts directly with spice consumption behaviour in determining food choice overall.

When clustering food items based on the effect of TT Cluster, groups did not align with food item categories. However, at least some of the groups whose consumption was most affected by TT Status overall shared textural similarities, as consumption of the group most clearly affected by TT Status were all “mushy” in nature, agreeing with previous findings (Bajec & Pickering, 2010). However, the effect of oral processing subtype on consumption prediction was unrelated to thermal taste, as variables from thermal taste related data blocks tended to not have high absolute loadings (strong relationships) to components extracted from the Oral Processing data block. This independence does not mean that the effect of thermal taste on food consumption is independent of texture, as oral processing behaviour is not directly related to texture liking (Kim & Vickers, 2020).

Although relationships could be extracted from the ROSA model, it was overfitted, as it was not equally effective at explaining consumption prediction of all items. This lack of fit points to the need for further development of the ROSA algorithm and validation of its optimal use in practice, as further explored in section 6.3.4.

## **6.3 Strengths and Limitations**

### **6.3.1 RapCoTT method**

Although the development of the RapCoTT method was a key objective of this PhD, and its merits have been previously discussed (Chapter 2 and this chapter), it is necessary to highlight specific strengths and limitations.

The efficiency and effectiveness of RapCoTT were demonstrated in Chapter 2, and the power of the insights that can be derived from RapCoTT data was demonstrated in Chapter 3. A significant strength of this method is that it does not rely on memory, or working memory and does not allow participants to reflect on their experiences and

potentially deceive the experimenter. RapCoTT has the potential to facilitate a paradigm shift about thermal taste – from a binary phenotype including a small portion of the population to a range of phenomena experienced, in some degree or another, by most people. This paradigm shift will also bring thermal taste out of relative obscurity and into its rightful focus with other orosensory experiences. The researcher hopes RapCoTT will become the gold standard for thermal taste phenotyping.

However, RapCoTT, as presented, has some limitations. The current setup only allows for reporting prototypical tastes, with an option for “other”. As thermal taste research has revealed that the scales (or options) presented to participants influence reported sensations (Thibodeau et al., 2020), it is likely that participants did not report all orosensations during RapCoTT and that some orosensations were “dumped” (Clark & Lawless, 1994) into similar categories. The mechanistic links with Sour and Bitter thermal taste highlight this phenomenon, as it is likely that taste-like orosensations like astringency and nociception are produced through thermal exposure. Therefore, it is recommended that additional orosensations be considered for inclusion in future versions of the RapCoTT method.

A second limitation is the use of English words, as read by participants, as the only means of data collection. Although no published data exists on the range or accuracy of different people interacting with a TCATA paradigm, individuals are expected to differ in their reading and sensory-verbal processing speed. Some factors to consider are neurodiversity (Leth-Steensen et al., 2000) and multilingualism (Hui & Godfroid, 2021). Although linguistic data was not collected, 16.8% of the local population of Palmerston North is multilingual (Stats NZ, 2018), indicating that a proportion of participants were responding to RapCoTT through a second language paradigm.

Responding to a TCATA paradigm in a secondary language is expected to be more challenging than responding in a first language. Therefore, further improvements to the RapCoTT paradigm should be investigated, potentially including manual translation or an option for participants to assign their own words or emoji (Reis Rocha et al., 2024) to orosensations during training.

### **6.3.2 Data collection**

This PhD presented the most comprehensive study to date on thermal taste. It was the first study to include oral processing behaviour, spicy food liking, food modification, fatty acid sensitivity and private body consciousness. It was the largest study to explore the connection between five-factor personality and food involvement and thermal taste, as the previous study exploring this (Yang, 2015) only included a small number of participants (50). It was also the first study to combine as many factors as other large studies (Thibodeau et al., 2020; Thibodeau et al., 2019; Yang et al., 2020) only explored a limited combination of factors. In addition, it was also the largest cohort to date in which thermal taste and food choice have been investigated.

Although the data collection aimed to be comprehensive regarding potential factors that can predict thermal taste (Chapter 4) or food choice (Chapter 5), some factors that would have helped explain discrepancies were omitted. The researcher chose to omit phenotyping for irritant bitter tasting (Green & Hayes, 2003) due to the prescribed method's complexity and the risks of handling a neurotoxin (capsaicin). In addition, risks associated with close personal contact during the emergency phase of the COVID-19 pandemic (WHO, 2023) were considered. Participants were also not assessed for sensation-seeking, previously linked to differences in spicy food behaviour (Byrnes & Hayes, 2013, 2015, 2016). Such information may have clarified the effects of spicy food liking on food choice.

In addition, several novel questionnaires assessing food behaviour were not included in data collection, as they were published after the study was designed or after data collection. These questionnaires include scales measuring the need for uniqueness (Cardello et al., 2019), food literacy (Park et al., 2020) the motivation to eat new foods (Nezlek et al., 2021), foodie scales (Pickering & Pickering, 2022; Setia et al., 2022), and an alternative scale to measure the liking of new foods (De Kock et al., 2022).

### **6.3.3 Timing**

One factor that was entirely outside the researcher's control was the research's historical timing, which overlapped with the start of the emergency phase of the COVID-19 pandemic (WHO, 2023). However, there were also some benefits to this timing.

In the period this PhD was ongoing, there were significant developments in fundamental taste research, enabling many of the mechanistic connections made in Chapter 3. Some examples are the role of TRPA1 in Sweet and Umami sensation (Matsumoto et al., 2023), the role of OTOP1 in Sour (Zhang et al., 2019) and ammonium chloride (Liang et al., 2023) taste, the role of TMC4 in Salt taste (Kasahara, Narukawa, Kanda, et al., 2021; Kasahara et al., 2022) and the role of TRPM4 in Sweet and other sensations (Dutta Banik & Medler, 2023). The COVID-19 pandemic may have partially influenced these developments; COVID-19 affects chemosensory sensitivity, including taste and chemesthesis (Parma et al., 2020), bringing fundamental taste research into global focus at the start of the pandemic.

However, the COVID-19 pandemic also negatively affected the research. One aforementioned limitation of the research due to the pandemic was the impact of the pandemic-related lockdowns on eating habits (Marty et al., 2021; Poelman et al., 2021;

Robinson et al., 2021). Although the lockdowns experienced in New Zealand during 2020 (when the food frequency data was collected) were relatively short and of low intensity, there is some evidence of dietary changes in New Zealand during that time (Gerritsen et al., 2021), which may have influenced the overall conclusions of this portion of the study.

A second challenge presented by the pandemic was the limitation of participants. Although this study is the largest study to collect information on such a complex set of variables surrounding thermal taste, a larger sample would have been preferred. However, recruiting participants for in-person studies, particularly studies involving extending the tongue, was challenging at the end of 2020.

#### **6.3.4 Modelling**

This research used several advanced data analysis approaches, which allowed for extracting insights about thermal taste, which would not have been possible without these approaches. In addition, it presents the first application in sensory science of many approaches (state sequences in the context of mixture hidden Markov models, interpretable machine learning, ROSA predictive models) and one of the first in some others (Markov modelling, fda).

Although not the first study to use Markov modelling for temporal data (Lecuelle et al., 2018), it is the first to use the concept of state sequences and hidden states. The novel data collected by RapCoTT necessitated these approaches, but they can also be applied to other temporal or TCATA data. For example, the concept of state sequences is relevant to the temporal consumption experience of multiple sips or bites (Weerawarna N. R. P. et al., 2021). In addition, although not presenting the first example of fda for temporal sensory data (Bi, 2015; Bi & Kuesten, 2013), the power

and usefulness of this approach were thoroughly demonstrated.

However, some improvements to the Markov modelling approach could be made. Very few constraints were imposed on the model when specifying the mixture hidden Markov model for the genetic algorithm and final fitting. For example, modelling could be improved by including an imposed “starting state” where all participants report no sensation. In addition, it was only possible to include a single success criterion for the genetic algorithm. The `ga` package (Scrucca, 2013) includes an experimental feature that allows for the inclusion of more than one success criterion. It could be possible to include successful co-clustering of different responses from the same participant (for example, as a percentage of similarly classified participants across replicates) as a secondary success criterion. This approach could result in model parameters that produce more consistent classification across replicates. However, the underlying cause of inconsistently classified participants may be inconsistent thermal taste experiences, as this phenomenon is described elsewhere (Skinner et al., 2018).

This thesis also presented the first study in sensory science to use interpretable machine learning to untangle many related factors. Although the researcher has observed the use of similar approaches in industry, this practice deserves wider adoption in sensory science research.

Although Smilde et al. (2022) includes a theoretical example of using ROSA with sensory data, this thesis presented the first application of ROSA in a sensory science context. It is the first study to explore using ROSA with multiple outcome variables. In addition, it is the first study to explore the use of common components with ROSA. However, although the ROSA model delivered insights into the relative importance of predictors, the ROSA modelling approach failed to arrive at a model which was not

overfitted. Therefore, the modelling approach in ROSA requires further development to optimise its fit for multiple outcome variables. One suggestion is using dual cross-validation (Smilde et al., 2022), by cross-validating the component selection and model fit simultaneously. The optimal use of common components also requires further exploration, as higher orders of common components (for example, components extracted from any combination of the 14 data blocks) may have resulted in a better model. However, best practices for dealing with common component models must be developed before this approach can be put into practice.

## **6.4 Future work**

Although a myriad of applications in sensory science and adjacent fields of the data analysis methods presented are imagined, this section focuses on thermal taste-related work.

### **6.4.1 Better thermal taste study**

#### *6.4.1.1 Inclusion of other thermal “tastes”*

Studies using RapCoTT with an expanded list of orosensations, including astringency, chemesthetic cooling (mint-like), fatty, tingling, numbing and burning, will help to demystify additional sensations that may result from temperature exposure, and clarify whether some thermal tastes are misreported astringent or tingling orosensations. Such a study would include extensive training with an expanded set of stimuli.

#### *6.4.1.2 Translation of RapCoTT method*

There is an opportunity to implement a participant-driven individual translation, where participants can assign their own words, associations, or emojis to specific

tastes and orosensations during orosensory training. These assigned tastes will then be presented back to the participants during the RapCoTT protocol. This adaptation would be specifically intended for use with the expanded orosensations (see section 6.4.1.1), as groups of people differ in their ability to describe chemesthetic stimuli (Byrnes et al., 2015), and allowing participants to use their own associations will improve data quality and reduce the possibility of taste confusions. In addition, the use of own associations may avoid the conflation of taste identity and hedonic “recognition” with specific tastes (Good = Sweet, Bad = Bitter) (Davis & Running, 2023) in untrained participants.

#### *6.4.1.3 More detailed data collection*

An additional improvement for this type of study would be to include more replicates of the RapCoTT protocol per participant, allowing individual data to be added together and modelled as time-intensity data (Jaeger et al., 2020; Vidal et al., 2021). For example, functional clustering (Jacques & Preda, 2014a, 2014b; Leroy et al., 2018) can be used to differentiate between different response patterns for specific tastes, assisting in further answering the question of how different mechanisms of thermal taste are distributed across thermal tasters. In addition, multiple replicates would assist in interrogating the effect of protocol order found in Chapter 2, as for some participants, thermal taste may become more apparent or intense with multiple temperature exposure replicates.

### **6.4.2 Food behavioural studies**

#### *6.4.2.1 Spicy food*

Chapter 4 highlighted some differences between TT and TnT regarding their overall liking of spices and their spice addition behaviour. In addition, Chapter 5 pointed to

an inherent interaction between the temperature at which thermal taste is reported and spicy food liking in determining food choice. As perception of spicy orosensations is linked to temperature-sensitive taste channels (Latorre et al., 2009), this mechanistic and behavioural interaction warrants further investigation. In particular, potential differences in cross-modal enhancement between basic tastes and spicy stimuli (Han et al., 2021; Wang et al., 2022), as differences in this phenomenon likely underlie behavioural and orosensory differences between TT and TnT.

#### *6.4.2.2 Taste enjoyment*

Chapter 4 showed preliminary evidence of differences in hedonic responses to basic tastes based on TT Cluster (Cluster 1 were more likely to be HSL). Chapter 5 also presented evidence that the temperature at which thermal taste is reported interacts with sweet liking in determining food choice. In addition, Chapter 4 presented evidence that TT groups (TT status and TT Cluster) differ in their salt addition behaviour. Overall, this suggests that the mechanisms through which thermal taste is perceived likely affects the hedonic response to basic tastes.

A study exploring hedonic responses to basic tastes would not only include differences in sweet and salty tastes, but also umami, sour, fatty and bitter. Groupings of baseline hedonic responses to sour taste were recently established (Spinelli et al., 2024), and although differences in Savoury/Umami liking have been shown to be a cross cultural effect (Cecchini et al., 2019; Prescott et al., 1992) the researcher observed spontaneous, polarised hedonic responses to the MSG solution during orosensory training with participants. Although such a study could be executed using taste solutions, the recently published Taste Liking Questionnaire (Haydar et al., 2024) which measures the liking of taste and oral sensations as represented by food items presents an interesting alternative.

#### *6.4.2.3 Food involvement and sensory sensitivity*

The curious predictive effect found between TT Clusters in the Set & Disposal scale of food involvement warrants further investigation, especially as it relates to recent findings. Pickering and Pickering (2022) found TT to be more likely to be “foodies”, and in a subsequent study (Pickering, 2024) found “foodies” to differ significantly from “non-foodies” in Set & Disposal, which was correlated with food literacy (Park et al., 2020). These elements and the hypotheses around differences in temperature and (unpleasant) wetness sensation as described in Chapter 4 warrant further investigation.

#### **6.4.3 Genetic study on thermal taste**

The findings in Chapter 3 point to the involvement of several taste channels in thermal taste, including TRPM5, TRPA1, TRPM4, and TRPV4. A previous study (Bering, 2013) on thermal taste and genetics explored variations in TRPM5 expression amongst thermal tasters, finding no differences. However, recent studies (Kringel et al., 2018) have demonstrated diversity in the expression of TRP channels related to sensory sensitivity. Therefore, there is an opportunity to explore genetic variation between groups of TT and TnT, phenotyped using RapCoTT, based on the additional mechanisms linked to thermal taste in this thesis. In particular, exploring whether the subgroups identified in this thesis share a genetic link or whether diversity in the genes encoding specific channels is related to thermal taste reporting at the activation temperature on those mechanisms would be interesting. This investigation could be coupled with a taste sensitivity investigation, as there is preliminary evidence that sensitivity for specific tastes relates to the perception of thermal taste through similar mechanisms (Thibodeau et al., 2020). In addition, it would be paired well with any of the modified RapCoTT data collection suggestions in section 6.4.1.

#### **6.4.4 Further factors to assess**

##### *6.4.4.1 Interoceptive data*

Differences in Private Body Consciousness between TT and TnT in the present study indicate differences in interoceptive sensibility (the awareness of bodily sensations). Two additional constructs in interoception exist, namely interoceptive accuracy (the ability to sense bodily sensations accurately) and interoceptive awareness (confidence in awareness of bodily sensations) (Garfinkel et al., 2015). In addition, alexithymia, found previously not to differ between TT and TnT (Yang, 2015), can be considered a generalised deficit in interoception (Brewer et al., 2016; Murphy et al., 2018). Exploring different interoceptive constructs and measures will likely further demystify thermal taste reporting patterns.

##### *6.4.4.2 Sensory Imagery*

Chapter 4 presented the idea that some personality differences between TT and TnT may be linked to differences in the ability to imagine sensations, as similar dimensions on the BFI have been linked to better abilities to imagine sensory sensations. (Hitsuwari & Nomura, 2023). In addition, links have been found between the ability to imagine visual imagery and sensory sensitivity across domains (Dance et al., 2021), possibly explaining historical findings on thermal taste and orosensitivity. Therefore, it makes sense to evaluate thermal tasters using the Plymouth Sensory Imagery Questionnaire (Andrade et al., 2014). However, in this case, traditional TT phenotyping may be combined with RapCoTT, to link this aspect with traditional phenotyping.

#### **6.5 Conclusion**

The RapCoTT method provides the key to unlocking knowledge around thermal taste.

A second key is the use of advanced modelling. Although this thesis only scratched the surface of what can be achieved through the use of RapCoTT, it presented some key novel findings. These findings include establishing different groups of thermal tasters and evidence around existing and novel mechanisms related to thermal taste. In addition, differences in behavioural patterns (personality and food involvement) between TT groups suggest a behavioural component to thermal taste, which may either be the cause (through interaction with underlying biological mechanisms) or the result of different thermal taste experiences. Overall, these findings question existing paradigms around thermal taste as a binary phenotype, but also solidify the importance of the study of thermal taste.

## References

- Agresti, A. (2002). *Categorical data analysis* (2nd ed ed.). Wiley-Interscience.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723. <https://doi.org/10.1109/TAC.1974.1100705>
- Allen, A. L., McGeary, J. E., & Hayes, J. E. (2014). Polymorphisms in TRPV1 and TAS2Rs Associate with Sensations from Sampled Ethanol. *Alcoholism-Clinical and Experimental Research*, 38(10), 2550-2560. <https://doi.org/10.1111/acer.12527>
- Andrade, J., May, J., Deepröse, C., Baugh, S. J., & Ganis, G. (2014). Assessing vividness of mental imagery: The Plymouth Sensory Imagery Questionnaire. *British Journal of Psychology*, 105(4), 547-563. <https://doi.org/10.1111/bjop.12050>
- Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 82(4), 1059-1086. <https://doi.org/10.1111/rssb.12377>
- Ardebili, A. T., & Rickertsen, K. (2020). Personality traits, knowledge, and consumer acceptance of genetically modified plant and animal products. *Food Quality and Preference*, 80, 103825. <https://doi.org/10.1016/j.foodqual.2019.103825>
- Ares, G., Alcaire, F., Antunez, L., Vidal, L., Gimenez, 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>
- Armitage, R. M., Iatridi, V., Thanh Vi, C., & Yeomans, M. R. (2023). Phenotypic differences in taste hedonics: The effects of sweet liking. *Food Quality and Preference*, 107, 104845. <https://doi.org/10.1016/j.foodqual.2023.104845>
- Armitage, R. M., Iatridi, V., & Yeomans, M. R. (2021). Understanding sweet-liking phenotypes and their implications for obesity: Narrative review and future directions. *Physiology & Behavior*, 235, 113398. <https://doi.org/10.1016/j.physbeh.2021.113398>
- Askwith, C. C., Benson, C. J., Welsh, M. J., & Snyder, P. M. (2001). DEG/ENaC ion channels involved in sensory transduction are modulated by cold temperature. *Proceedings of the National Academy of Sciences of the United States of America*, 98(11), 6459-6463. <https://doi.org/10.1073/pnas.111155398>
- Azadi, G. G. (2021). *Multi block analysis of gastrointestinal neuroendocrine tumors data using response oriented sequential alternation (ROSA)* [MSc Thesis, Norwegian University of Life Sciences]. Norges miljø- og biovitenskapelige universitet Ås, Norway. <https://hdl.handle.net/11250/2771366>
- Azodi, C. B., Tang, J., & Shiu, S.-H. (2020). Opening the Black Box: Interpretable Machine Learning for Geneticists. *Trends in Genetics*, 36(6), 442-455. <https://doi.org/10.1016/j.tig.2020.03.005>
- Bajec, M. R., & Pickering, G. J. (2008). Thermal taste, PROP responsiveness, and perception of oral sensations. *Physiology & Behavior*, 95(4), 581-590. <https://doi.org/10.1016/j.physbeh.2008.08.009>
- Bajec, M. R., & Pickering, G. J. (2010). Association of thermal taste and PROP responsiveness with food liking, neophobia, body mass index, and waist circumference. *Food Quality and Preference*, 21(6), 589-601. <https://doi.org/10.1016/j.foodqual.2010.03.007>
- Bajec, M. R., Pickering, G. J., & DeCourville, N. (2012). Influence of Stimulus Temperature on Orosensory Perception and Variation with Taste Phenotype. *Chemosensory Perception*, 5(3-4), 243-265. <https://doi.org/10.1007/s12078-012-9129-5>
- Bangdiwala, S. (1985, 1985). A graphical test for observer agreement. 45th International Statistical Institute Meeting,
- Bangdiwala, S. I. (2017). Graphical aids for visualizing and interpreting patterns in departures from agreement in ordinal categorical observer agreement data. *Journal of Biopharmaceutical Statistics*, 27(5), 773-783. <https://doi.org/10.1080/10543406.2016.1273941>

- Bangdiwala, S. I., & Shankar, V. (2013). The agreement chart. *BMC medical research methodology*, 13, 97. <https://doi.org/10.1186/1471-2288-13-97>
- Banik, D. D., Martin, L. E., Freichel, M., Torregrossa, A. M., & Medler, K. F. (2018). TRPM4 and TRPM5 are both required for normal signaling in taste receptor cells. *Proceedings of the National Academy of Sciences of the United States of America*, 115(4), E772-E781. <https://doi.org/10.1073/pnas.1718802115>
- Barańczuk, U. (2019). The five factor model of personality and alexithymia: A meta-analysis. *Journal of Research in Personality*, 78, 227-248. <https://doi.org/10.1016/j.jrp.2018.12.005>
- Bartoshuk, L. M., Duffy, V. B., Green, B. G., Hoffman, H. J., Ko, C. W., Lucchina, L. A., Marks, L. E., Snyder, D. J., & Weiffenbach, J. M. (2004). Valid across-group comparisons with labeled scales: the gLMS versus magnitude matching. *Physiology & Behavior*, 82(1), 109-114. <https://doi.org/10.1016/j.physbeh.2004.02.033>
- Bell, R., & Marshall, D. W. (2003). The construct of food involvement in behavioral research: scale development and validation. *Appetite*, 40(3), 235-244. [https://doi.org/10.1016/s0195-6663\(03\)00009-6](https://doi.org/10.1016/s0195-6663(03)00009-6)
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1), 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Bering, A. (2013). *Genetic Analysis of Taste in Homo Sapiens* [MSc Thesis, Brock University]. Brock University Digital Repository. St. Catharines, Ontario, Canada. <http://hdl.handle.net/10464/4220>
- Bering, A. B., Pickering, G., & Liang, P. (2013). TAS2R38 Single Nucleotide Polymorphisms Are Associated with PROP—but Not Thermal—Tasting: a Pilot Study. *Chemosensory Perception*, 7(1), 23-30. <https://doi.org/10.1007/s12078-013-9160-1>
- Berry, D. N., & Simons, C. T. (2020). Do Caucasian American and South Asian Indian cultural groups differ in sensitivity to capsaicin? A study designed to control for chili pepper affinity. *Journal of Food Science*, 85(9), 2896-2901. <https://doi.org/10.1111/1750-3841.15369>
- Bi, J. (2015). Measurements of time intensity. In *Sensory Discrimination Tests and Measurements* (pp. 365-397). <https://doi.org/10.1002/9781118994863.ch14>
- Bi, J., & Chung, J. (2011). Identification of Drivers of Overall Liking - Determination of Relative Importances of Regressor Variables. *Journal of Sensory Studies*, 26(4), 245-254. <https://doi.org/10.1111/j.1745-459X.2011.00340.x>
- Bi, J., & Kuesten, C. (2013). Using Functional Data Analysis (FDA) Methodology and the R Package “fda” for Sensory Time-Intensity Evaluation. *Journal of Sensory Studies*, 28(6), 474-482. <https://doi.org/10.1111/joss.12072>
- Biancolillo, A., Liland, K. H., Måge, I., Næs, T., & Bro, R. (2016). Variable selection in multi-block regression. *Chemometrics and Intelligent Laboratory Systems*, 156, 89-101. <https://doi.org/10.1016/j.chemolab.2016.05.016>
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197-227. <https://doi.org/10.1007/s11749-016-0481-7>
- Biecek, P. (2018). DALEX: Explainers for Complex Predictive Models in R. *Journal of Machine Learning Research*, 19(84), 1-5. <https://jmlr.org/papers/v19/18-416.html>
- Biecek, P., & Burzykowski, T. (2021). *Explanatory Model Analysis*. Chapman and Hall/CRC, New York. <https://pbiecek.github.io/ema/>
- Bilder, C. R., & Loughin, T. M. (2014). *Analysis of categorical data with R*. CRC Press.
- Birder, L. A., Nakamura, Y., Kiss, S., Nealen, M. L., Barrick, S., Kanai, A. J., Wang, E., Ruiz, G., de Groat, W. C., Apodaca, G., Watkins, S., & Caterina, M. J. (2002). Altered urinary bladder function in mice lacking the vanilloid receptor TRPV1. *Nature Neuroscience*, 5(9), 856-860. <https://doi.org/10.1038/nn902>
- Blakeslee, A. F., & Fox, A. L. (1932). Our different taste worlds - P T C as a demonstration of genetic differences in taste. *Journal of Heredity*, 23(3), 97-107. <https://doi.org/DOI>

[10.1093/oxfordjournals.jhered.a103585](https://doi.org/10.1093/oxfordjournals.jhered.a103585)

- 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>
- Bouron, A., Kiselyov, K., & Oberwinkler, J. (2015). Permeation, regulation and control of expression of TRP channels by trace metal ions. *Pflügers Archiv-European Journal of Physiology*, 467(6), 1143-1164. <https://doi.org/10.1007/s00424-014-1590-3>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/a:1010933404324>
- Breslin, P. A. S., Izumi, A., Tharp, A., Ohkuri, T., Yokoo, Y., Flammer, L. J., Rawson, N. E., & Margolskee, R. F. (2021). Evidence that human oral glucose detection involves a sweet taste pathway and a glucose transporter pathway. *Plos One*, 16(10), e0256989-e0256989. <https://doi.org/10.1371/journal.pone.0256989>
- Brewer, R., Cook, R., & Bird, G. (2016). Alexithymia: a general deficit of interoception. *Royal Society Open Science*, 3(10), 150664. <https://doi.org/10.1098/rsos.150664>
- Byrnes, N., Loss, C. R., & Hayes, J. E. (2015). Perception of chemesthetic stimuli in groups who differ by food involvement and culinary experience. *Food Quality and Preference*, 46, 142-150. <https://doi.org/10.1016/j.foodqual.2015.07.017>
- Byrnes, N. K., Nestrud, M. A., & Hayes, J. E. (2015). Perceptual Mapping of Chemesthetic Stimuli in Naive Assessors. *Chemosensory Perception*, 8(1), 19-32. <https://doi.org/10.1007/s12078-015-9178-7>
- Byrnes, N. K., & Hayes, J. E. (2013). Personality factors predict spicy food liking and intake. *Food Quality and Preference*, 28(1), 213-221. <https://doi.org/10.1016/j.foodqual.2012.09.008>
- Byrnes, N. K., & Hayes, J. E. (2015). Gender differences in the influence of personality traits on spicy food liking and intake. *Food Quality and Preference*, 42, 12-19. <https://doi.org/10.1016/j.foodqual.2015.01.002>
- Byrnes, N. K., & Hayes, J. E. (2016). Behavioral measures of risk tasking, sensation seeking and sensitivity to reward may reflect different motivations for spicy food liking and consumption. *Appetite*, 103, 411-422. <https://doi.org/10.1016/j.appet.2016.04.037>
- Cao, G. Y., Li, M., Han, L., Tayie, F., Yao, S. S., Huang, Z., Ai, P., Liu, Y. Z., Hu, Y. H., & Xu, B. (2019). Dietary fat intake and cognitive function among older populations: a systematic review and meta-analysis. *The Journal Of Prevention of Alzheimer's Disease*, 1-8. <https://doi.org/10.14283/jpad.2019.9>
- Cao, X., Ma, L. L., Yang, F., Wang, K. W., & Zheng, J. (2014). Divalent cations potentiate TRPV1 channel by lowering the heat activation threshold. *Journal of General Physiology*, 143(1), 75-90. <https://doi.org/10.1085/jgp.201311025>
- Cardello, A. V., Chheang, S. L., Hedderley, D. I., Guo, L. F., Hunter, D. C., & Jaeger, S. R. (2019). Toward a new scale to measure consumers' "need for uniqueness" in foods and beverages: The 31-item FBNFU scale. *Food Quality and Preference*, 72, 159-171. <https://doi.org/10.1016/j.foodqual.2018.10.008>
- Cardinal, M., Chaussy, M., Donnay-Moreno, C., Cornet, J., Rannou, C., Fillonneau, C., Prost, C., Baron, R., & Courcoux, P. (2020). Use of random forest methodology to link aroma profiles to volatile compounds: Application to enzymatic hydrolysis of Atlantic salmon (*Salmo salar*) by-products combined with Maillard reactions. *Food Research International*, 134, 11, Article 109254. <https://doi.org/10.1016/j.foodres.2020.109254>
- 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. <https://doi.org/10.1111/rssc.12356>
- Carreiro, L. R. R., Haddad, H., & Baldo, M. V. C. (2011). Effects of intensity and positional predictability of a visual stimulus on simple reaction time. *Neuroscience Letters*, 487(3), 345-349.

- <https://doi.org/10.1016/j.neulet.2010.10.053>
- Castiglioni, A. J., & García-Añoveros, J. (2007). Chapter 8 - MechanoTRPs and TRPA1. In O. P. Hamill (Ed.), *Current Topics in Membranes* (Vol. 59, pp. 171-189). Academic Press. [https://doi.org/10.1016/S1063-5823\(06\)59008-8](https://doi.org/10.1016/S1063-5823(06)59008-8)
- Castura, J. C. (2018). Dynamics of Consumer Perception. In G. Ares & P. Varela (Eds.), *Methods in Consumer Research, Volume 1* (pp. 211-240). Woodhead Publishing. <https://doi.org/10.1016/b978-0-08-102089-0.00009-1>
- Castura, J. C. (2022). tempR: Temporal Sensory Data Analysis.
- 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. <https://doi.org/10.1016/j.foodqual.2015.06.017>
- 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, 104370. <https://doi.org/10.1016/j.foodqual.2021.104370>
- Cecchini, M. P., Knaapila, A., Hoffmann, E., Boschi, F., Hummel, T., & Iannilli, E. (2019). A cross-cultural survey of umami familiarity in European countries. *Food Quality and Preference*, 74, 172-178. <https://doi.org/10.1016/j.foodqual.2019.01.017>
- Chambers, D. H., Allison, A.-M. A., & Chambers, E. (2004). Training Effects on Performance of Descriptive Panelists. *Journal of Sensory Studies*, 19(6), 486-499. <https://doi.org/10.1111/j.1745-459X.2004.082402.x>
- Chamoun, E., Mutch, D. M., Allen-Vercoe, E., Buchholz, A. C., Duncan, A. M., Spriet, L. L., Haines, J., Ma, D. W. L., & Guelph Family Health, S. (2018). A review of the associations between single nucleotide polymorphisms in taste receptors, eating behaviors, and health. *Critical reviews in food science and nutrition*, 58(2), 194-207. <https://doi.org/10.1080/10408398.2016.1152229>
- Chao, Y. T., Woosch, D., Pieniak, M., & Hummel, T. (2022). Gender difference in ratings of odor intensity during olfactory training. *Journal of Sensory Studies*. <https://doi.org/10.1111/joss.12784>
- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, 61(6), 1 - 36. <https://doi.org/10.18637/jss.v061.i06>
- Choi, S. E., & Chan, J. (2015). Relationship of 6-n-propylthiouracil taste intensity and chili pepper use with body mass index, energy intake, and fat intake within an ethnically diverse population. *Journal of the Academy of Nutrition and Dietetics*, 115(3), 389-396. <https://doi.org/10.1016/j.jand.2014.09.001>
- Chong, J. S. X., Ng, G. J. P., Lee, S. C., & Zhou, J. (2017). Salience network connectivity in the insula is associated with individual differences in interoceptive accuracy. *Brain Structure and Function*, 222(4), 1635-1644. <https://doi.org/10.1007/s00429-016-1297-7>
- Clapperton, J. F., & Piggott, J. R. (1979). Flavour Characterization by Trained and Untrained Assessors. *Journal of the Institute of Brewing*, 85(5), 275-277. <https://doi.org/10.1002/j.2050-0416.1979.tb03922.x>
- 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>
- Clark, R. A. (2011). *Multimodal flavour perception: The effect of sweetness, bitterness, alcohol content and carbonation level on flavour perception* [PhD Thesis, University of Nottingham]. Nottingham eTheses. <https://eprints.nottingham.ac.uk/13432/>
- Cleaver, G., & Wedel, M. (2001). Identifying random-scoring respondents in sensory research using finite mixture regression models. *Food Quality and Preference*, 12(5-7), 373-384. [https://doi.org/10.1016/s0950-3293\(01\)00028-3](https://doi.org/10.1016/s0950-3293(01)00028-3)
- 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-s](https://doi.org/10.1016/0963-9969(93)90081-s)
- Cohen, A. (1980). On the graphical display of the significant components in two-way contingency tables. *Communications in Statistics - Theory and Methods*, 9(10), 1025-1041. <https://doi.org/10.1080/03610928008827940>
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and psychological measurement*, 20(1), 37-46. <https://doi.org/10.1177/001316446002000104>
- Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70(4), 213-220. <https://doi.org/10.1037/h0026256>
- Corneille, O., & Gawronski, B. (2024). Self-reports are better measurement instruments than implicit measures. *Nature Reviews Psychology*. <https://doi.org/10.1038/s44159-024-00376-z>
- Costanzo, A., Nowson, C., Orellana, L., Bolhuis, D., Duesing, K., & Keast, R. (2018). Effect of dietary fat intake and genetics on fat taste sensitivity: a co-twin randomized controlled trial. *American Journal of Clinical Nutrition*, 107(5), 683-694. <https://doi.org/10.1093/ajcn/nqy022>
- Cox, D. N., Hendrie, G. A., & Carty, D. (2016). Sensitivity, hedonics and preferences for basic tastes and fat amongst adults and children of differing weight status: A comprehensive review. *Food Quality and Preference*, 48, 359-367. <https://doi.org/10.1016/j.foodqual.2015.01.006>
- Cruz, A., & Green, B. G. (2000). Thermal stimulation of taste. *Nature*, 403(6772), 889-892. <https://doi.org/10.1038/35002581>
- da Veiga, G. C., Johann, G., Lima, V. A., Kaushik, N., & Mitterer - Daltoé, M. L. (2021). Food Choice Questionnaire and PLS - Path modeling as tools to understand interest in low sugar products. *Journal of Sensory Studies*, 36(5), e12667. <https://doi.org/10.1111/joss.12667>
- Dance, C. J., Ward, J., & Simner, J. (2021). What is the Link Between Mental Imagery and Sensory Sensitivity? Insights from Aphantasia. *Perception*, 50(9), 757-782. <https://doi.org/10.1177/03010066211042186>
- 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>
- Daoudi, H., Plesnik, J., Sayed, A., Sery, O., Rouabah, A., Rouabah, L., & Khan, N. A. (2015). Oral Fat Sensing and CD36 Gene Polymorphism in Algerian Lean and Obese Teenagers. *Nutrients*, 7(11), 9096-9104. <https://doi.org/10.3390/nu7115455>
- Davis, L. A., & Running, C. A. (2023). Good is sweet and bad is bitter: Conflation of affective value of aromas with taste qualities in untrained participants. *Journal of Sensory Studies*, 38(3). <https://doi.org/10.1111/joss.12820>
- Davison, J., Share, M., Hennessy, M., Bunting, B., Markovina, J., & Stewart-Knox, B. (2015). Correlates of food choice in unemployed young people: The role of demographic factors, self-efficacy, food involvement, food poverty and physical activity. *Food Quality and Preference*, 46, 40-47. <https://doi.org/10.1016/j.foodqual.2015.06.014>
- Dawes, C., O'Connor, A. M., & Aspen, J. M. (2000). The effect on human salivary flow rate of the temperature of a gustatory stimulus. *Arch Oral Biol*, 45(11), 957-961. [https://doi.org/10.1016/s0003-9969\(00\)00070-4](https://doi.org/10.1016/s0003-9969(00)00070-4)
- De Kock, H. L., Nkhabutlane, P., Kobue-Lekalake, R. I., Kriek, J., Steyn, A., Clarissa, V. H., Purdon, L., Kruger, C., Kinnear, M., Taljaard-Swart, H., & Tuorila, H. (2022). An alternative food neophobia scale (FNS-A) to quantify responses to new foods. *Food Quality and Preference*, 101, 104626. <https://doi.org/https://doi.org/10.1016/j.foodqual.2022.104626>
- de la Roche, J., Eberhardt, M. J., Klinger, A. B., Stanslowsky, N., Wegner, F., Koppert, W., Reeh, P. W., Lampert, A., Fischer, M. J. M., & Leffler, A. (2013). The Molecular Basis for Species-specific Activation of Human TRPA1 Protein by Protons Involves Poorly Conserved Residues within Transmembrane Domains 5 and 6. *Journal of Biological Chemistry*, 288(28), 20280-20292. <https://doi.org/10.1074/jbc.M113.479337>
- De Toffoli, A., Spinelli, S., Monteleone, E., Arena, E., Di Monaco, R., Endrizzi, I., Toschi, T. G., Laureati, M., Napolitano, F., Torri, L., & Dinnella, C. (2019). Influences of Psychological Traits and PROP

- Taster Status on Familiarity with and Choice of Phenol-Rich Foods and Beverages. *Nutrients*, 11(6), 23, Article 1329. <https://doi.org/10.3390/nu11061329>
- Deliza, R., & MacFie, H. J. H. (1996). The generation of sensory expectation by external cues and its effect on sensory perception and hedonic ratings: A review. *Journal of Sensory Studies*, 11(2), 103-128. <https://doi.org/10.1111/j.1745-459X.1996.tb00036.x>
- Deshaware, S., & Singhal, R. (2017). Genetic variation in bitter taste receptor gene TAS2R38, PROP taster status and their association with body mass index and food preferences in Indian population. *Gene*, 627, 363-368. <https://doi.org/10.1016/j.gene.2017.06.047>
- Dias, A. G., Rousseau, D., Duizer, L., Cockburn, M., Chiu, W., Nielsen, D., & El-Sohemy, A. (2013). Genetic Variation in Putative Salt Taste Receptors and Salt Taste Perception in Humans. *Chemical Senses*, 38(2), 137-145. <https://doi.org/10.1093/chemse/bjs090>
- Dirks, M., Turner, D., & Poole, D. (2023). Spectral sensor fusion for prediction of Li and Zr in rocks: Neural network and PLS methods. *Chemometrics and Intelligent Laboratory Systems*, 240, 104915. <https://doi.org/10.1016/j.chemolab.2023.104915>
- Doty, R. L., Chen, J. H., & Overend, J. (2017). Taste Quality Confusions: Influences of Age, Smoking, PTC Taster Status, and other Subject Characteristics. *Perception*, 46(3-4), 257-267. <https://doi.org/10.1177/0301006616685577>
- 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>
- Drummond, H. A. (2007). Chapter 6 - ENaC Proteins in Vascular Smooth Muscle Mechanotransduction. In O. P. Hamill (Ed.), *Current Topics in Membranes* (Vol. 59, pp. 127-153). Academic Press. [https://doi.org/10.1016/S1063-5823\(06\)59006-4](https://doi.org/10.1016/S1063-5823(06)59006-4)
- Dunn, J. C. (1974). Well-Separated Clusters and Optimal Fuzzy Partitions. *Journal of Cybernetics*, 4(1), 95-104. <https://doi.org/10.1080/01969727408546059>
- Dutta Banik, D., & Medler, K. F. (2023). Defining the role of TRPM4 in broadly responsive taste receptor cells. *Frontiers in Cellular Neuroscience*, 17. <https://doi.org/10.3389/fncel.2023.1148995>
- Dwyer, P., Ferrer, E., Saron, C. D., & Rivera, S. M. (2022). Exploring Sensory Subgroups in Typical Development and Autism Spectrum Development Using Factor Mixture Modelling. *Journal of Autism and Developmental Disorders*, 52(9), 21. <https://doi.org/10.1007/s10803-021-05256-6>
- Dwyer, P., Saron, C. D., & Rivera, S. M. (2020). Identification of Longitudinal Sensory Subtypes in Typical Development and Autism Spectrum Development Using Growth Mixture Modelling. *Research in Autism Spectrum Disorders*, 78, 16, Article 101645. <https://doi.org/10.1016/j.rasd.2020.101645>
- Ecarma, M. J. Y., & Nolden, A. A. (2021). A review of the flavor profile of metal salts: understanding the complexity of metallic sensation. *Chemical Senses*, 46, 8, Article bjab043. <https://doi.org/10.1093/chemse/bjab043>
- Eertmans, A., Victoir, A., Vansant, G., & Van den Bergh, O. (2005). Food-related personality traits, food choice motives and food intake: Mediator and moderator relationships. *Food Quality and Preference*, 16(8), 714-726. <https://doi.org/10.1016/j.foodqual.2005.04.007>
- Eggink, P. M., Maliepaard, C., Tikunov, Y., Haanstra, J. P. W., Bovy, A. G., & Visser, R. G. F. (2012). A taste of sweet pepper: Volatile and non-volatile chemical composition of fresh sweet pepper (*Capsicum annuum*) in relation to sensory evaluation of taste. *Food Chemistry*, 132(1), 301-310. <https://doi.org/10.1016/j.foodchem.2011.10.081>
- Eldeghaidy, S., Thomas, D., Skinner, M., Ford, R., Giesbrecht, T., Thomas, A., Hort, J., & Francis, S. (2018). An automated method to detect and quantify fungiform papillae in the human tongue: Validation and relationship to phenotypical differences in taste perception. *Physiology & Behavior*, 184, 226-234. <https://doi.org/10.1016/j.physbeh.2017.12.003>
- Eldeghaidy, S., Yang, Q., Abualait, T., Williamson, A. M., Hort, J., & Francis, S. T. (2021). Thermal taster

- status: Temperature modulation of cortical response to sweetness perception. *Physiology & Behavior*, 230, 113266. <https://doi.org/10.1016/j.physbeh.2020.113266>
- Falbel, D., Damiani, A., Hogervorst, R. M., Kuhn, M., & Couch, S. (2022). bonsai: Model Wrappers for Tree-Based Models. <https://CRAN.R-project.org/package=bonsai>
- Fay, M. P. (2010). Two-sided Exact Tests and Matching Confidence Intervals for Discrete Data. *The R Journal*, 2(1), 53 - 58. <https://doi.org/10.32614/rj-2010-008>
- Fenko, A., Leufkens, J.-M., & van Hoof, J. J. (2015). New product, familiar taste: Effects of slogans on cognitive and affective responses to an unknown food product among food neophobics and neophilics. *Food Quality and Preference*, 39, 268-276. <https://doi.org/10.1016/j.foodqual.2014.07.021>
- Ferentzi, E., Köteles, F., Csala, B., Drew, R., Tihanyi, B. T., Pulay-Kottlár, G., & Doering, B. K. (2017). What makes sense in our body? Personality and sensory correlates of body awareness and somatosensory amplification. *Personality and Individual Differences*, 104, 75-81. <https://doi.org/10.1016/j.paid.2016.07.034>
- Ferguson, R. J., & Ahles, T. A. (1998). Private body consciousness, anxiety and pain symptom reports of chronic pain patients. *Behaviour Research and Therapy*, 36(5), 527-535. [https://doi.org/10.1016/s0005-7967\(98\)00048-5](https://doi.org/10.1016/s0005-7967(98)00048-5)
- Ferrandiz-Huertas, C., Mathivanan, S., Wolf, C. J., Devesa, I., & Ferrer-Montiel, A. (2014). Trafficking of ThermoTRP Channels. *Membranes (Basel)*, 4(3), 525-564. <https://doi.org/10.3390/membranes4030525>
- Franks, E. M., Jeltema, M., Luck, P. J., Beckley, J., Foegeding, E. A., & Vinyard, C. J. (2020). Morphological and masticatory performance variation of mouth behavior groups [Article]. *Journal of Texture Studies*, 51(2), 343-351. <https://doi.org/10.1111/jtxs.12483>
- Friendly, M., & Meyer, D. (2016). *Discrete data analysis with R : visualization and modeling techniques for categorical and count data*. Taylor & Francis.
- Gabardinho, A., Ritschard, G., Muller, N. S., & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1-37. <https://doi.org/10.18637/jss.v040.i04>
- García-Esquinas, E., Ortolá, R., Banegas, J. R., Lopez-García, E., & Rodríguez-Artalejo, F. (2019). Dietary n-3 polyunsaturated fatty acids, fish intake and healthy ageing. *International Journal of Epidemiology*, 48(6), 1914-1924. <https://doi.org/10.1093/ije/dyz196>
- Garfinkel, S. N., Seth, A. K., Barrett, A. B., Suzuki, K., & Critchley, H. D. (2015). Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness. *Biological Psychology*, 104, 65-74. <https://doi.org/10.1016/j.biopsycho.2014.11.004>
- Garrity, P. A. (2011). Weakly acidic, but strongly irritating: TRPA1 and the activation of nociceptors by cytoplasmic acidification. *Journal of General Physiology*, 137(6), 489-491. <https://doi.org/10.1085/jgp.201110657>
- Gauthier, J.-A., Widmer, E. D., Bucher, P., & Notredame, C. (2010). 1. Multichannel Sequence Analysis Applied to Social Science Data. *Sociological Methodology*, 40(1), 1-38. <https://doi.org/10.1111/j.1467-9531.2010.01227.x>
- Gerritsen, S., Egli, V., Roy, R., Haszard, J., Backer, C. D., Teunissen, L., Cuykx, I., Decorte, P., Pabian, S. P., Van Royen, K., & Te Morenga, L. (2021). Seven weeks of home-cooked meals: changes to New Zealanders' grocery shopping, cooking and eating during the COVID-19 lockdown. *Journal of the Royal Society of New Zealand*, 51(sup1), S4-S22. <https://doi.org/10.1080/03036758.2020.1841010>
- Gibson, E. L., Held, I., Khawnekar, D., & Rutherford, P. (2016). Differences in Knowledge, Stress, Sensation Seeking, and Locus of Control Linked to Dietary Adherence in Hemodialysis Patients. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.01864>
- Graham, C. A. M., Pilic, L., King, A., Nixon, J. E., Pipe, J., Holton, J., Tamba, K., Hearne, G., Pedlar, C. R., Lorente-Cebrián, S., González Muniesa, P., & Mavrommatis, Y. (2021). Genetic differences in fat taste sensitivity and dietary intake in a UK female cohort. *Food Quality and Preference*, 92,

104202. <https://doi.org/10.1016/j.foodqual.2021.104202>
- Granic, A., Davies, K., Adamson, A., Kirkwood, T., Hill, T. R., Siervo, M., Mathers, J. C., & Jagger, C. (2016). Dietary Patterns High in Red Meat, Potato, Gravy, and Butter Are Associated with Poor Cognitive Functioning but Not with Rate of Cognitive Decline in Very Old Adults. *The Journal of Nutrition*, 146(2), 265-274. <https://doi.org/10.3945/jn.115.216952>
- Granitto, P. M., Biasioli, F., Endrizzi, I., & Gasperi, F. (2008). Discriminant models based on sensory evaluations: Single assessors versus panel average. *Food Quality and Preference*, 19(6), 589-595. <https://doi.org/10.1016/j.foodqual.2008.03.006>
- Granitto, P. M., Gasperi, F., Biasioli, F., Trainotti, E., & Furlanello, C. (2007). Modern data mining tools in descriptive sensory analysis: A case study with a Random forest approach. *Food Quality and Preference*, 18(4), 681-689. <https://doi.org/10.1016/j.foodqual.2006.11.001>
- Green, B., & Hayes, J. (2003). Capsaicin as a probe of the relationship between bitter taste and chemesthesis. *Physiology & Behavior*, 79(4-5), 811-821. [https://doi.org/10.1016/s0031-9384\(03\)00213-0](https://doi.org/10.1016/s0031-9384(03)00213-0)
- Green, B. G., Alvarez-Reeves, M., George, P., & Akirav, C. (2005). Chemesthesis and taste: evidence of independent processing of sensation intensity. *Physiology & Behavior*, 86(4), 526-537. <https://doi.org/10.1016/j.physbeh.2005.08.038>
- Green, B. G., Dalton, P., Cowart, B., Shaffer, G., Rankin, K., & Higgins, J. (1996). Evaluating the 'Labeled Magnitude Scale' for measuring sensations of taste and smell. *Chemical Senses*, 21(3), 323-334. <https://doi.org/10.1093/chemse/21.3.323>
- Green, B. G., & Frankmann, S. P. (1987). The effect of cooling the tongue on the perceived intensity of taste. *Chemical Senses*, 12(4), 609-619. <https://doi.org/10.1093/chemse/12.4.609>
- Green, B. G., & George, P. (2004). 'Thermal taste' predicts higher responsiveness to chemical taste and flavor. *Chemical Senses*, 29(7), 617-628. <https://doi.org/10.1093/chemse/bjh065>
- Green, B. G., & Nachtigal, D. (2015). Temperature Affects Human Sweet Taste via At Least Two Mechanisms. *Chemical Senses*, 40(6), 391-399. <https://doi.org/10.1093/chemse/bjv021>
- Green, B. G., & Schullery, M. T. (2003). Stimulation of bitterness by capsaicin and menthol: differences between lingual areas innervated by the glossopharyngeal and chorda tympani nerves. *Chemical Senses*, 28(1), 45-55. <https://doi.org/10.1093/chemse/28.1.45>
- Greenwell, B. M., & Boehmke, B. C. (2020). Variable Importance Plots—An Introduction to the vip Package. *The R Journal*, 12(1), 343-366. <https://doi.org/10.32614/rj-2020-013>
- Hamming, R. W. (1950). Error detecting and error correcting codes. *The Bell System Technical Journal*, 29(2), 147-160. <https://doi.org/10.1002/j.1538-7305.1950.tb00463.x>
- Han, P., Müller, L., & Hummel, T. (2021). Peri-threshold Trigeminal Stimulation with Capsaicin Increases Taste Sensitivity in Humans. *Chemosensory Perception*. <https://doi.org/10.1007/s12078-021-09285-4>
- Han, P., Penzler, M., Jonathan, W., & Hummel, T. (2020). Frequent minty chewing gum use is associated with increased trigeminal sensitivity: An fMRI study. *Brain Research*, 1730, 146663. <https://doi.org/10.1016/j.brainres.2020.146663>
- Haryono, R. Y., Sprajcer, M. A., & Keast, R. S. (2014). Measuring oral fatty acid thresholds, fat perception, fatty food liking, and papillae density in humans. *Jove-Journal of Visualized Experiments*(88), Article e51236. <https://doi.org/10.3791/51236>
- Hasbroucq, T., & Guiard, Y. (1992). The effects of intensity and irrelevant location of a tactile stimulation in a choice reaction time task. *Neuropsychologia*, 30(1), 91-94. [https://doi.org/10.1016/0028-3932\(92\)90017-g](https://doi.org/10.1016/0028-3932(92)90017-g)
- Haydar, S., Karlsson, C. C., Linneberg, A., Kårhus, L. L., Ångquist, L., Hansen, T., Bredie, W., & Grarup, N. (2024). The Taste Liking Questionnaire (TasteLQ) – A tool for assessment of taste liking in the Danish population. *Food Quality and Preference*, 117, 105176. <https://doi.org/10.1016/j.foodqual.2024.105176>
- Hayes, J. E., & Keast, R. S. (2011). Two decades of supertasting: where do we stand? *Physiology & Behavior*, 104(5), 1072-1074. <https://doi.org/10.1016/j.physbeh.2011.08.003>

- Hazley, D., McCarthy, S. N., Stack, M., Walton, J., McNulty, B. A., Flynn, A., & Kearney, J. M. (2022). Food neophobia and its relationship with dietary variety and quality in Irish adults: Findings from a national cross-sectional study. *Appetite*, *169*, 105859. <https://doi.org/10.1016/j.appet.2021.105859>
- Helske, S. (2018). Examples and Tips for Estimating Markovian Models with seqHMM. *R package vignette*. [https://CRAN.R-project.org/web/packages/seqHMM/vignettes/seqHMM\\_estimation.pdf](https://CRAN.R-project.org/web/packages/seqHMM/vignettes/seqHMM_estimation.pdf) .
- Helske, S., & Helske, J. (2019). Mixture Hidden Markov Models for Sequence Data: The seqHMM Package in R. *Journal of Statistical Software*, *88*(3), 1-32. <https://doi.org/10.18637/jss.v088.i03>
- Henriques, A. S., King, S. C., & Meiselman, H. L. (2009). Consumer segmentation based on food neophobia and its application to product development. *Food Quality and Preference*, *20*(2), 83-91. <https://doi.org/10.1016/j.foodqual.2008.01.003>
- Hettinger, T. P., Gent, J. F., Marks, L. E., & Frank, M. E. (1999). A confusion matrix for the study of taste perception. *Perception & Psychophysics*, *61*(8), 1510-1521. <https://doi.org/10.3758/bf03213114>
- Heyer, E., Balaesque, P., Jobling, M. A., Quintana-Murci, L., Chaix, R., Segurel, L., Aldashev, A., & Hegay, T. (2009). Genetic diversity and the emergence of ethnic groups in Central Asia. *BMC Genetics*, *10*(1), 49. <https://doi.org/10.1186/1471-2156-10-49>
- Higgins, M. J., Bakke, A. J., & Hayes, J. E. (2020). Personality traits and bitterness perception influence the liking and intake of pale ale style beers. *Food Quality and Preference*, *86*, Article 103994. <https://doi.org/10.1016/j.foodqual.2020.103994>
- Hitsuwari, J., & Nomura, M. (2023). Developing and validating a Japanese version of the Plymouth Sensory Imagery Questionnaire. *Frontiers in Psychology*, *14*, 1166543. <https://doi.org/10.3389/fpsyg.2023.1166543>
- Holt, S. H. A., Cobiac, L., Beaumont-Smith, N. E., Easton, K., & Best, D. J. (2000). Dietary habits and the perception and liking of sweetness among Australian and Malaysian students: A cross-cultural study. *Food Quality and Preference*, *11*(4), 299-312. [https://doi.org/10.1016/s0950-3293\(99\)00076-2](https://doi.org/10.1016/s0950-3293(99)00076-2)
- Honnens de Lichtenberg Broge, E., Wendin, K., Rasmussen, M. A., & Bredie, W. L. P. (2021). Changes in perception and liking for everyday food odors among older adults. *Food Quality and Preference*, *93*, 104254. <https://doi.org/10.1016/j.foodqual.2021.104254>
- Hort, J., Ford, R. A., Eldeghaidy, S., & Francis, S. T. (2016). Thermal taster status: Evidence of cross-modal integration. *Human Brain Mapping*, *37*(6), 2263-2275. <https://doi.org/10.1002/hbm.23171>
- Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., & Van Der Laan, M. (2006). Survival Ensembles. *Biostatistics*, *7*(3), 355-373. <https://doi.org/10.1093/biostatistics/kxj011>
- Hothorn, T., & Zeileis, A. (2015). partykit: A modular toolkit for recursive partytioning in R. *The Journal of Machine Learning Research*, *16*(1), 3905-3909.
- Hu, C., Huang, K., Hu, X., Liu, Y., Yuan, F., Wang, Q., & Fu, G. (2015). Measuring the cognitive resources consumed per second for real-time lie-production and recollection: a dual-tasking paradigm. *Frontiers in Psychology*, *6*(596). <https://doi.org/10.3389/fpsyg.2015.00596>
- Huang, X., Shang, H. L., & Pitt, D. (2022). Permutation entropy and its variants for measuring temporal dependence. *Australian & New Zealand Journal of Statistics*, *64*(4), 442-477. <https://doi.org/10.1111/anzs.12376>
- Hui, B., & Godfroid, A. (2021). Testing the Role of Processing Speed and Automaticity in Second Language Listening. *Applied Psycholinguistics*, *42*(5), 1089-1115. <https://doi.org/10.1017/S0142716420000193>
- Hvitfeldt, E. (2022). themis: Extra Recipes Steps for Dealing with Unbalanced Data. <https://CRAN.R-project.org/package=themis>
- Iatridi, V., Armitage, R. M., Yeomans, M. R., & Hayes, J. E. (2020). Effects of Sweet-Liking on Body

- Composition Depend on Age and Lifestyle: A Challenge to the Simple Sweet-Liking-Obesity Hypothesis. *Nutrients*, 12(9), 2702. <https://doi.org/10.3390/nu12092702>
- Iatridi, V., Quadt, L., Hayes, J. E., Garfinkel, S. N., & Yeomans, M. R. (2021). Female sweet-likers have enhanced cross-modal interoceptive abilities. *Appetite*, 165, 105290. <https://doi.org/10.1016/j.appet.2021.105290>
- Jacques, J., & Preda, C. (2014a). Functional data clustering: a survey. *Advances in Data Analysis and Classification*, 8(3), 231-255. <https://doi.org/10.1007/s11634-013-0158-y>
- Jacques, J., & Preda, C. (2014b). Model-based clustering for multivariate functional data. *Computational Statistics & Data Analysis*, 71, 92-106. <https://doi.org/10.1016/j.csda.2012.12.004>
- 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., Andani, Z., Wakeling, I. N., & MacFie, H. J. H. (1998). Consumer preferences for fresh and aged apples: A cross-cultural comparison. *Food Quality and Preference*, 9(5), 355-366. [https://doi.org/10.1016/s0950-3293\(98\)00031-7](https://doi.org/10.1016/s0950-3293(98)00031-7)
- 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., Chheang, S. L., Jin, D., Roigard, C. M., & Ares, G. (2020). Check-all-that-apply (CATA) questions: Sensory term citation frequency reflects rated term intensity and applicability. *Food Quality and Preference*, 86, 103986. <https://doi.org/10.1016/j.foodqual.2020.103986>
- Jaeger, S. R., Chheang, S. L., Roigard, C. M., & Cardello, A. V. (2022). Individual differences in food neophobia and private body consciousness influence product-elicited emotional valence and arousal. *Food Quality and Preference*, 99, 104566. <https://doi.org/10.1016/j.foodqual.2022.104566>
- Jaeger, S. R., & Hedderley, D. I. (2013). Impact of individual differences in emotional intensity and private body consciousness on EsSense Profile® responses. *Food Quality and Preference*, 27(1), 54-62. <https://doi.org/10.1016/j.foodqual.2012.06.004>
- Jaeger, S. R., Rasmussen, M. A., & Prescott, J. (2017). Relationships between food neophobia and food intake and preferences: Findings from a sample of New Zealand adults. *Appetite*, 116, 410-422. <https://doi.org/10.1016/j.appet.2017.05.030>
- James, B. (2018). Oral processing and texture perception influences satiation. *Physiology & Behavior*, 193(Pt B), 238-241. <https://doi.org/10.1016/j.physbeh.2018.03.015>
- Jeltema, M., Beckley, J., & Vahalik, J. (2015). Model for understanding consumer textural food choice. *Food Science & Nutrition*, 3(3), 202-212. <https://doi.org/10.1002/fsn3.205>
- Jeltema, M., Beckley, J., & Vahalik, J. (2016). Food texture assessment and preference based on Mouth Behavior. *Food Quality and Preference*, 52, 160-171. <https://doi.org/10.1016/j.foodqual.2016.04.010>
- 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>
- Jeltema, M., Beckley, J. H., & Vahalik, J. (2014). Importance of understanding mouth behavior when optimizing product texture now and in the future. In *Food Texture Design and Optimization* (pp. 423-442). <https://doi.org/10.1002/9781118765616.ch17>
- Jezewska-Zychowicz, M., Gębski, J., & Kobylińska, M. (2020). Food Involvement, Eating Restrictions and Dietary Patterns in Polish Adults: Expected Effects of Their Relationships (LifeStyle Study). *Nutrients*, 12(4), 1200. <https://doi.org/10.3390/nu12041200>
- Jilani, H., Intemann, T., Buchecker, K., Charalambos, H., Gianfagna, F., De Henauw, S., Lauria, F., Molnar, D., Moreno, L. A., Lissner, L., Pala, V., Siani, A., Veidebaum, T., Ahrens, W., &

- Hebestreit, A. (2022). Correlates of bitter, sweet, salty and umami taste sensitivity in European children: Role of sex, age and weight status - The IDEFICS study. *Appetite*, 175, 106088. <https://doi.org/10.1016/j.appet.2022.106088>
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research*, 3rd ed. (pp. 114-158). The Guilford Press.
- John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research*, 2nd ed. (pp. 102-138). Guilford Press.
- Karami, A., Moosavi, A. A., Pourghasemi, H. R., Ronaghi, A., Ghasemi-Fasaei, R., Vidal, E., & Lado, M. (2024). Proximal sensing approach for characterization of calcareous soils using multiblock data analysis. *Geoderma Regional*, 36, e00752. <https://doi.org/10.1016/j.geodrs.2023.e00752>
- Karmous, I., Doggui, R., Sayed Khan, A., Ben Amor, N., Khan, N. A., & Jamoussi, H. (2022). Is fat taste associated with diet quality? A cross-sectional study conducted among Tunisian adults. *Appetite*, 176, 106138. <https://doi.org/10.1016/j.appet.2022.106138>
- Kasahara, Y., Narukawa, M., Ishimaru, Y., Kanda, S., Umatani, C., Takayama, Y., Tominaga, M., Oka, Y., Kondo, K., Kondo, T., Takeuchi, A., Misaka, T., Abe, K., & Asakura, T. (2021). TMC4 is a novel chloride channel involved in high-concentration salt taste sensation. *The Journal of Physiological Sciences*, 71(1), 23, Article 23. <https://doi.org/10.1186/s12576-021-00807-z>
- Kasahara, Y., Narukawa, M., Kanda, S., Tominaga, M., Abe, K., Misaka, T., & Asakura, T. (2021). Transmembrane channel-like 4 is involved in pH and temperature-dependent modulation of salty taste. *Bioscience Biotechnology and Biochemistry*, 85(11), 2295-2299. <https://doi.org/10.1093/bbb/zbab152>
- Kasahara, Y., Narukawa, M., Takeuchi, A., Tominaga, M., Abe, K., & Asakura, T. (2022). Molecular logic of salt taste reception in special reference to transmembrane channel-like 4 (TMC4). *The Journal of Physiological Sciences*, 72(1). <https://doi.org/10.1186/s12576-022-00856-y>
- Kassambara, A. (2017). *Practical guide to Cluster Analysis in R : unsupervised machine learning*. [CreateSpace Independent Publishing Platform].
- Kavaliauskaite, G., Thibodeau, M., Ford, R., & Yang, Q. (2023). Using correlation matrices to standardise sweet liking status classification. *Food Quality and Preference*, 104, 104759. <https://doi.org/10.1016/j.foodqual.2022.104759>
- Kay, K. E., Martin, L. E., James, K. F., Haygood, S. M., & Torregrossa, A.-M. (2020). Temperature Is Sufficient to Condition a Flavor Preference for a Cold-Paired Solution in Rats. *Chemical Senses*, 45(7), 563-572. <https://doi.org/10.1093/chemse/bjaa050>
- Keast, R. S. J. (2016). Effects of sugar and fat consumption on sweet and fat taste. *Current Opinion in Behavioral Sciences*, 9, 55-60. <https://doi.org/10.1016/j.cobeha.2015.12.003>
- Keesman, M., Aarts, H., Vermeent, S., Häfner, M., & Papies, E. K. (2016). Consumption Simulations Induce Salivation to Food Cues. *Plos One*, 11(11), e0165449. <https://doi.org/10.1371/journal.pone.0165449>
- Keller, A., Gerkin, R. C., Guan, Y. F., Dhurandhar, A., Turu, G., Szalai, B., Mainland, J. D., Ihara, Y., Yu, C. W., Wolfinger, R., Vens, C., Schietgat, L., De Grave, K., Norel, R., Stolovitzky, G., Cecchi, G. A., Vosshall, L. B., Meyer, P., & Prediction, D. O. (2017). Predicting human olfactory perception from chemical features of odor molecules. *Science*, 355(6327), 820-+. <https://doi.org/10.1126/science.aal2014>
- Keller, C., & Siegrist, M. (2015). Does personality influence eating styles and food choices? Direct and indirect effects. *Appetite*, 84, 128-138. <https://doi.org/10.1016/j.appet.2014.10.003>
- Kelling, S., & Halpern, B. (1983). Taste flashes: reaction times, intensity, and quality. *Science*, 219(4583), 412-414. <https://doi.org/10.1126/science.6849142>
- Kelling, S. T., & Halpern, B. P. (1987). Taste judgments and gustatory stimulus duration: simple taste reaction times. *Chemical Senses*, 12(4), 543-562. <https://doi.org/10.1093/chemse/12.4.543>

- Ketel, E. C., de Wijk, R. A., de Graaf, C., & Stieger, M. (2020). Relating oral physiology and anatomy of consumers varying in age, gender and ethnicity to food oral processing behavior. *Physiology & Behavior*, *215*, 112766. <https://doi.org/10.1016/j.physbeh.2019.112766>
- Ketel, E. C., Zhang, Y., Jia, J., Wang, X., de Wijk, R. A., Chen, J., & Stieger, M. (2021). Comparison of and relationships between oral physiology, anatomy and food oral processing behavior of Chinese (Asian) and Dutch (Caucasian) consumers differing in age. *Physiology & Behavior*, *232*, 113284. <https://doi.org/10.1016/j.physbeh.2020.113284>
- Kievit, R., Frankenhuys, W., Waldorp, L., & Borsboom, D. (2013). Simpson's paradox in psychological science: a practical guide. *Frontiers in Psychology*, *4*. <https://doi.org/10.3389/fpsyg.2013.00513>
- Kim, E. H. J., Paredes, D., Motoi, L., Eckert, M., Wadamori, Y., Tartaglia, J., Green, C., Hedderley, D. I., & Morgenstern, M. P. (2019). Dynamic flavor perception of encapsulated flavors in a soft chewable matrix. *Food Research International*, *123*, 241-250. <https://doi.org/10.1016/j.foodres.2019.04.038>
- Kim, M. J., Son, H. J., Kim, Y., Kweon, H. J., Suh, B. C., Lyall, V., & Rhyu, M. R. (2014). Selective Activation of hTRPV1 by N-Geranyl Cyclopropylcarboxamide, an Amiloride-Insensitive Salt Taste Enhancer. *Plos One*, *9*(2), 10, Article e89062. <https://doi.org/10.1371/journal.pone.0089062>
- Kim, S.-E., Lee, S. M., & Kim, K.-O. (2016). Consumer acceptability of coffee as affected by situational conditions and involvement. *Food Quality and Preference*, *52*, 124-132. <https://doi.org/10.1016/j.foodqual.2016.04.008>
- Kim, S., & Vickers, Z. (2020). Liking of food textures and its relationship with oral physiological parameters and mouth-behavior groups. *Journal of Texture Studies*, *51*(3), 412-425. <https://doi.org/10.1111/jtxs.12504>
- Knaapila, A., Hwang, L. D., Lysenko, A., Duke, F. F., Fesi, B., Khoshnevisan, A., James, R. S., Wysocki, C. J., Rhyu, M., Tordoff, M. G., Bachmanov, A. A., Mura, E., Nagai, H., & Reed, D. R. (2012). Genetic analysis of chemosensory traits in human twins. *Chemical Senses*, *37*(9), 869-881. <https://doi.org/10.1093/chemse/bjs070>
- Knaapila, A., Silventoinen, K., Broms, U., Rose, R. J., Perola, M., Kaprio, J., & Tuorila, H. M. (2011). Food neophobia in young adults: genetic architecture and relation to personality, pleasantness and use frequency of foods, and body mass index--a twin study. *Behavior Genetics*, *41*(4), 512-521. <https://doi.org/10.1007/s10519-010-9403-8>
- Kornelis, M., Herpen, E. v., Lans, I. v. d., & Aramyan, L. (2010). Using non-food information to identify food-choice segment membership. *Food Quality and Preference*, *21*(5), 512-520. <https://doi.org/10.1016/j.foodqual.2010.01.007>
- Kringel, D., Geisslinger, G., Resch, E., Oertel, B. G., Thrun, M. C., Heinemann, S., & Lötsch, J. (2018). Machine-learned analysis of the association of next-generation sequencing-based human TRPV1 and TRPA1 genotypes with the sensitivity to heat stimuli and topically applied capsaicin. *Pain*, *159*(7), 1366-1381. <https://doi.org/10.1097/j.pain.0000000000001222>
- Kuhn, M. (2022). finetune: Additional Functions for Model Tuning. <https://CRAN.R-project.org/package=finetune>
- Kuhn, M., & Wickham, H. (2020). Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. <https://www.tidymodels.org>
- Kumar, S., & Kumar, S. (2021). Absolute Temperature Thresholds for Detection of Skin Wetness and Dampness on the Hand and their Variation with Sex and Age. *Perception*, *30*10066211043467. <https://doi.org/10.1177/03010066211043467>
- Kurogi, M., Kawai, Y., Nagatomo, K., Tateyama, M., Kubo, Y., & Saitoh, O. (2015). Auto-oxidation Products of Epigallocatechin Gallate Activate TRPA1 and TRPV1 in Sensory Neurons. *Chemical Senses*, *40*(1), 27-46. <https://doi.org/10.1093/chemse/bju057>
- Kwiecien, R., Kopp-Schneider, A., & Blettner, M. (2011). Concordance analysis: part 16 of a series on evaluation of scientific publications. *Deutsches Arzteblatt international*, *108*(30), 515-521. <https://doi.org/10.3238/arztebl.2011.0515>

- Labbe, D., Rytz, A., & Hugi, A. (2004). Training is a critical step to obtain reliable product profiles in a real food industry context. *Food Quality and Preference*, *15*(4), 341-348. [https://doi.org/10.1016/s0950-3293\(03\)00081-8](https://doi.org/10.1016/s0950-3293(03)00081-8)
- Lahne, J., Trubek, A. B., & Pelchat, M. L. (2014). Consumer sensory perception of cheese depends on context: A study using comment analysis and linear mixed models. *Food Quality and Preference*, *32*, 184-197. <https://doi.org/10.1016/j.foodqual.2013.10.007>
- Landau, W. (2021). The targets R package: a dynamic Make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. *Journal of Open Source Software*, *6*(57), 2959. <https://doi.org/10.21105/joss.02959>
- Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., & Morales, J. M. (2012). Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. *Ecology*, *93*(11), 2336-2342. <https://doi.org/10.1890/11-2241.1>
- Latorre, R., Zaelzer, C., & Brauchi, S. (2009). Structure-functional intimacies of transient receptor potential channels. *Quarterly Reviews of Biophysics*, *42*(3), 201-246. <https://doi.org/10.1017/S0033583509990072>
- Laugerette, F., Gaillard, D., Passilly-Degrace, P., Niot, I., & Besnard, P. (2007). Do we taste fat? *Biochimie*, *89*(2), 265-269. <https://doi.org/10.1016/j.biochi.2006.10.011>
- Laugerette, F., Passilly-Degrace, P., Patris, B., Niot, I., Febbraio, M., Montmayeur, J. P., & Besnard, P. (2005). CD36 involvement in orosensory detection of dietary lipids, spontaneous fat preference, and digestive secretions. *The Journal of Clinical Investigation*, *115*(11), 3177-3184. <https://doi.org/10.1172/JCI25299>
- Laureati, M., Sandvik, P., Almli, V. L., Sandell, M., Zeinstra, G. G., Methven, L., Wallner, M., Jilani, H., Alfaro, B., & Proserpio, C. (2020). Individual differences in texture preferences among European children: Development and validation of the Child Food Texture Preference Questionnaire (CFTPQ). *Food Quality and Preference*, *80*, Article Unsp 103828. <https://doi.org/10.1016/j.foodqual.2019.103828>
- Laureati, M., Spinelli, S., Monteleone, E., Dinnella, C., Prescott, J., Cattaneo, C., Proserpio, C., De Toffoli, A., Gasperi, F., Endrizzi, I., Torri, L., Peparajo, M., Arena, E., Bonello, F., Condelli, N., Di Monaco, R., Gatti, E., Piasentier, E., Tesini, F., & Pagliarini, E. (2018). Associations between food neophobia and responsiveness to “warning” chemosensory sensations in food products in a large population sample. *Food Quality and Preference*, *68*, 113-124. <https://doi.org/10.1016/j.foodqual.2018.02.007>
- Laursen, W. J., Anderson, E. O., Hoffstaetter, L. J., Bagriantsev, S. N., & Gracheva, E. O. (2015). Species-specific temperature sensitivity of TRPA1. *Temperature*, *2*(2), 214-226. <https://doi.org/10.1080/23328940.2014.1000702>
- Lawless, H. T. (2013). Basics of Signal Detection Theory. In *Quantitative Sensory Analysis* (pp. 47-70). <https://doi.org/10.1002/9781118684818.ch3>
- Lawless, H., Rozin, P., & Shenker, J. (1985). Effects of oral capsaicin on gustatory, olfactory and irritant sensations and flavor identification in humans who regularly or rarely consume chili pepper. *Chemical Senses*, *10*(4), 579-589. <https://doi.org/10.1093/chemse/10.4.579>
- Lawless, H. T., & Heymann, H. (2010). *Sensory evaluation of food : principles and practices* (2nd ed ed.). Springer.
- Lecuelle, G., Visalli, M., Cardot, H., & Schlich, P. (2018). Modeling Temporal Dominance of Sensations with semi-Markov chains. *Food Quality and Preference*, *67*, 59-66. <https://doi.org/10.1016/j.foodqual.2017.06.003>
- Lee, A., Guest, S., & Essick, G. (2006). Thermally evoked parotid salivation. *Physiology & Behavior*, *87*(4), 757-764. <https://doi.org/10.1016/j.physbeh.2006.01.021>
- Lemon, C. H. (2017). Modulation of taste processing by temperature. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *313*(4), R305-R321. <https://doi.org/10.1152/ajpregu.00089.2017>
- Leroy, A., Marc, A., Dupas, O., Rey, J. L., & Gey, S. (2018). Functional Data Analysis in Sport Science:

- Example of Swimmers' Progression Curves Clustering. *Applied Sciences*, 8(10), 1766. <https://doi.org/10.3390/app8101766>
- Leth-Steensen, C., Elbaz, Z. K., & Douglas, V. I. (2000). Mean response times, variability and skew in the responding of ADHD children: A response time distributional approach. *Acta Psychologica*, 104(2), 167-190. [https://doi.org/10.1016/S0001-6918\(00\)00019-6](https://doi.org/10.1016/S0001-6918(00)00019-6)
- Li, Q., Cui, Y., Jin, R., Lang, H., Yu, H., Sun, F., He, C., Ma, T., Li, Y., Zhou, X., Liu, D., Jia, H., Chen, X., & Zhu, Z. (2017). Enjoyment of Spicy Flavor Enhances Central Salty-Taste Perception and Reduces Salt Intake and Blood Pressure. *Hypertension*, 70(6), 1291-1299. <https://doi.org/10.1161/hypertensionaha.117.09950>
- Li, Y., & James, B. (2021). Oral processing preference affects flavor perception in dark chocolate with added ingredients. *Journal of Food Science*, 86(1), 177-183. <https://doi.org/10.1111/1750-3841.15557>
- Liang, P., Jiang, J., Ding, Q., Tang, X., & Roy, S. (2018). Memory Load Influences Taste Sensitivities. *Frontiers in Psychology*, 9(2533), 2533. <https://doi.org/10.3389/fpsyg.2018.02533>
- Liang, Z., Wilson, C. E., Teng, B., Kinnamon, S. C., & Liman, E. R. (2023). The proton channel OTO1 is a sensor for the taste of ammonium chloride. *Nature Communications*, 14(1). <https://doi.org/10.1038/s41467-023-41637-4>
- Liland, K. H. (2022). multiblock: Multiblock Data Fusion in Statistics and Machine Learning. <https://CRAN.R-project.org/package=multiblock>
- Liland, K. H., Naes, T., & Indahl, U. G. (2016). ROSA-a fast extension of partial least squares regression for multiblock data analysis. *Journal of Chemometrics*, 30(11), 651-662. <https://doi.org/10.1002/cem.2824>
- Lim, A. J. Y., Teo, P. S., Tan, V. W. K., & Forde, C. G. (2020). Associations between Psycho-Hedonic Responses to Sweet and Savoury Tastes with Diet and Body Composition in a Sample of Asian Females. *Foods*, 9(9), 1318. <https://doi.org/10.3390/foods9091318>
- Lim, J., Urban, L., & Green, B. G. (2008). Measures of individual differences in taste and creaminess perception. *Chemical Senses*, 33(6), 493-501. <https://doi.org/10.1093/chemse/bjn016>
- Lisofsky, N., Kazzer, P., Heekeren, H. R., & Prehn, K. (2014). Investigating socio-cognitive processes in deception: A quantitative meta-analysis of neuroimaging studies. *Neuropsychologia*, 61, 113-122. <https://doi.org/10.1016/j.neuropsychologia.2014.06.001>
- Liu, J., Shen, Q., Zhang, J., Beyens, U., Cai, W., Decety, J., & Li, H. (2021). The Difference Spotting Task: A new nonverbal measure of cheating behavior. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-020-01526-w>
- Liu, P., Shah, B. P., Croasdell, S., & Gilbertson, T. A. (2011). Transient receptor potential channel type M5 is essential for fat taste. *The Journal of Neuroscience*, 31(23), 8634-8642. <https://doi.org/10.1523/JNEUROSCI.6273-10.2011>
- Liu, X. B., Ong, H. L., & Ambudkar, I. (2018). TRP Channel Involvement in Salivary Glands-Some Good, Some Bad. *Cells*, 7(7), 18, Article 74. <https://doi.org/10.3390/cells7070074>
- Lo Piano, S. (2020). Ethical principles in machine learning and artificial intelligence: cases from the field and possible ways forward. *Humanities and Social Sciences Communications*, 7(1). <https://doi.org/10.1057/s41599-020-0501-9>
- Looy, H., Callaghan, S., & Weingarten, H. P. (1992). Hedonic response of sucrose likers and dislikers to other gustatory stimuli. *Physiology & Behavior*, 52(2), 219-225. [https://doi.org/10.1016/0031-9384\(92\)90261-y](https://doi.org/10.1016/0031-9384(92)90261-y)
- Lötsch, J., Kringel, D., Geisslinger, G., Oertel, B. G., Resch, E., & Malkusch, S. (2020). Machine-Learned Association of Next-Generation Sequencing-Derived Variants in Thermosensitive Ion Channels Genes with Human Thermal Pain Sensitivity Phenotypes. *International Journal of Molecular Sciences*, 21(12), 4367. <https://doi.org/10.3390/ijms21124367>
- Lovell, A. L., Roy, R., Klein, A., Cavadino, A., Foster, M., Krebs, J. D., Braakhuis, A., & Merry, T. L. (2023). Habitual Dietary Patterns, Nutrient Intakes, and Adherence to the Mediterranean Diet among New Zealand Adults: The NZ MED Cross-Sectional Study. *Nutrients*, 15(12), 2663.

- <https://doi.org/10.3390/nu15122663>
- Lu, N.-V., Vuong, T.-N., & Dinh, D.-T. (2020). Combining Correlation-Based Feature and Machine Learning for Sensory Evaluation of Saigon Beer. *International Journal of Knowledge and Systems Science*, 11(2), 71-85. <https://doi.org/10.4018/ijkss.2020040104>
- Lubbert, M., Kyereme, J., Schobel, N., Beltran, L., Wetzel, C. H., & Hatt, H. (2013). Transient receptor potential channels encode volatile chemicals sensed by rat trigeminal ganglion neurons. *PLoS One*, 8(10), e77998, Article e77998. <https://doi.org/10.1371/journal.pone.0077998>
- Ludy, M. J., & Mattes, R. D. (2012). Comparison of sensory, physiological, personality, and cultural attributes in regular spicy food users and non-users. *Appetite*, 58(1), 19-27. <https://doi.org/10.1016/j.appet.2011.09.018>
- Luebbert, M., Radtke, D., Wodarski, R., Damann, N., Hatt, H., & Wetzel, C. H. (2010). Direct activation of transient receptor potential V1 by nickel ions. *Pflügers Archiv-European Journal of Physiology*, 459(5), 737-750. <https://doi.org/10.1007/s00424-009-0782-8>
- Lundahl, D. S., Lukes, B. K., McDaniel, M. R., & Henderson, L. A. (1986). A semi-ascending paired difference method for determining sensory thresholds of added substances to background media. *Journal of Sensory Studies*, 1(3/4), 291-306. <https://doi.org/10.1111/j.1745-459X.1986.tb00179.x>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.
- Lunn, T. E., Nowson, C. A., Worsley, A., & Torres, S. J. (2014). Does personality affect dietary intake? *Nutrition*, 30(4), 403-409. <https://doi.org/10.1016/j.nut.2013.08.012>
- Lyll, V., Heck, G. L., Vinnikova, A. K., Ghosh, S., Phan, T. H., Alam, R. I., Russell, O. F., Malik, S. A., Bigbee, J. W., & DeSimone, J. A. (2004). The mammalian amiloride-insensitive non-specific salt taste receptor is a vanilloid receptor-1 variant. *The Journal of Physiology*, 558(Pt 1), 147-159. <https://doi.org/10.1113/jphysiol.2004.065656>
- Lyu, C., Schijvens, D., Hayes, J. E., & Stieger, M. (2021). Capsaicin burn increases thickness discrimination thresholds independently of chronic chili intake [Article]. *Food Res Int*, 149, 110702, Article 110702. <https://doi.org/10.1016/j.foodres.2021.110702>
- MacLaren, V. V., & Best, L. A. (2009). Female students' disordered eating and the big five personality facets. *Eating Behaviors*, 10(3), 192-195. <https://doi.org/10.1016/j.eatbeh.2009.04.001>
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2021). cluster: Cluster Analysis Basics and Extensions. <https://CRAN.R-project.org/package=cluster>
- Marin, A. B., Barnard, J., Darlington, R. B., & Acree, T. E. (1991). Sensory Thresholds: Estimation from Dose-Response Curves. *Journal of Sensory Studies*, 6(4), 205-225. <https://doi.org/10.1111/j.1745-459X.1991.tb00515.x>
- Marshall, D., & Bell, R. (2004). Relating the food involvement scale to demographic variables, food choice and other constructs. *Food Quality and Preference*, 15(7-8), 871-879. <https://doi.org/10.1016/j.foodqual.2004.06.003>
- Marty, L., de Lauzon-Guillain, B., Labesse, M., & Nicklaus, S. (2021). Food choice motives and the nutritional quality of diet during the COVID-19 lockdown in France. *Appetite*, 157, 105005. <https://doi.org/10.1016/j.appet.2020.105005>
- Matsumoto, K., Kamide, M., Uchida, K., Takahata, M., Shichiri, R., Hida, Y., Taniguchi, Y., Ohishi, A., Tominaga, M., & Nagasawa, K. (2023). Transient Receptor Potential Ankyrin 1 in Taste Nerve Contributes to the Sense of Sweet Taste in Mice. *Biological and Pharmaceutical Bulletin*, 46(7), 939-945. <https://doi.org/10.1248/bpb.b23-00091>
- Matsumoto, K., & Kato, S. (2019). [Physiological and pathophysiological roles of TRPV4 channel in gastrointestinal tract]. *Nihon Yakurigaku Zasshi*, 154(3), 92-96. <https://doi.org/10.1254/fpj.154.92>
- Matsumoto, K., Ohishi, A., Iwatsuki, K., Yamazaki, K., Takayanagi, S., Tsuji, M., Aihara, E., Utsumi, D., Tsukahara, T., Tominaga, M., Nagasawa, K., & Kato, S. (2019). Transient receptor potential vanilloid 4 mediates sour taste sensing via type III taste cell differentiation. *Scientific Reports*,

- 9(1), 6686. <https://doi.org/10.1038/s41598-019-43254-y>
- McClintock, B. T., Langrock, R., Gimenez, O., Cam, E., Borchers, D. L., Glennie, R., & Patterson, T. A. (2020). Uncovering ecological state dynamics with hidden Markov models. *Ecology Letters*, 23(12), 1878-1903. <https://doi.org/10.1111/ele.13610>
- Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, 118, 62-69. <https://doi.org/10.1016/j.chemolab.2012.07.010>
- Mehmood, T., Sæbø, S., & Liland, K. H. (2020). Comparison of variable selection methods in partial least squares regression. *Journal of Chemometrics*, 34(6). <https://doi.org/10.1002/cem.3226>
- Men, H., Jiao, Y. N., Shi, Y., Gong, F. R., Chen, Y. Z., Fang, H. R., & Liu, J. J. (2018). Odor Fingerprint Analysis Using Feature Mining Method Based on Olfactory Sensory Evaluation. *Sensors*, 18(10), 21, Article 3387. <https://doi.org/10.3390/s18103387>
- Mevik, B.-H., & Cederkvist, H. R. (2004). Mean squared error of prediction (MSEP) estimates for principal component regression (PCR) and partial least squares regression (PLSR). *Journal of Chemometrics*, 18(9), 422-429. <https://doi.org/10.1002/cem.887>
- Meyer, D., Zeileis, A., & Hornik, K. (2006). The Strucplot Framework: Visualizing Multi-way Contingency Tables with vcd. *Journal of Statistical Software*, 17(3), 48. <https://doi.org/10.18637/jss.v017.i03>
- 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>
- Miller, L. C., Murphy, R., & Buss, A. H. (1981). Consciousness of Body - Private and Public. *Journal of Personality and Social Psychology*, 41(2), 397-406. <https://doi.org/10.1037/0022-3514.41.2.397>
- 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>
- Moayedi, Y., Michlig, S., Park, M., Koch, A., & Lumpkin, E. A. (2022). Localization of TRP Channels in Healthy Oral Mucosa from Human Donors. *Eneuro*, 9(6), ENEURO.0328-0321.2022. <https://doi.org/10.1523/eneuro.0328-21.2022>
- Mojet, J., Christ-Hazelhof, E., & Heidema, J. (2001). Taste perception with age: generic or specific losses in threshold sensitivity to the five basic tastes? *Chemical Senses*, 26(7), 845-860. <https://doi.org/10.1093/chemse/26.7.845>
- Mojet, J., Heidema, J., & Christ-Hazelhof, E. (2003). Taste perception with age: generic or specific losses in supra-threshold intensities of five taste qualities? *Chemical Senses*, 28(5), 397-413. <https://doi.org/10.1093/chemse/28.5.397>
- Molnar, C., Casalicchio, G., & Bischl, B. (2020). Interpretable Machine Learning – A Brief History, State-of-the-Art and Challenges. In (pp. 417-431). Springer International Publishing. [https://doi.org/10.1007/978-3-030-65965-3\\_28](https://doi.org/10.1007/978-3-030-65965-3_28)
- Mora, M., Urdaneta, E., & Chaya, C. (2019). Effect of personality on the emotional response elicited by wines. *Food Quality and Preference*, 76, 39-46. <https://doi.org/10.1016/j.foodqual.2019.03.015>
- Motter, A. L., & Ahern, G. P. (2012). TRPA1 is a polyunsaturated fatty acid sensor in mammals. *Plos One*, 7(6), e38439, Article e38439. <https://doi.org/10.1371/journal.pone.0038439>
- Murphy, J., Catmur, C., & Bird, G. (2018). Alexithymia is associated with a multidomain, multidimensional failure of interoception: Evidence from novel tests. *Journal of Experimental Psychology: General*, 147(3), 398-408. <https://doi.org/10.1037/xge0000366>
- Nachtigal, D., & Green, B. G. (2018). Receptor specificity of thermal sweet taste and its interaction with chemical taste. *Chemical Senses*, 43(7), E211-E211.
- Nachtigal, D., & Green, B. G. (2020). Sweet Thermal Taste: Perceptual Characteristics in Water and Dependence on TAS1R2/TAS1R3. *Chemical Senses*, 45(3), 219-230. <https://doi.org/10.1093/chemse/bjaa009>

- Nagatomo, K., & Kubol, Y. (2008). Caffeine activates mouse TRPA1 channels but suppresses human TRPA1 channels. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(45), 17373-17378. <https://doi.org/10.1073/pnas.0809769105>
- Nederkoorn, C., Houben, K., & Havermans, R. C. (2019). Taste the texture. The relation between subjective tactile sensitivity, mouthfeel and picky eating in young adults. *Appetite*, *136*, 58-61. <https://doi.org/10.1016/j.appet.2019.01.015>
- Newman, B. H., & Martin, C. A. (2001). The effect of hot beverages, cold beverages, and chewing gum on oral temperature. *Transfusion*, *41*(10), 1241-1243. <https://doi.org/10.1046/j.1537-2995.2001.41101241.x>
- Nezlek, J. B., & Forestell, C. A. (2019). Food neophobia and the Five Factor Model of personality. *Food Quality and Preference*, *73*, 210-214. <https://doi.org/10.1016/j.foodqual.2018.11.007>
- Nezlek, J. B., Forestell, C. A., & Cyprianska, M. (2021). Approach and avoidance motivation and interest in new foods: Introducing a measure of the motivation to eat new foods. *Food Quality and Preference*, *88*, 104111. <https://doi.org/10.1016/j.foodqual.2020.104111>
- Ng, M., Chaya, C., & Hort, J. (2013). Beyond liking: Comparing the measurement of emotional response using EsSense Profile and consumer defined check-all-that-apply methodologies. *Food Quality and Preference*, *28*(1), 193-205. <https://doi.org/10.1016/j.foodqual.2012.08.012>
- Nguyen, Q. C., Næs, T., Almøy, T., & Varela, P. (2020). Portion size selection as related to product and consumer characteristics studied by PLS path modelling. *Food Quality and Preference*, *79*, 10, Article Unsp 103613. <https://doi.org/10.1016/j.foodqual.2018.11.020>
- Nolden, A. A., & Hayes, J. E. (2017). Perceptual and Affective Responses to Sampled Capsaicin Differ by Reported Intake. *Food Quality and Preference*, *55*, 26-34. <https://doi.org/10.1016/j.foodqual.2016.08.003>
- Nystrand, B. T., Olsen, S. O., & Tudoran, A. A. (2021). Individual differences in functional food consumption: The role of time perspective and the Big Five personality traits. *Appetite*, *156*, 104979. <https://doi.org/10.1016/j.appet.2020.104979>
- Olabi, A., Neuhaus, T., Bustos, R., Cook-Camacho, M., Corvi, T., & Abdouni, L. (2015). An investigation of flavor complexity and food neophobia. *Food Quality and Preference*, *42*, 123-129. <https://doi.org/10.1016/j.foodqual.2015.01.004>
- Oleszkiewicz, A., Schultheiss, T., Schriever, V. A., Linke, J., Cuevas, M., Hähner, A., & Hummel, T. (2018). Effects of “trigeminal training” on trigeminal sensitivity and self-rated nasal patency. *European Archives of Oto-Rhino-Laryngology*, *275*(7), 1783-1788. <https://doi.org/10.1007/s00405-018-4993-5>
- Pages, J., & Tenenhaus, M. (2001). Multiple factor analysis combined with PLS path modelling. Application to the analysis of relationships between physicochemical variables, sensory profiles and hedonic judgements. *Chemometrics and Intelligent Laboratory Systems*, *58*(2), 261-273. [https://doi.org/10.1016/s0169-7439\(01\)00165-4](https://doi.org/10.1016/s0169-7439(01)00165-4)
- Pangborn, R. M. (1970). Individual variation in affective responses to taste stimuli. *Psychonomic Science*, *21*(2), 125-126. <https://doi.org/10.3758/bf03335798>
- Pangborn, R. M., Chrisp, R. B., & Bertolero, L. L. (1970). Gustatory, salivary, and oral thermal responses to solutions of sodium chloride at four temperatures. *Perception & Psychophysics*, *8*(2), 69-75. <https://doi.org/10.3758/bf03210177>
- Pangborn, R. M., & Giovanni, M. E. (1984). Dietary intake of sweet foods and of dairy fats and resultant gustatory responses to sugar in lemonade and to fat in milk. *Appetite*, *5*(4), 317-327. [https://doi.org/10.1016/s0195-6663\(84\)80004-5](https://doi.org/10.1016/s0195-6663(84)80004-5)
- Panigutti, C., Perotti, A., & Pedreschi, D. (2020). Doctor XAI: an ontology-based approach to black-box sequential data classification explanations. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*.
- Pankowska, P., Bakker, B., Oberski, D. L., & Pavlopoulos, D. (2018). Reconciliation of inconsistent data sources by correction for measurement error: The feasibility of parameter re-use. *Statistical Journal of the IAOS*, *34*(3), 317-329. <https://doi.org/10.3233/sji-170368>

- Panurywanti, E. E., Wiboworini, B., & Indarto, D. (2021). The effect of banana dose and duration on the decrease of sleep disorders in the elderly. *Journal of Medical & Allied Sciences*, 11(1), 71-76.
- Park, D., Park, Y. K., Park, C. Y., Choi, M.-K., & Shin, M.-J. (2020). Development of a Comprehensive Food Literacy Measurement Tool Integrating the Food System and Sustainability. *Nutrients*, 12(11), 3300. <https://doi.org/10.3390/nu12113300>
- Parma, V., Ohla, K., Veldhuizen, M. G., Niv, M. Y., Kelly, C. E., Bakke, A. J., Cooper, K. W., Bouysset, C., Pirastu, N., Dibattista, M., Kaur, R., Liuzza, M. T., Pepino, M. Y., Schöpf, V., Pereda-Loth, V., Olsson, S. B., Gerkin, R. C., Rohlfs Domínguez, P., Albayay, J., . . . Hayes, J. E. (2020). More Than Smell—COVID-19 Is Associated With Severe Impairment of Smell, Taste, and Chemesthesis. *Chemical Senses*, 45(7), 609-622. <https://doi.org/10.1093/chemse/bjaa041>
- Peng, M., Ginieis, R., Abeywickrema, S., McCormack, J., & Prescott, J. (2023). Rejection thresholds for sweetness reduction in a model drink predict dietary sugar intake. *Food Quality and Preference*, 110, 104965. <https://doi.org/10.1016/j.foodqual.2023.104965>
- Petit, C. E. F., Hollowood, T. A., Wulfert, F., & Hort, J. (2007). Colour–coolant–aroma interactions and the impact of congruency and exposure on flavour perception. *Food Quality and Preference*, 18(6), 880-889. <https://doi.org/https://doi.org/10.1016/j.foodqual.2007.02.003>
- Pfeiler, T. M., & Egloff, B. (2020). Personality and eating habits revisited: Associations between the big five, food choices, and Body Mass Index in a representative Australian sample. *Appetite*, 149, 104607. <https://doi.org/10.1016/j.appet.2020.104607>
- Philipp, M., Rusch, T., Hornik, K., & Strobl, C. (2018). Measuring the Stability of Results from Supervised Statistical Learning. *Journal of Computational and Graphical Statistics*, 27(4), 685-700. <https://doi.org/10.1080/10618600.2018.1473779>
- Philipp, M., Zeileis, A., & Strobl, C. (2016). A Toolkit for Stability Assessment of Tree-Based Learners. 315-325.
- Pickering, G. J. (2024). On being a foodie: Food literacy, involvement, and disgust. *Food Quality and Preference*, 118, 105210. <https://doi.org/https://doi.org/10.1016/j.foodqual.2024.105210>
- Pickering, G. J., Bartolini, J. A., & Bajec, M. R. (2010). Perception of Beer Flavour Associates with Thermal Taster Status. *Journal of the Institute of Brewing*, 116(3), 239-244. <https://doi.org/10.1002/j.2050-0416.2010.tb00426.x>
- Pickering, G. J., & Klodnicki, C. E. (2016). Does Liking and Orosensation Intensity Elicited by Sampled Foods Vary with Thermal Tasting? *Chemosensory Perception*, 9(2), 47-55. <https://doi.org/10.1007/s12078-016-9207-1>
- Pickering, G. J., & Kvas, R. (2016). Thermal Tasting and Difference Thresholds for Prototypical Tastes in Wine. *Chemosensory Perception*, 9(1), 37-46. <https://doi.org/10.1007/s12078-016-9203-5>
- Pickering, G. J., Lucas, S., & Gaudette, N. J. (2016). Variation in orosensation and liking of sampled foods with thermal tasting phenotype. *Flavour*, 5(1), 2. <https://doi.org/10.1186/s13411-016-0049-1>
- Pickering, G. J., Moyes, A., Bajec, M. R., & Decourville, N. (2009). Thermal taster status associates with oral sensations elicited by wine. *Australian Journal of Grape and Wine Research*, 16(2), 361-367. <https://doi.org/10.1111/j.1755-0238.2010.00098.x>
- Pickering, G. J., & Pickering, H. M. G. (2022). On being a foodie: development of the Foodie Index and association with personality and taste phenotype. *Food Quality and Preference*, 96, 104451. <https://doi.org/10.1016/j.foodqual.2021.104451>
- Pierguidi, L., Spinelli, S., Prescott, J., Monteleone, E., & Dinnella, C. (2023). Responsiveness to warning sensations and anxiety-related psychological traits modulate individual differences in preference for vegetable foods with varied sensory properties. *Food Research International*, 173, Article 113342. <https://doi.org/10.1016/j.foodres.2023.113342>
- Pilic, L., Lubasinski, N. J., Berk, M., Ward, D., Graham, C. A.-M., Da Silva Anastacio, V., King, A., & Mavrommatis, Y. (2020). The associations between genetics, salt taste perception and salt intake in young adults. *Food Quality and Preference*, 84.

- <https://doi.org/10.1016/j.foodqual.2020.103954>
- Piochi, M., Cabrino, G., Morini, G., & Torri, L. (2020). Individual differences in the perception of orthonasal irritation induced by food. *Appetite*, *144*, 8, Article Unsp 104460. <https://doi.org/10.1016/j.appet.2019.104460>
- Pliner, P., & Hobden, K. (1992). Development of a scale to measure the trait of food neophobia in humans. *Appetite*, *19*(2), 105-120. [https://doi.org/10.1016/0195-6663\(92\)90014-w](https://doi.org/10.1016/0195-6663(92)90014-w)
- Poelman, M. P., Gillebaart, M., Schlinkert, C., Dijkstra, S. C., Derksen, E., Mensink, F., Hermans, R. C. J., Aardening, P., de Ridder, D., & de Vet, E. (2021). Eating behavior and food purchases during the COVID-19 lockdown: A cross-sectional study among adults in the Netherlands. *Appetite*, *157*, 105002. <https://doi.org/10.1016/j.appet.2020.105002>
- Poole, R. L., & Tordoff, M. G. (2017). The Taste of Caffeine. *Journal of caffeine research*. <https://doi.org/10.1089/jcr.2016.0030>
- Popper, R., Rosenstock, W., Schraidt, M., & Kroll, B. J. (2004). The effect of attribute questions on overall liking ratings. *Food Quality and Preference*, *15*(7-8), 853-858. <https://doi.org/10.1016/j.foodqual.2003.12.004>
- Powers, J. J. (1977). Heterogeneity and Temperature Effects on Flavor Sensing of Ethyl Caprylate. *Journal of Food Science*, *42*(1), 275-276. <https://doi.org/10.1111/j.1365-2621.1977.tb01271.x>
- Prescott, J., Chheang, S. L., & Jaeger, S. R. (2022). Food neophobia: Higher responsiveness to sensory properties but low engagement with foods generally. *Journal of Sensory Studies*. <https://doi.org/10.1111/joss.12771>
- Prescott, J., Laing, D., Bell, G., Yoshida, M., Gillmore, R., Allen, S., Yamazaki, K., & Ishii, R. (1992). Hedonic responses to taste solutions: a cross-cultural study of Japanese and Australians. *Chemical Senses*, *17*(6), 801-809. <https://doi.org/10.1093/chemse/17.6.801>
- Probst, P., Wright, M. N., & Boulesteix, A.-L. (2019). Hyperparameters and tuning strategies for random forest. *WIREs Data Mining and Knowledge Discovery*, *9*(3), e1301. <https://doi.org/10.1002/widm.1301>
- Puputti, S., Aisala, H., Hoppu, U., & Sandell, M. (2019). Factors explaining individual differences in taste sensitivity and taste modality recognition among Finnish adults. *Journal of Sensory Studies*, *34*(4), e12506. <https://doi.org/10.1111/joss.12506>
- Puputti, S., Hoppu, U., & Sandell, M. (2019). Taste Sensitivity Is Associated with Food Consumption Behavior but not with Recalled Pleasantness. *Foods*, *8*(10), 444. <https://doi.org/10.3390/foods8100444>
- R. Core Team. (2021). R: A Language and Environment for Statistical Computing. <https://www.R-project.org/>
- Ramsay, J. O., & Dalzell, C. J. (1991). Some Tools for Functional Data Analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, *53*(3), 539-572. <http://www.jstor.org/stable/2345586>
- Ramsay, J. O., Graves, S., & Hooker, G. (2021). fda: Functional Data Analysis. <https://CRAN.R-project.org/package=fda>
- Ramsay, J. O., Hooker, G., & Graves, S. (2009). *Functional data analysis with R and MATLAB*. Springer.
- 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>
- Reed, D. R. (2008). Birth of a new breed of supertaster. *Chemical Senses*, *33*(6), 489-491. <https://doi.org/10.1093/chemse/bjn031>
- Reis Rocha, L. C., Passos, L. P., Silas Souza, A. H., Amorim, K. A., Rodrigues, J. F., & Pinheiro, A. C. M. (2024). Association of emojis to the five basic tastes (sweet, salty, sour, bitter and umami). *Food Quality and Preference*, *117*, 105177. <https://doi.org/10.1016/j.foodqual.2024.105177>
- Ren, Z., Rhyu, M. R., Phan, T. H., Mummalaneni, S., Murthy, K. S., Grider, J. R., DeSimone, J. A., & Lyall, V. (2013). TRPM5-dependent amiloride- and benzamil-insensitive NaCl chorda tympani taste

- nerve response. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 305(1), G106-117. <https://doi.org/10.1152/ajpgi.00053.2013>
- 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>
- Rhyu, M.-R., Kim, Y., & Lyall, V. (2021). Interactions between Chemesthesis and Taste: Role of TRPA1 and TRPV1. *International Journal of Molecular Sciences*, 22(7), 3360. <https://doi.org/10.3390/ijms22073360>
- Ribeiro, M. N., Carvalho, I. A., Fonseca, G. A., Lago, R. C., Rocha, L. C., Ferreira, D. D., Vilas Boas, E. V., & Pinheiro, A. C. (2021). Quality control of fresh strawberries by a random forest model. *J Sci Food Agric*, 9. <https://doi.org/10.1002/jsfa.11092>
- Riera, C. E., Vogel, H., Simon, S. A., Damak, S., & le Coutre, J. (2008). The capsaicin receptor participates in artificial sweetener aversion. *Biochemical and Biophysical Research Communications*, 376(4), 653-657. <https://doi.org/10.1016/j.bbrc.2008.09.029>
- Riera, C. E., Vogel, H., Simon, S. A., Damak, S., & le Coutre, J. (2009). Sensory attributes of complex tasting divalent salts are mediated by TRPM5 and TRPV1 channels. *The Journal of Neuroscience*, 29(8), 2654-2662. <https://doi.org/10.1523/JNEUROSCI.4694-08.2009>
- Rippon, G. (2019). *The Gendered Brain: The new neuroscience that shatters the myth of the female brain*. Random House.
- Roberts, G., & Spadafora, N. D. (2020). Analysis of Apple Flavours: The Use of Volatile Organic Compounds to Address Cultivar Differences and the Correlation between Consumer Appreciation and Aroma Profiling. *Journal of Food Quality*, 2020, 7, Article 8497259. <https://doi.org/10.1155/2020/8497259>
- Robinson, E., Boyland, E., Chisholm, A., Harrold, J., Maloney, N. G., Marty, L., Mead, B. R., Noonan, R., & Hardman, C. A. (2021). Obesity, eating behavior and physical activity during COVID-19 lockdown: A study of UK adults. *Appetite*, 156, 104853. <https://doi.org/10.1016/j.appet.2020.104853>
- Rocha, R. S., Calvalcanti, R. N., Silva, R., Guimarães, J. T., Balthazar, C. F., Pimentel, T. C., Esmerino, E. A., Freitas, M. Q., Granato, D., Costa, R. G. B., Silva, M. C., & Cruz, A. G. (2020). Consumer acceptance and sensory drivers of liking of Minas Frescal Minas cheese manufactured using milk subjected to ohmic heating: Performance of machine learning methods. *LWT - Food Science and Technology*, 126, 8, Article 109342. <https://doi.org/10.1016/j.lwt.2020.109342>
- Rodrigues, H., Gómez-Corona, C., & Valentin, D. (2020). Femininities & Masculinities: Sex, gender, and stereotypes in food studies. *Current Opinion in Food Science*, 33, 156-164. <https://doi.org/10.1016/j.cofs.2020.05.002>
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7)
- Rypma, B., Prabhakaran, V., Desmond, J. E., & Gabrieli, J. D. E. (2001). Age differences in prefrontal cortical activity in working memory. *Psychology and Aging*, 16(3), 371-384. <https://doi.org/10.1037/0882-7974.16.3.371>
- Sam, C. H. Y. (2012). *The validation of a semi-quantitative, multi-nutrient food frequency questionnaire for assessing selected nutrient intakes in New Zealand adults* [MSc Thesis, University of Otago]. Dunedin, New Zealand. <https://ourarchive.otago.ac.nz/esploro/outputs/graduate/The-validation-of-a-semi-quantitative-multi-nutrient/9926478623701891#file-0>
- Sam, C. H. Y., Skidmore, P., Skeaff, S., Parackal, S., Wall, C., & Bradbury, K. E. (2020). Relative validity and reproducibility of a short food frequency questionnaire to assess nutrient intakes of New Zealand adults. *Nutrients*, 12(3). <https://doi.org/10.3390/nu12030619>
- Sato, H., Wada, H., Matsumoto, H., Takagiwa, M., & Goto, T. K. (2022). Differences in dynamic perception of salty taste intensity between young and older adults. *Scientific Reports*, 12(1), 7558. <https://doi.org/10.1038/s41598-022-11442-y>

- Schnettler, B., Grunert, K. G., Lobos, G., Miranda-Zapata, E., Denegri, M., Lapo, M., Hueche, C., & Rojas, J. (2019). Maternal well-being, food involvement and quality of diet: Profiles of single mother-adolescent dyads. *Children and Youth Services Review*, *96*, 336-345. <https://doi.org/10.1016/j.childyouth.2018.11.020>
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461-464. <http://www.jstor.org/stable/2958889>
- Sclafani, A., & Ackroff, K. (2022). Fat preference deficits and experience-induced recovery in global taste-deficient Trpm5 and Calhm1 knockout mice. *Physiology & Behavior*, *246*, 113695, Article 113695. <https://doi.org/10.1016/j.physbeh.2022.113695>
- Sclafani, A., Zukerman, S., Glendinning, J. I., & Margolskee, R. F. (2007). Fat and carbohydrate preferences in mice: the contribution of  $\alpha$ -gustducin and Trpm5 taste-signaling proteins. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *293*(4), R1504-R1513. <https://doi.org/10.1152/ajpregu.00364.2007>
- Scott, C. L., & Downey, R. G. (2007). Types of food aversions: Animal, vegetable, and texture. *Journal of Psychology*, *141*(2), 127-134. <https://doi.org/10.3200/jrlp.141.2.127-134>
- Scott, N. O., Burgess, B., & Tepper, B. J. (2019). Perception and liking of soups flavored with chipotle chili and ginger extracts: Effects of PROP taster status, personality traits and emotions. *Food Quality and Preference*, *73*, 192-201. <https://doi.org/10.1016/j.foodqual.2018.11.009>
- Scrucca, L. (2013). GA: A Package for Genetic Algorithms in R. *Journal of Statistical Software*, *53*(4), 1 - 37. <https://doi.org/10.18637/jss.v053.i04>
- Señarís, R., Ordás, P., Reimúndez, A., & Viana, F. (2018). Mammalian cold TRP channels: impact on thermoregulation and energy homeostasis. *Pflügers Archiv - European Journal of Physiology*, *470*(5), 761-777. <https://doi.org/10.1007/s00424-018-2145-9>
- Setia, A., Malani, V. D., Goel, V., Mishra, A., & Chan, K. Q. (2022). An exploratory study on the development of a 'foodie' scale: Evidence from a study on students in India. *Food Quality and Preference*, *97*, 104494. <https://doi.org/10.1016/j.foodqual.2021.104494>
- Shahbake, M., Hutchinson, I., Laing, D. G., & Jinks, A. L. (2005). Rapid quantitative assessment of fungiform papillae density in the human tongue. *Brain Research*, *1052*(2), 196-201. <https://doi.org/10.1016/j.brainres.2005.06.031>
- Sijben, R., Hensel, S. M., Rodriguez-Raecke, R., & Freiherr, J. (2021). Gender-Dependent Crossmodal Interactions Between Olfactory and Tactile Stimulation Revealed Using the Unimodal Tactile Stimulation Device (UniTaSD). *Chemical Senses*, *46*. <https://doi.org/10.1093/chemse/bjaa077>
- Skinner, M. (2017). *Exploring individual variation in oral perception* [PhD Thesis, University of Nottingham]. Nottingham eTheses. <https://eprints.nottingham.ac.uk/id/eprint/48736>
- Skinner, M., Eldeghaidy, S., Ford, R., Giesbrecht, T., Thomas, A., Francis, S., & Hort, J. (2018). Variation in thermally induced taste response across thermal tasters. *Physiology & Behavior*, *188*, 67-78. <https://doi.org/10.1016/j.physbeh.2018.01.017>
- Skinner, M., Lim, M., Tarrega, A., Ford, R., Linforth, R., Thomas, A., & Hort, J. (2017). Investigating the oronasal contributions to metallic perception. *International Journal of Food Science & Technology*, *52*(6), 1299-1306. <https://doi.org/10.1111/ijfs.13417>
- Small-Kelly, S. (2018). *Taste responsiveness and beer behaviour* [MSc Thesis, Brock University]. Brock University Digital Repository. St. Catharines, Ontario. <https://dr.library.brocku.ca/handle/10464/13772>
- Small-Kelly, S., & Pickering, G. (2019). Variation in Orosensory Responsiveness to Alcoholic Beverages and Their Constituents—the Role of the Thermal Taste Phenotype. *Chemosensory Perception*, *13*(1), 45-58. <https://doi.org/10.1007/s12078-019-09266-8>
- Smilde, A. K., Måge, I., Naes, T., Hankemeier, T., Lips, M. A., Kiers, H. A. L., Acar, E., & Bro, R. (2017). Common and distinct components in data fusion. *Journal of Chemometrics*, *31*(7), e2900. <https://doi.org/10.1002/cem.2900>
- Smilde, A. K., Naes, T., & Liland, K. H. (2022). *Multiblock Data Fusion in Statistics and Machine Learning: Applications in the Natural and Life Sciences*. John Wiley & Sons.

- Smith, S. D., Nadeau, C., Sorokopud-Jones, M., & Kornelsen, J. (2022). The Relationship Between Functional Connectivity and Interoceptive Sensibility. *Brain Connectivity*, 12(5), 417-431. <https://doi.org/10.1089/brain.2020.0777>
- Solheim, R., & Lawless, H. T. (1996). Consumer purchase probability affected by attitude towards low-fat foods, liking, private body consciousness and information on fat and price. *Food Quality and Preference*, 7(2), 137-143. [https://doi.org/10.1016/0950-3293\(95\)00045-3](https://doi.org/10.1016/0950-3293(95)00045-3)
- Spence, C. (2019). Do men and women really live in different taste worlds? *Food Quality and Preference*, 73, 38-45. <https://doi.org/10.1016/j.foodqual.2018.12.002>
- Spence, C. (2020). Temperature-Based Crossmodal Correspondences: Causes and Consequences. *Multisensory Research*, 33(6), 645-682. <https://doi.org/10.1163/22134808-20191494>
- Spencer, M., & Dalton, P. (2019). The third dimension of flavor: A chemesthetic approach to healthier eating (a review). *Journal of Sensory Studies*, n/a(n/a), e12551. <https://doi.org/10.1111/joss.12551>
- Spinelli, S., De Toffoli, A., Dinnella, C., Laureati, M., Pagliarini, E., Bendini, A., Braghieri, A., Gallina Toschi, T., Sinesio, F., Torri, L., Gasperi, F., Endrizzi, I., Magli, M., Borgogno, M., di Salvo, R., Favotto, S., Prescott, J., & Monteleone, E. (2018). Personality traits and gender influence liking and choice of food pungency. *Food Quality and Preference*, 66, 113-126. <https://doi.org/10.1016/j.foodqual.2018.01.014>
- Spinelli, S., Dinnella, C., Tesini, F., Bendini, A., Braghieri, A., Proserpio, C., Torri, L., Miele, N. A., Aprea, E., Mazzaglia, A., Gallina Toschi, T., & Monteleone, E. (2020). Gender Differences in Fat-Rich Meat Choice: Influence of Personality and Attitudes. *Nutrients*, 12(5), 1374. <https://doi.org/10.3390/nu12051374>
- Spinelli, S., Hopfer, H., Moulinier, V., Prescott, J., Monteleone, E., & Hayes, J. E. (2024). Distinct sensory hedonic functions for sourness in adults. *Food Quality and Preference*, 116, 105152. <https://doi.org/https://doi.org/10.1016/j.foodqual.2024.105152>
- Spinelli, S., Prescott, J., Pierguidi, L., Dinnella, C., Arena, E., Braghieri, A., Di Monaco, R., Toschi, T. G., Endrizzi, I., Proserpio, C., Torri, L., & Monteleone, E. (2021). Phenol-rich food acceptability: The influence of variations in sweetness optima and sensory-liking patterns. *Nutrients*, 13(3), 1-21, Article 866. <https://doi.org/10.3390/nu13030866>
- Sporer, S. L. (2016). Deception and Cognitive Load: Expanding Our Horizon with a Working Memory Model. *Frontiers in Psychology*, 7(420). <https://doi.org/10.3389/fpsyg.2016.00420>
- Staniak, M., & Biecek, P. (2019). Explanations of Model Predictions with live and breakDown Packages. *The R Journal*, 10(2), 395. <https://doi.org/10.32614/rj-2018-072>
- Startek, J. B., Boonen, B., Lopez-Requena, A., Talavera, A., Alpizar, Y. A., Ghosh, D., Van Ranst, N., Nilius, B., Voets, T., & Talavera, K. (2019). Mouse TRPA1 function and membrane localization are modulated by direct interactions with cholesterol. *eLife*, 8, e46084. <https://doi.org/10.7554/eLife.46084>
- Stats NZ. (2018). 2018 Census place summaries - Palmerston North. Retrieved 4 May 2024, from <https://www.stats.govt.nz/tools/2018-census-place-summaries/palmerston-north-city#ethnicity-culture-and-identity>
- Stats NZ. (2019). Gender identity V1.0.0. In.
- Stats NZ. (2020). Retrieved 17 May 2024, from <https://www.stats.govt.nz/tools/2018-census-ethnic-group-summaries>
- Story, G. M., Peier, A. M., Reeve, A. J., Eid, S. R., Mosbacher, J., Hricik, T. R., Earley, T. J., Hergarden, A. C., Andersson, D. A., Hwang, S. W., McIntyre, P., Jegla, T., Bevan, S., & Patapoutian, A. (2003). ANKTM1, a TRP-like Channel Expressed in Nociceptive Neurons, Is Activated by Cold Temperatures. *Cell*, 112(6), 819-829. [https://doi.org/10.1016/s0092-8674\(03\)00158-2](https://doi.org/10.1016/s0092-8674(03)00158-2)
- Strani, L., Vitale, R., Tanzilli, D., Bonacini, F., Perolo, A., Mantovani, E., Ferrando, A., & Cocchi, M. (2022). A Multiblock Approach to Fuse Process and Near-Infrared Sensors for On-Line Prediction of Polymer Properties. *Sensors*, 22(4), 1436. <https://doi.org/10.3390/s22041436>
- Streiner, D. L. (2015). Best (but oft-forgotten) practices: the multiple problems of multiplicity—

- whether and how to correct for many statistical tests<sup>1</sup>. *The American Journal of Clinical Nutrition*, 102(4), 721-728. <https://doi.org/https://doi.org/10.3945/ajcn.115.113548>
- Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., & Zeileis, A. (2008). Conditional Variable Importance for Random Forests. *BMC Bioinformatics*, 9(307). <https://doi.org/10.1186/1471-2105-9-307>
- Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8(1), 25. <https://doi.org/10.1186/1471-2105-8-25>
- Studer, M., & Ritschard, G. (2016). What matters in differences between life trajectories: a comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(2), 481-511. <https://doi.org/10.1111/rssa.12125>
- Su, T., Gao, X., Li, H., Zhang, L., Han, P., & Chen, H. (2022). Frequent spicy food consumption is associated with reduced capsaicin and salty taste sensitivity but unchanged sour taste or intranasal trigeminal sensitivity. *Food Quality and Preference*, 96, 104411. <https://doi.org/10.1016/j.foodqual.2021.104411>
- Suzuki, M., Mizuno, A., Kodaira, K., & Imai, M. (2003). Impaired pressure sensation in mice lacking TRPV4. *Journal of Biological Chemistry*, 278(25), 22664-22668. <https://doi.org/10.1074/jbc.M302561200>
- Takahashi, S., Kurogi, M., & Saitoh, O. (2021). The diversity in sensitivity of TRPA1 and TRPV1 of various animals to polyphenols. *Biomedical Research-Tokyo*, 42(2), 43-+. <https://doi.org/10.2220/biomedres.42.43>
- Talavera, K., Ninomiya, Y., Winkel, C., Voets, T., & Nilius, B. (2007). Influence of temperature on taste perception. *Cellular and Molecular Life Sciences*, 64(4), 377-381. <https://doi.org/10.1007/s00018-006-6384-0>
- Talavera, K., Yasumatsu, K., Voets, T., Droogmans, G., Shigemura, N., Ninomiya, Y., Margolskee, R. F., & Nilius, B. (2005). Heat activation of TRPM5 underlies thermal sensitivity of sweet taste. *Nature*, 438(7070), 1022-1025. <https://doi.org/10.1038/nature04248>
- Talavera, K., Yasumatsu, K., Yoshida, R., Margolskee, R. F., Voets, T., Ninomiya, Y., & Nilius, B. (2008). The taste transduction channel TRPM5 is a locus for bitter-sweet taste interactions. *The FASEB Journal*, 22(5), 1343-1355. <https://doi.org/10.1096/fj.07-9591com>
- 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>
- Tanaka, A., Mochizuki, T., Ishibashi, T., Akamizu, T., Matsuoka, T.-a., & Nishi, M. (2022). Reduced Fat Taste Sensitivity in Obese Japanese Patients and Its Recovery after a Short-Term Weight Loss Program. *Journal of Nutritional Science and Vitaminology*, 68(6), 504-512. <https://doi.org/10.3177/jnsv.68.504>
- Tapanee, P., Tidwell, D. K., Schilling, M. W., Peterson, D. G., & Tolar-Peterson, T. (2021). Genetic Variation in Taste Receptor Genes (SCNN1B, TRPV1) and Its Correlation with the Perception of Saltiness in Normotensive and Hypertensive Adults. *International Journal of Hypertension*, 2021, 7, Article 5559831. <https://doi.org/10.1155/2021/5559831>
- Tay, E., Barnett, D., Leilua, E., Kerse, N., Rowland, M., Rolleston, A., Waters, D. L., Edlin, R., Connolly, M., Hale, L., Pillai, A., & Teh, R. (2021). The Diet Quality and Nutrition Inadequacy of Pre-Frail Older Adults in New Zealand. *Nutrients*, 13(7), 2384. <https://doi.org/10.3390/nu13072384>
- Thibodeau, M., Bajec, M., Saliba, A., & Pickering, G. (2020). Homogeneity of thermal tasters and implications for mechanisms and classification. *Physiology & Behavior*, 227, 113160. <https://doi.org/10.1016/j.physbeh.2020.113160>
- Thibodeau, M., Ford, R., Pickering, G., & Yang, Q. (2023). Thermal tasters and thermal non-tasters differ in food disgust. Pangborn, Dijon, France.
- Thibodeau, M., & Pickering, G. (2021). Perception of Aqueous Ethanol Binary Mixtures Containing Alcohol-Relevant Taste and Chemesthetic Stimuli. *Beverages*, 7(2), 23. <https://doi.org/10.3390/beverages7020023>

- Thibodeau, M., & Pickering, G. J. (2019). The role of taste in alcohol preference, consumption and risk behavior. *Critical reviews in food science and nutrition*, 59(4), 676-692. <https://doi.org/10.1080/10408398.2017.1387759>
- Thibodeau, M., Saliba, A., Bajec, M., & Pickering, G. (2019). Examination and Validation of Classification Schema for Determining Thermal Taste Status. *Chemosensory Perception*, 12(2), 69-89. <https://doi.org/10.1007/s12078-019-09264-w>
- Thibodeau, P. H. (2016). A moist crevice for word aversion: In semantics not sounds. *Plos One*, 11(4), Article e0153686. <https://doi.org/10.1371/journal.pone.0153686>
- Tian, L., Zhang, H., Yang, S., Luo, A., Kamau, P. M., Hu, J., Luo, L., & Lai, R. (2023). Vertebrate OTOP1 is also an alkali-activated channel. *Nature Communications*, 14(1). <https://doi.org/10.1038/s41467-022-35754-9>
- Tornwall, O., Silventoinen, K., Hiekkalinna, T., Perola, M., Tuorila, H., & Kaprio, J. (2014). Identifying flavor preference subgroups. Genetic basis and related eating behavior traits. *Appetite*, 75, 1-10. <https://doi.org/10.1016/j.appet.2013.11.020>
- Touw, W. G., Bayjanov, J. R., Overmars, L., Backus, L., Boekhorst, J., Wels, M., & van Hijum, S. A. F. T. (2012). Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle? *Briefings in Bioinformatics*, 14(3), 315-326. <https://doi.org/10.1093/bib/bbs034>
- Tu, Y.-H., Liu, N., Xiao, C., Gavrilova, O., & Reitman, M. L. (2023). Loss of Otopetrin 1 affects thermoregulation during fasting in mice. *Plos One*, 18(10), e0292610. <https://doi.org/10.1371/journal.pone.0292610>
- Tucker, R. M., & Mattes, R. D. (2013). Influences of Repeated Testing on Nonesterified Fatty Acid Taste. *Chemical Senses*, 38(4), 325-332. <https://doi.org/10.1093/chemse/bjt002>
- Turner, H. N., & Liman, E. R. (2022). The Cellular and Molecular Basis of Sour Taste. *Annual Review of Physiology*, 84, 41-58. <https://doi.org/10.1146/annurev-physiol-060121-041637>
- Ullah, A., Liu, Y., Wang, Y., Gao, H., Wang, H., Zhang, J., & Li, G. (2022). E-Taste: Taste Sensations and Flavors Based on Tongue's Electrical and Thermal Stimulation. *Sensors*, 22(13), 4976. <https://doi.org/10.3390/s22134976>
- van der Wal, R. C., & van Dillen, L. F. (2013). Leaving a Flat Taste in Your Mouth: Task Load Reduces Taste Perception. *Psychological Science*, 24(7), 1277-1284. <https://doi.org/10.1177/0956797612471953>
- van der Werf, S. P., de Vree, B., van der Meer, J. W. M., & Bleijenberg, G. (2002). The relations among body consciousness, somatic symptom report, and information processing speed in chronic fatigue syndrome. *Neuropsychiatry Neuropsychology and Behavioral Neurology*, 15(1), 2-9.
- van Opstal, A. M., van den Berg-Huysmans, A. A., Hoeksma, M., Blonk, C., Pijl, H., Rombouts, S., & van der Grond, J. (2018). The effect of consumption temperature on the homeostatic and hedonic responses to glucose ingestion in the hypothalamus and the reward system [Article]. *Am J Clin Nutr*, 107(1), 20-25. <https://doi.org/10.1093/ajcn/nqx023>
- Velázquez, A. L., Vidal, L., Varela, P., & Ares, G. (2020a). 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>
- Velázquez, A. L., Vidal, L., Varela, P., & Ares, G. (2020b). Cross-modal interactions as a strategy for sugar reduction in products targeted at children: Case study with vanilla milk desserts. *Food Research International*, 130, 108920. <https://doi.org/10.1016/j.foodres.2019.108920>
- Vidal, L., Ares, G., & Jaeger, S. R. (2021). Differences in citation proportions in CATA questions can be interpreted as differences perceived intensity of sensory attributes. *Journal of Sensory Studies*, 36(5), e12695. <https://doi.org/10.1111/joss.12695>
- Vig, L., Koteles, F., & Ferentzi, E. (2022). Questionnaires of interoception do not assess the same construct. *Plos One*, 17(8 August), Article e0273299. <https://doi.org/10.1371/journal.pone.0273299>
- Vigneau, E., Courcoux, P., Symoneaux, R., Guérin, L., & Villière, A. (2018). Random forests: A machine learning methodology to highlight the volatile organic compounds involved in olfactory

- perception. *Food Quality and Preference*, 68, 135-145. <https://doi.org/10.1016/j.foodqual.2018.02.008>
- 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/https://doi.org/10.1016/j.foodqual.2023.104942>
- Voets, T., Owsianik, G., Janssens, A., Talavera, K., & Nilius, B. (2007). TRPM8 voltage sensor mutants reveal a mechanism for integrating thermal and chemical stimuli. *Nature Chemical Biology*, 3(3), 174-182. <https://doi.org/10.1038/nchembio862>
- von Molitor, E., Riedel, K., Krohn, M., Rudolf, R., Hafner, M., & Cesetti, T. (2020). An alternative pathway for sweet sensation: possible mechanisms and physiological relevance. *Pflugers Archiv-European Journal of Physiology*, 472(12), 1667-1691. <https://doi.org/10.1007/s00424-020-02467-1>
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, 17(2), 228-243. <https://doi.org/10.1037/a0027127>
- Wang, Y., Zhong, K., Shi, B., Wang, H., Liu, L., Zhang, L., Zhao, L., & Gao, H. (2022). Cross-modal effect of capsaicin and pepper oleoresin on the enhancement of saltiness perception in a NaCl model solution. *Food Quality and Preference*, 98, 104542. <https://doi.org/10.1016/j.foodqual.2022.104542>
- Wang, Y. Y., Chang, R. B., Allgood, S. D., Silver, W. L., & Liman, E. R. (2011). A TRPA1-dependent mechanism for the pungent sensation of weak acids. *The Journal of General Physiology*, 137(6), 493-505. <https://doi.org/10.1085/jgp.201110615>
- Wang, Y. Y., Chang, R. B., & Liman, E. R. (2010). TRPA1 is a component of the nociceptive response to CO<sub>2</sub>. *The Journal of Neuroscience*, 30(39), 12958-12963. <https://doi.org/10.1523/JNEUROSCI.2715-10.2010>
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician*, 70(2), 129-133. <https://doi.org/10.1080/00031305.2016.1154108>
- Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a World Beyond “p<0.05”. *The American Statistician*, 73(sup1), 1-19. <https://doi.org/10.1080/00031305.2019.1583913>
- 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>
- Wendin, K., Janestad, H., & Hall, G. (2003). Modelling and analysis of dynamic sensory data. *Food Quality and Preference*, 14(8), 663-671. [https://doi.org/10.1016/s0950-3293\(02\)00208-2](https://doi.org/10.1016/s0950-3293(02)00208-2)
- Weyer-Menkoff, I., Thrun, M. C., & Lotsch, J. (2018). Machine-learned analysis of quantitative sensory testing responses to noxious cold stimulation in healthy subjects. *European Journal of Pain*, 22(5), 862-874. <https://doi.org/10.1002/ejp.1173>
- WHO. (2023). Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic. Retrieved 6 May 2024, from [https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-coronavirus-disease-\(covid-19\)-pandemic](https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic)
- Wickham, H. (2016). *Ggplot2 : elegant graphics for data analysis* (Second edition ed.). Springer. <https://ggplot2.tidyverse.org>
- Wickham, H. (2017). *tidyverse: Easily Install and Load the 'Tidyverse'* <https://CRAN.R-project.org/package=tidyverse>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., . . . Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open*

- Source Software, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Wilson, A., Jeltama, M., Morgenstern, M. P., Motoi, L., Kim, E., & Hedderley, D. (2018). Comparison of physical chewing measures to consumer typed Mouth Behavior. *Journal of Texture Studies*, 49(3), 262-273. <https://doi.org/10.1111/jtxs.12328>
- Wilton, M., Stancak, A., Giesbrecht, T., Thomas, A., & Kirkham, T. (2019). Intensity expectation modifies gustatory evoked potentials to sweet taste: Evidence of bidirectional assimilation in early perceptual processing. *Psychophysiology*, 56(3), e13299, Article UNSP e13299. <https://doi.org/10.1111/psyp.13299>
- Wold, S., Johansson, E., & Cocchi, M. (1993). PLS: partial least squares projections to latent structures. In *3D QSAR in Drug Design: Theory, Methods and Applications*. (pp. 523-550). Kluwer ESCOM Science Publisher.
- Yang, K., Li, Y., Mao, Z., Liu, X., Zhang, H., Liu, R., Xue, Y., Tu, R., Liu, X., Zhang, X., Li, W., & Wang, C. (2018). Relationship between spicy flavor, spicy food intake frequency, and general obesity in a rural adult Chinese population: The RuralDiab study. *Nutrition, Metabolism and Cardiovascular Diseases*, 28(3), 252-261. <https://doi.org/10.1016/j.numecd.2017.10.021>
- Yang, Q. (2015). *Individual variation across PROP and Thermal Taste phenotypes* [PhD Thesis, University of Nottingham]. Nottingham eTheses.
- Yang, Q., Dorado, R., Chaya, C., & Hort, J. (2018). The impact of PROP and thermal taster status on the emotional response to beer. *Food Quality and Preference*, 68, 420-430. <https://doi.org/10.1016/j.foodqual.2018.03.001>
- Yang, Q., Hollowood, T., & Hort, J. (2014). Phenotypic variation in oronasal perception and the relative effects of PROP and Thermal Taster Status. *Food Quality and Preference*, 38, 83-91. <https://doi.org/10.1016/j.foodqual.2014.05.013>
- Yang, Q., Kraft, M., Shen, Y., MacFie, H., & Ford, R. (2019). Sweet Liking Status and PROP Taster Status impact emotional response to sweetened beverage. *Food Quality and Preference*, 75, 133-144. <https://doi.org/10.1016/j.foodqual.2019.02.016>
- Yang, Q., Williamson, A.-M., Hasted, A., & Hort, J. (2020). Exploring the relationships between taste phenotypes, genotypes, ethnicity, gender and taste perception using Chi-square and regression tree analysis. *Food Quality and Preference*, 83, 103928. <https://doi.org/10.1016/j.foodqual.2020.103928>
- Yang, T., Zhang, L., Xu, G., Yang, Z., Luo, Y., Li, Z., Zhong, K., Shi, B., Zhao, L., & Sun, P. (2022). Investigating taste sensitivity, chemesthetic sensation and their relationship with emotion perception in Chinese young and older adults. *Food Quality and Preference*, 96, 104406. <https://doi.org/10.1016/j.foodqual.2021.104406>
- Yokoo, T., Sirovich, L., & Knight, B. W. (2001). An optimization approach to signal extraction from noisy multivariate data. *Neuroimage*, 14(6), 1309-1326-1326. <https://doi.org/10.1006/nimg.2001.0950>
- Zeileis, A., Meyer, D., & Hornik, K. (2007). Residual-Based Shadings for Visualizing (Conditional) Independence. *Journal of Computational and Graphical Statistics*, 16(3), 507-525. <https://doi.org/10.1198/106186007x237856>
- Zeileis, A., Meyer, D., & Hornik, K. (2017). *vcd: Visualizing Categorical Data*. In
- Zhang, J., Jin, H., Zhang, W. Y., Ding, C., O'Keefe, S., Ye, M. Y., & Zuker, C. S. (2019). Sour Sensing from the Tongue to the Brain. *Cell*, 179(2), 392-+. <https://doi.org/10.1016/j.cell.2019.08.031>
- Zhang, Y., Hoon, M. A., Chandrashekar, J., Mueller, K. L., Cook, B., Wu, D., Zuker, C. S., & Ryba, N. J. (2003). Coding of sweet, bitter, and umami tastes: different receptor cells sharing similar signaling pathways. *Cell*, 112(3), 293-301. [https://doi.org/10.1016/s0092-8674\(03\)00071-0](https://doi.org/10.1016/s0092-8674(03)00071-0)
- Zhou, L., Ma, Y., Chen, H., & Han, P. (2023). Sex-specific association between regional gray matter volume and spicy food craving or consumption. *Appetite*, 190, 107038. <https://doi.org/10.1016/j.appet.2023.107038>
- Zhou, X., Yeomans, M., Thomas, A., Wilde, P., Linter, B., & Methven, L. (2021). Individual differences in oral tactile sensitivity and gustatory fatty acid sensitivity and their relationship with

fungiform papillae density, mouth behaviour and texture perception of a food model varying in fat. *Food Quality and Preference*, 90, 104116. <https://doi.org/10.1016/j.foodqual.2020.104116>

Zhou, X. R., Shen, Y. C., Parker, J. K., Kennedy, O. B., & Methven, L. (2016). Relative Effects of Sensory Modalities and Importance of Fatty Acid Sensitivity on Fat Perception in a Real Food Model. *Chemosensory Perception*, 9(3), 105-119. <https://doi.org/10.1007/s12078-016-9211-5>

Zimmermann, K., Lennerz, J. K., Hein, A., Link, A. S., Kaczmarek, J. S., Delling, M., Uysal, S., Pfeifer, J. D., Riccio, A., & Clapham, D. E. (2011). Transient receptor potential cation channel, subfamily C, member 5 (TRPC5) is a cold-transducer in the peripheral nervous system. *Proceedings of the National Academy of Sciences*, 108(44), 18114-18119. <https://doi.org/10.1073/pnas.1115387108>

## Appendix A. Ethics approval SOA 19/26

### SOA 19/26 Thermal taster phenotyping: Protocol validation and participant phenotyping

Janita Botha (HEC: Southern A Application SOA 19/26)

Department: School of Food & Advanced Technologies

Supervisor: Dr Peter Cannon; Prof Joanne Hort

The application was provisionally approved, subject to the fulfilment of the conditions below to the satisfaction of Dr Negar Partow (Chair).

Please note that the Committee is always willing to enter into dialogue with applicants over the points made. There may be information that has not been made available to the Committee, or aspects of the research may not have been fully understood.

#### SECTION C

##### C1

- How has the Thermode been tested and verified as safe for use with research participants?

#### C3/INFORMATION SHEET

- C3 mentions 2 protocols/procedures to be tested; however, the information sheet states 5 protocols (Project procedures, bullet 3). Please review the details and ensure information is consistent in the application and public documentation.

#### SECTION D

##### D4

- Is an email list to be used in the process of recruitment? If so, provide further details, e.g. how are potential participants emails obtained; is it through an organisational mailing list; is permission required to access (or is an advertisement sent on behalf of the researcher)?

#### FLYER/INFORMATION SHEET

- The flyer notes participants will be compensated for their time; however, no further detail is provided in the application form or the information sheet. Please provide further detail in the documentation.

#### INFORMATION SHEET

- Refer to C3 and flyer/information sheet points above – include details regarding reimbursement for time.
- Data management, bullet 4 – this point notes participants will be provided with a link to access summary findings. Suggest an indication of when the findings are likely to be available would also be beneficial.
- Please include the correct committee approval statement, as follows: *This project has been reviewed and approved by the Massey University Human Ethics Committee: Southern A, Application 19/26. If you have any concerns about the conduct of this research, please contact Dr Negar Partow, Chair, Massey University Human Ethics Committee: Southern A, telephone 04 801 5799 x 63363, email [humanethicsoutha@massey.ac.nz](mailto:humanethicsoutha@massey.ac.nz).*
- Please upload a revised information sheet.

## Appendix B. Ethics approval SOA 20/40

### SOA 20/40 Investigating relationships between Thermal Taster Status and other phenotypes and behavioural measures

Janita Botha (HEC: Southern A Application SOA 20/40)

Department: School of Food & Advanced Technologies

Supervisor: Prof Joanne Hort; Dr Peter Cannon

The application was provisionally approved, subject to the fulfilment of the conditions below to the satisfaction of Dr Negar Partow (Chair).

Please note that the Committee is always willing to enter into dialogue with applicants over the points made. There may be information that has not been made available to the Committee, or aspects of the research may not have been fully understood.

- Note: Please ensure adherence to current national COVID-19 restrictions and the latest guidelines from Massey University regarding research during this time. Note: This may necessitate the need for researchers to consider amendments to the data collection methods that are currently proposed or delaying the start of the research.
- The committee noted that some participants will be approached based on their participation in earlier research. Please clarify the breadth of their earlier consent to ensure that this further research falls within the terms of the original consent, i.e. was consent given by participants to be approached for future research.
- Is there a health screening questionnaire? The committee noted that allergy to dairy and Graves disease are mentioned in the flow chart. Therefore, there should be a health screening questionnaire that includes these (and any other criteria that would exclude a participant from the study).

#### SECTION D

##### D2

- Please provide a power calculation to justify the number of participants being recruited to the study.

##### D7/E8

- Please outline what Massey University approvals are required where information is sent to departmental lists via their administrators, e.g. HoD?

#### D8/ADVERTISEMENT/INFORMATION SHEETS

- Please clarify the reason for putting the upper age limit of 65 on participants. Note: This upper age limit should also be clear in the advertisement and information sheet.

##### D10

- Please clarify the safety management plan in place for the taste testing.

#### SECTION G

##### G3a

- If more than one participant is attending the Riddet Complex to take taste tests, what arrangements are in place for safe social distancing (if the Covid-19 Alert Level requires appropriate distancing)?
- If participants are attending in the company of others, what steps are envisaged to protect the privacy of participants in relation to one another?

- Also, what precautions are envisaged for the researcher? A face-shield will be considered (should this be mandatory) but what of a mask, gloves, distancing to the extent feasible?

#### FLYER

- Ensure inclusion/exclusion criteria are clear.
- Please include the correct committee approval statement, as follows: *This project has been reviewed and approved by the Massey University Human Ethics Committee: Southern A, Application 20/40. If you have any concerns about the conduct of this research, please contact Dr Negar Partow, Chair, Massey University Human Ethics Committee: Southern A, telephone 04 801 5799 x 63363, email humanethicsoutha@massey.ac.nz.*
- Upload revised flyer.

#### INFORMATION SHEETS

- The first two state the 150 participants will take part in the research; however, the third states 225. Please ensure consistency or clarify the difference if the numbers in each information sheet are correct.
- The first two Information Sheets (intended for those who have participated in earlier more limited studies) are confusing as they run through elements of the whole study and then say the reader will not need to do the first and second etc. It may be more comprehensible if the sheets for these first explained that the study would, with their approval, draw on the results from their earlier participation in TT tests, but also ask them to now undertake certain additional tests/questionnaires. The combined results would then be joined by results from others recruited for the first time who would complete matching tests.
- The consent to participate in further studies seems very broad. It should be narrowed to studies of a similar kind e.g. relating to tasting phenotypes.
- The committee noted the information sheets are light on detail about the nature and purpose of the research (to determine the participants thermal taster status by investigating the response to temperature stimulation of the tongue). However, the research seeks detail relating to different tasting phenotypes and any relationship (at least in relation to TTS?) with psychological and behavioural variables. Therefore, the information should provide greater clarity about the nature and extent of the research.
- Provide further detail about the nature and purpose of completing the online questionnaires.
- “Data Management” - participants are told “the data obtained during this study will be used to understand if taster status phenotypes affect food choice behaviours”. Could this be included earlier on and made clearer how the nature of the tests undertaken may help toward such an understanding?
- “Participant’s rights” – include a timeframe for withdrawal from the study.
- Please include the correct committee approval statement, as follows: *This project has been reviewed and approved by the Massey University Human Ethics Committee: Southern A, Application 20/40. If you have any concerns about the conduct of this research, please contact Dr Negar Partow, Chair, Massey University Human Ethics Committee: Southern A, telephone 04 801 5799 x 63363, email humanethicsoutha@massey.ac.nz.*
- Please upload revised information sheets.

## Appendix C. RapCoTT method design documentation

Note: Figures for which the author does not own the copyright have been redacted

### Thermode practical planning – Version 3

12 April 2019

For “Individual variation in Taste perception: Investigating sensory perception in Thermal Tasters and its impact on food choice behaviour”

Janita J Botha

#### 1 Context

This document serves as a guide to capture the future steps/milestones for the building/creation of the thermode. This is a working document which outlines the outcomes in three main areas:

- 1) Functional programming,
- 2) Hardware
- 3) Software.

#### 2 Programming/protocols

The different programs are required for the validation of the new protocol, since we would like to compare protocols in order to validate the new/updated protocol. The older protocols use two phases (a heating and a cooling phase) and generally start at 35°C and try to return the tongue to 35°C. All protocols use a heating/cooling rate of 1°C per second.

Note that the thermode needs to be able to run all the protocols described, but these will not necessarily form part of the final protocol required for the validation.

##### 2.1 Protocol 1 (Green)

In the original paper by Cruz and Green (2000), a very simple protocol is used, which simply involved heating or cooling the tongue in one direction. However, this protocol was refined for future work (Green, Alvarez-Reeves, George, & Akirav, 2005; Green & George, 2004). In the updated protocol, the warming trial starts at 35°C, cools the tongue to 15°C and then heats the tongue to 35°C. The cooling trial also starts at 35°C, cools the tongue to 15°C and then keeps the tongue at this temperature for 10 seconds. These protocols are summarised in Figure 1.

Note that this protocol will not be used during the validation/phenotyping process.

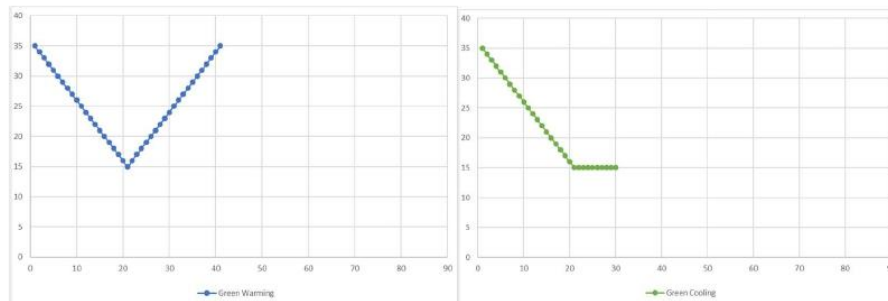


Figure 1 A visual representation of the Green protocols a) the warming trial b) the cooling trial

## 2.2 Protocol 2 (Pickering)

Bajec and Pickering (2008) updated the previous protocols by increasing the maximum temperature of the warming trial to 40 °C. They also modified the cooling trial by reducing the minimum temperature to 5°C. This is summarised in Figure 2. Verbally, the warming trial starts at 35°C, cooling the tongue to 15°C, and heating the tongue to 40°C. The cooling trial starts at 35° and cools the tongue to 5 °, where it is held for 10 seconds. This protocol was used in several other publications by the same research group (Pickering & Klodnicki, 2016; Pickering & Kvas, 2016; Pickering, Lucas, & Gaudette, 2016).

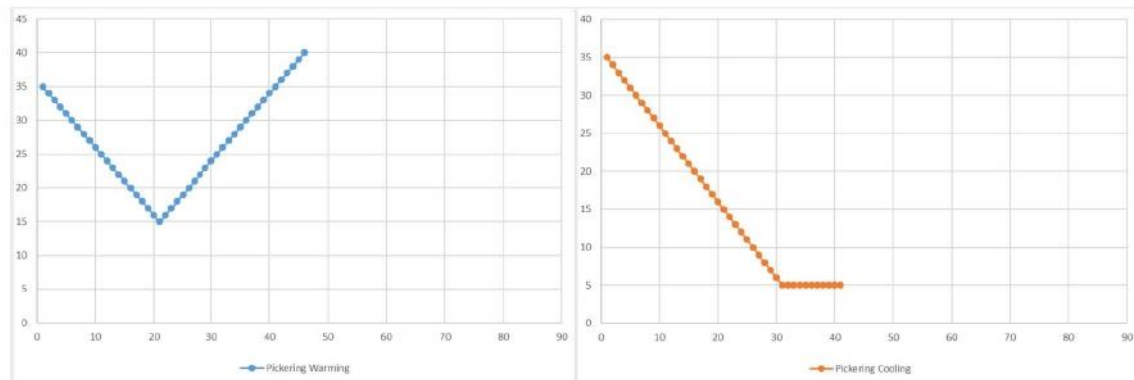


Figure 2 A visual representation of the Pickering protocols a) the warming trial b) the cooling trial

## 2.3 Protocol 3 (Skinner)

The protocol described by Skinner et al. (2018) is an adaptation of the previous protocols as it includes the return of the tongue to 35°C after both trials. These protocols can be seen in Figure 3 . In full, the “warming” trial start at 35°C, cools the tongue to 15°C, warms the tongue to 40°C and returns it to 35°C. The “cooling” trial starts at 35°C, cools the tongue to 5°C and re-warms it to 35°C. Skinner proposed that future protocols could include both the warming and the cooling phases as part of one extended trial.

Note that this protocol will not be used during the validation/phenotyping process.

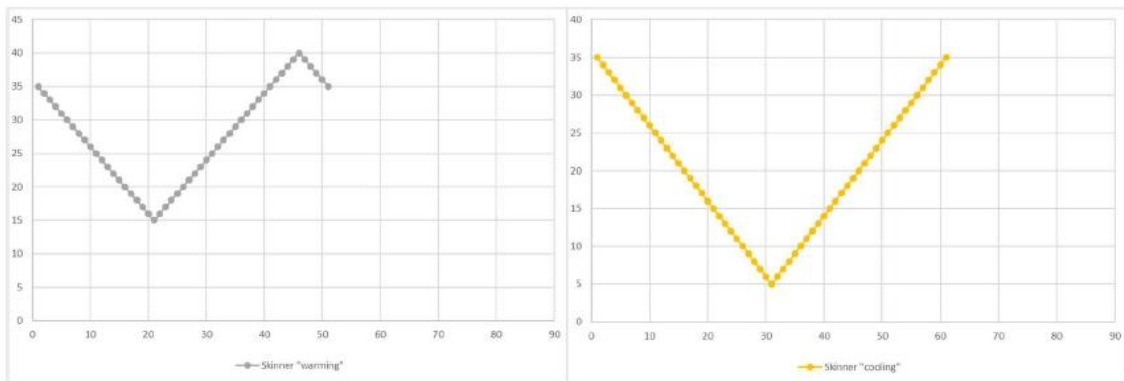


Figure 3 A visual representation of the Skinner protocols a) the warming trial b) the cooling trial

## 2.4 Protocol 4 (Botha)

Following from Skinner et al. (2018), the protocol for this project should satisfy the following conditions:

- 1) Start at 35°C
- 2) Cover a range of at least 5°C to 40°C
- 3) Have both the cooling and warming phases included in the protocol

The simplest way to achieve this is as illustrated in Figure 4.

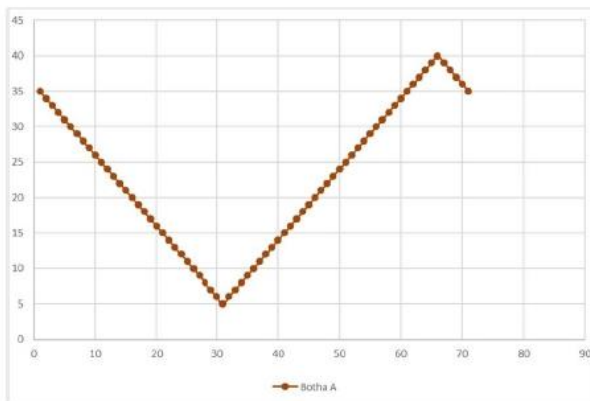


Figure 4 Graphical representation of proposed protocol

## 2.5 Summary of protocols

The three protocols are summarised in Table 1

Table 1 Summary of Themode protocols

Author	Known as	Warming	Cooling	Used in validation
Green	Protocol 1	35 -> 15 -> 35	35 -> 15 (10 second hold)	No
Pickering	Protocol 2	35 -> 15 -> 40	35 -> 5 (10 second hold)	Yes
Skinner	Protocol 3	35 -> 15 -> 40 -> 35	35 -> 5 -> 35	No
Proposed	Protocol 4	35 -> 5 -> 40 -> 35		Yes

### 3 Hardware

The function of hardware on the thermode is threefold:

- 1) To protect the circuitry from accidental damage
- 2) To allow for correct functioning of thermode
- 3) To allow for ease of interaction with thermode
- 4) To ensure safety of participants (protections from electric shock)

The first two items are somewhat self-explanatory, and will not be detailed here. However, the requirements for interface hardware will be explored.

#### 3.1 Interface hardware

There are two examples from literature as to how the human interface to the thermode might look. Firstly, Yang (2015) shows an image which shows the Peltier device as hand-held and as held with a nonconductive (plastic) handle – which can be seen in Figure 5. Skinner et al. (2018) shows a different setup in which the thermode is held in a plastic casing, which is inserted into the mouth and has a space for the tongue to be held, as can be seen in Figure 6.



*Figure 5* The Peltier device as described by Yang a) the thermode probe b) The complete evaluation system c) The thermode in use (showing hand-held setup)

Either of these approaches would be suitable. What is, however important is that the participants should be comfortable and should be able to hold (either orally or by hand) the thermode in a manner which disallows thermal conductivity by hand.

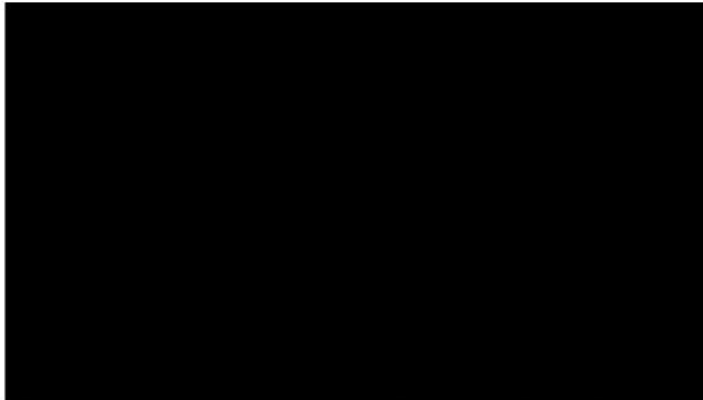


Figure 6 The Thermode as described by Skinner a) The Thermode in use b) The plastic mouth-piece

Figure 7 show a basic diagram of the final design used for experimentation, design, as well as a photograph of the device. .In short, the device consists of a small peltier device connected to a thermocouple probe. This is mounted on an aluminium heat-sink, which is covered by a non-conductive handle. The device also connected to a microchip (Arduino device) and runs off software specifically designed for this purpose, and is powered using a 9V battery. This will allow participants to comfortably hold the thermode up to their tongue without experiencing the temperature change digitally.

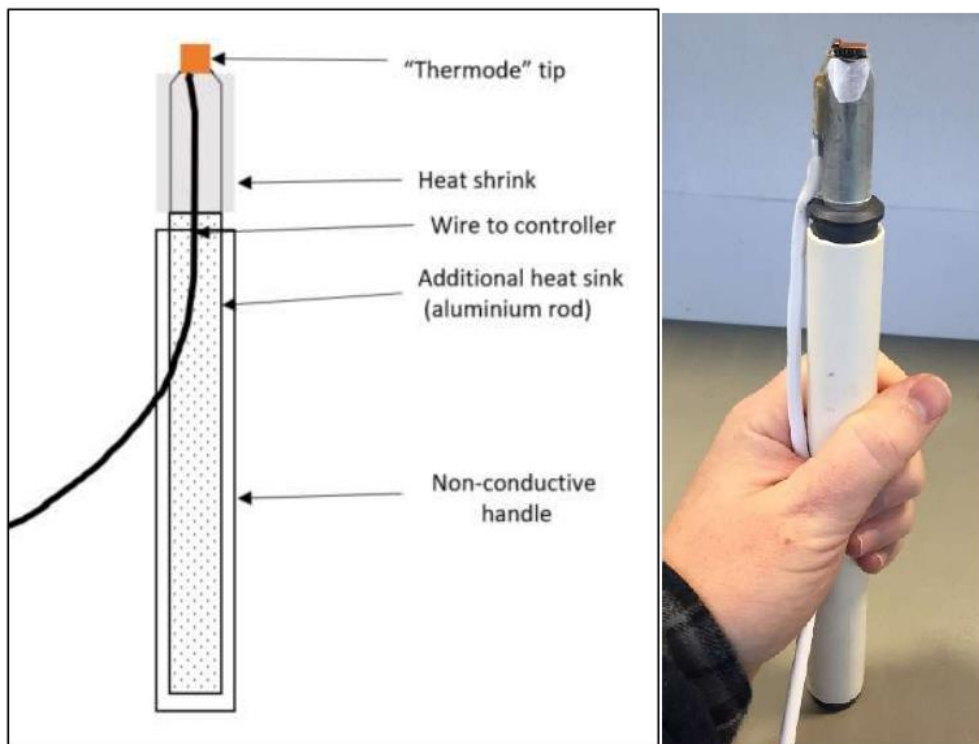


Figure 7 a) A basic diagram of the thermode design. B) A photograph of apparatus

## **4 Software**

Software refers to the additional software/interface requirements, which will allow for experimental data collection. As data collection and thermode control are highly interlinked, a system that communicates with the thermode in real time was designed,

### **4.1 Thermode control**

A basic interface was required which allows the setting of the thermode protocol and start/stop functions.

### **4.2 Data Capturing**

Thermode (time/temperature) data will be captured automatically. This will also be visually displayed to the experimenter.

### **4.3 Interface with data capturing software**

Due to the fact that the functioning of the thermode and the interaction of participants need to be finely synchronised, one piece of software needs to have the capability to trigger both the start of the data collection protocol, as well as the running of the thermode protocol. This means that a custom piece of software is required. A simple overview of the setup is shown in Figure 8, which shows how the different pieces of hardware and software will be connect to one another. This setup enables both the data collection and the thermode to be controlled simultaneously. Note that the thermode is not connected to the mains electricity supply, and will communicate with the laptop computer via Bluetooth.

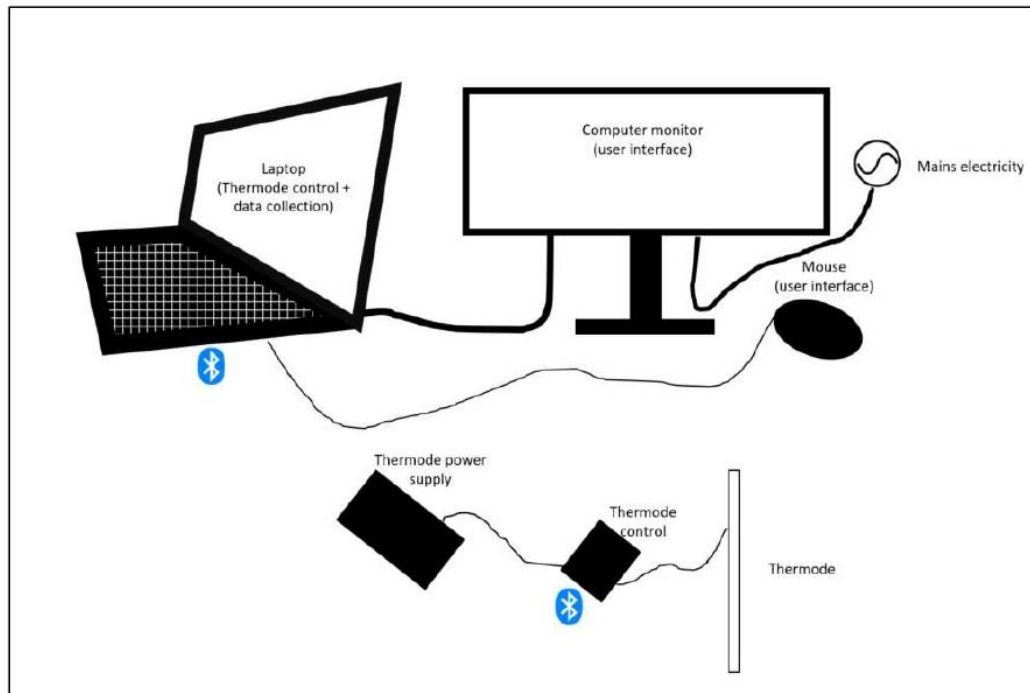


Figure 8 An overview of the final experimental setup

For the validation of the experimental protocol, the custom-built software needed to have the ability to run two sets of user interface – one which collects data in a similar manner to Bajec and Pickering (2008), and one which can collect data with the updated protocol, and the data collection requirements for it. An overview of the different data collection requirements can be seen in Table 2.

Table 2 Overview of data collection requirements for the 2 different protocols

Protocol	Thermode steps	Type of data collected
2 (Pickering)	2 phases: Warming followed by cooling	Taste perceived (Yes/No) Detail of taste (CATA from list of basic tastes) Intensity of tastes (gLMS scale)
4 (Botha)	1 phase: warming and cooling happens in sequence	TCATA of tastes perceived

The details of the sequencing of the two protocols can be seen in Table 3 and Table 4. Protocol 2 runs in two phases, with data collection between the two phases. Three types of data should be collected, namely Yes/No data, Details of tastes (if perceived), as well as data on the gLMS (as described in Green, Shaffer, and Gilmore (1993)). Two examples of the gLMS can be seen in Figure 9. Protocol 4 runs in a single phase, with continuous TCATA data collection.

There were two additional requirements which are independent of the protocols. Firstly, the software should indicate at the beginning of the test that the thermode is still trying to reach setpoint with a display that indicates that the test is not yet ready to start. The software should also have the capacity to display a warning message for participants to remove the thermode from their tongue should excessive temperatures be reached. This temperature is to be determined.

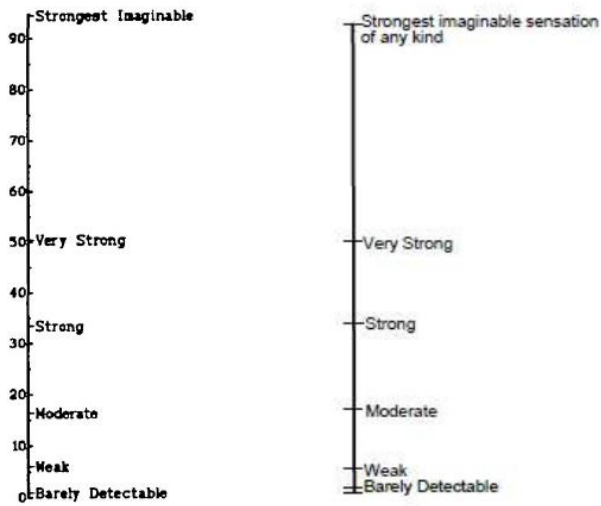


Figure 9 Examples of the gLMS. Left – with markings as first proposed by Green et al. (1993) Right Unmarked as used by Yang (2015)

Table 4 Details of interaction/capturing of the Botha protocol (Protocol 4)

Step	Description	Thermode	Display text	Display requirements	Trigger	Data collection
1	"Welcome"	At 35°C	Welcome to the test. Once you are holding the thermode in place and are comfortable to start the test, please tick next	"Next" button	Researcher starts the test	None
2	Test starts	At 35°C	Please check start when you are ready to start the test	"Start" button - which triggers the protocol	"Next" button	Absolute time of test start (Participant and thermode)
3	Cooling portion of protocol	35°C->5°C -> 40°C	Please check any tastes which you perceive. If you perceive no taste, please check only cooling/warming.	Checkboxes (several can be checked) for basics tastes. List to be confirmed but should include: Sweet, Bitter, Sour, Salty, Savoury, Other	Start button	TCATA data
4	Break	Returning to 35°C, resting	Please take a short break to recover	Timer?	Previous screen completed	None
5	Test ends or repeats		End of test message		Previous screen completed OR "No" on question 5	None

Table 3 Details of interaction/capturing of the Pickering protocol (Protocol 2)

Step	Description	Thermode	Display text	Display requirements	Trigger	Data collection
1	"Welcome"	At 35°C	Welcome to the test. Once you are holding the thermode in place and are comfortable to start the test, please tick next	"Next" button	Researcher starts the test	None
2	Test starts	At 35°C	Please check start when you are ready to start the test	"Start" button - which triggers the protocol	"Next" button	Absolute time of test start (Participant and thermode)
3	Cooling of tongue to run warming protocol	35°C -> 15°C	Please wait while your tongue is cooling		"Start button"	None
4	Warming protocol	15° -> 40°C	Please focus on any taste perceived. Are you perceiving a taste?	Checkboxes (only one can be checked) Yes/No (only temperature)	Time/temperature change	TCATA Yes/No
5	Confirming of tastes	Returning to 35 °C, resting	Please remove your tongue from the thermode. Did you perceive taste during the warming protocol?	Checkboxes (only one can be checked) Yes/No (only temperature)	End of protocol (communicated from the thermode?)	Yes/No
6	Asking detail about tastes		Which tastes did you perceive? Please check all that apply	Checkboxes (several can be checked) for basics tastes. List to be confirmed but should include: Sweet, Bitter, Sour, Salty, Savoury, Other	"Yes" response on previous question	CATA data on each of the
7	Ranking tastes on gLMS		Please rank the intensity of the taste perceived	gLMS for each of the tastes checked in the previous screen	Tastes checked in previous portion	gLMS intensity on each taste
8	Break		Please take a short break to recover	Timer?	Previous screen completed OR "No" on question 5	None
9	Test starts	At 35°C	Please check start when you are ready to start the test. In this part of the test, we will begin collecting data immediately	"Start" button - which triggers the protocol	Specific time of countdown	Absolute time of test start (Participant and thermode)
10	Cooling protocol	35°C -> 5°C + rest	Please focus on any taste perceived. Are you perceiving a taste?	Checkboxes (only one can be checked) Yes/No (only temperature)	"Start button"	TCATA Yes/No
11	Confirming of tastes	Returning to 35 °C, resting	Please remove your tongue from the thermode. Did you perceive taste during the cooling protocol?	Checkboxes (only one can be checked) Yes/No (only temperature)	End of protocol (communicated from the thermode?)	Yes/No
12	Asking detail about tastes		Which tastes did you perceive? Please check all that apply	Checkboxes (several can be checked) for basics tastes. List to be confirmed but should include: Sweet, Bitter, Sour, Salty, Savoury, Other	"Yes" response on previous question	CATA data on each of the
13	Ranking tastes on gLMS		Please rank the intensity of the taste perceived	gLMS for each of the tastes checked in the previous screen	Tastes checked in previous portion	gLMS intensity on each taste
14	Test ends or repeats		End of test message		Previous screen completed OR "No" on question 11	None

## 5 Final thoughts

This document is intended to be a working document and to act as a guide for further development of the thermode – and is not intended as a final or definite description of all future developments – as these should be guided both by what is practical and available as well as the experimental requirements.

## 6 References:

- Bajec, M. R., & Pickering, G. J. (2008). Thermal taste, PROP responsiveness, and perception of oral sensations. *Physiol Behav*, *95*(4), 581-590. 10.1016/j.physbeh.2008.08.009
- Cruz, A., & Green, B. G. (2000). Thermal stimulation of taste. *Nature*, *403*(6772), 889-892.
- Green, B. G., Alvarez-Reeves, M., George, P., & Akirav, C. (2005). Chemesthesis and taste: Evidence of independent processing of sensation intensity. *Physiology & Behavior*, *86*(4), 526-537. 10.1016/j.physbeh.2005.08.038
- Green, B. G., & George, P. (2004). 'Thermal taste' predicts higher responsiveness to chemical taste and flavor. *Chemical Senses*, *29*(7), 617-628. 10.1093/chemse/bjh065
- Green, B. G., Shaffer, G. S., & Gilmore, M. M. (1993). DERIVATION AND EVALUATION OF A SEMANTIC SCALE OF ORAL SENSATION MAGNITUDE WITH APPARENT RATIO PROPERTIES. *Chemical Senses*, *18*(6), 683-702. 10.1093/chemse/18.6.683
- Pickering, G. J., & Klodnicki, C. E. (2016). Does Liking and Orosensation Intensity Elicited by Sampled Foods Vary with Thermal Tasting? *Chemosensory Perception*, *9*(2), 47-55. 10.1007/s12078-016-9207-1
- Pickering, G. J., & Kvas, R. (2016). Thermal Tasting and Difference Thresholds for Prototypical Tastes in Wine. *Chemosensory Perception*, *9*(1), 37-46. 10.1007/s12078-016-9203-5
- Pickering, G. J., Lucas, S., & Gaudette, N. J. (2016). Variation in orosensation and liking of sampled foods with thermal tasting phenotype. *Flavour*, *5*(1), 2. 10.1186/s13411-016-0049-1
- Skinner, M., Eldeghaidy, S., Ford, R., Giesbrecht, T., Thomas, A., Francis, S., & Hort, J. (2018). Variation in thermally induced taste response across thermal tasters. *Physiol Behav*, *188*, 67-78. 10.1016/j.physbeh.2018.01.017
- Yang, Q. (2015). *Individual variation across PROP and Thermal Taste phenotypes*. (PhD), University of Nottingham,

## Appendix D. Additional tables RapCoTT

Table D-1 Summary of contingency tables comparing total numbers of classified ( $Y = TT + TnT$ ) and non-classified ( $N = UC$ ) individuals for each pair of methods and results and associated Odds Ratio (OR) and probability values

All data											
Scheme A	RapCoTT			Scheme B	RapCoTT			Scheme B	Scheme A		
	N	Y	Total		N	Y	Total		N	Y	Total
N	3	40	43	N	3	24	26	N	19	7	26
Y	7	82	89	Y	7	99	106	Y	24	82	106
Total	10	122	132	Total	10	122	132	Total	43	89	132
	OR		5.7		OR		3.29		OR		0.29
	$p$		<0.001		$p$		0.005		$p$		0.003

Traditional first											
Scheme A	RapCoTT			Scheme B	RapCoTT			Scheme B	Scheme A		
	N	Y	Total		N	Y	Total		N	Y	Total
N	3	24	27	N	3	15	18	N	13	5	18
Y	2	36	38	Y	2	45	47	Y	14	33	47
Total	5	60	65	Total	5	60	65	Total	27	38	65
	OR		12		OR		7.5		OR		0.36
	$p$		<0.001		$p$		0.002		$p$		0.064

Combined first											
Scheme A	RapCoTT			Scheme B	RapCoTT			Scheme B	Scheme A		
	N	Y	Total		N	Y	Total		N	Y	Total
N	0	16	16	N	0	8	8	N	6	2	8
Y	5	46	51	Y	25	54	79	Y	10	49	59
Total	5	62	67	Total	25	62	87	Total	16	51	67
	OR		3.2		OR		1.6		OR		0.2
	$p$		0.027		$p$		0.581		$p$		0.039

Table D–2 Numbers of participants (percentages in brackets) of agreement and disagreement between pairs of methods, taking protocol order into account. Cells where classifications agree are underlined. Cells where classifications disagree are italicised. Zero frequencies are highlighted in bold.

Scheme\_A vs Scheme\_C

Total set					Traditional first					Combined first				
<b>C</b>					<b>C</b>					<b>C</b>				
<b>A</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total	<b>A</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total	<b>A</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total
<b>TnT</b>	<u>23 (17%)</u>	2 (2%)	5 (4%)	30	<b>TnT</b>	<u>11 (17%)</u>	1 (2%)	5 (8%)	17	<b>TnT</b>	<u>12 (18%)</u>	1 (1%)	<b>0 (0%)</b>	13
<b>UC</b>	10 (8%)	<u>3 (2%)</u>	30 (23%)	43	<b>UC</b>	2 (3%)	<u>3 (5%)</u>	22 (34%)	27	<b>UC</b>	8 (12%)	<b>0 (0%)</b>	8 (12%)	16
<b>TT</b>	5 (4%)	5 (4%)	<u>49 (37%)</u>	59	<b>TT</b>	<b>0 (0%)</b>	1 (2%)	<u>20 (31%)</u>	21	<b>TT</b>	5 (7%)	4 (6%)	<u>29 (43%)</u>	38
Total	38	10	84	132	Total	13	5	47	65	Total	25	5	37	67

Scheme\_B vs Scheme\_C

Total set					Traditional first					Combined first				
<b>C</b>					<b>C</b>					<b>C</b>				
<b>B</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total	<b>B</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total	<b>B</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total
<b>TnT</b>	<u>19 (14%)</u>	1 (1%)	<b>0 (0%)</b>	20	<b>TnT</b>	<u>10 (15%)</u>	<b>0 (0%)</b>	<b>0 (0%)</b>	10	<b>TnT</b>	<u>9 (13%)</u>	1 (1%)	<b>0 (0%)</b>	10
<b>UC</b>	8 (6%)	<u>3 (2%)</u>	15 (11%)	26	<b>UC</b>	3 (5%)	<u>3 (5%)</u>	12 (18%)	18	<b>UC</b>	5 (7%)	<b>0 (0%)</b>	3 (4%)	8
<b>TT</b>	11 (8%)	6 (5%)	<u>69 (52%)</u>	86	<b>TT</b>	<b>0 (0%)</b>	2 (3%)	<u>35 (54%)</u>	37	<b>TT</b>	11 (16%)	4 (6%)	<u>34 (51%)</u>	49
Total	38	10	84	132	Total	13	5	47	65	Total	25	5	37	67

Scheme\_A vs Scheme\_B

Total set					Traditional first					Combined first				
<b>B</b>					<b>B</b>					<b>B</b>				
<b>A</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total	<b>A</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total	<b>A</b>	<b>TnT</b>	<b>UC</b>	<b>TT</b>	Total
<b>TnT</b>	<u>20 (15%)</u>	7 (5%)	3 (2%)	30	<b>TnT</b>	<u>10 (15%)</u>	5 (8%)	2 (3%)	17	<b>TnT</b>	<u>10 (15%)</u>	2 (3%)	1 (1%)	13
<b>UC</b>	<b>0 (0%)</b>	<u>19 (14%)</u>	24 (18%)	43	<b>UC</b>	<b>0 (0%)</b>	<u>13 (20%)</u>	14 (22%)	27	<b>UC</b>	<b>0 (0%)</b>	<u>6 (9%)</u>	10 (15%)	16
<b>TT</b>	<b>0 (0%)</b>	<b>0 (0%)</b>	<u>59 (45%)</u>	59	<b>TT</b>	<b>0 (0%)</b>	<b>0 (0%)</b>	<u>21 (32%)</u>	21	<b>TT</b>	<b>0 (0%)</b>	<b>0 (0%)</b>	<u>38 (57%)</u>	38
Total	20	26	86	132	Total	10	18	37	65	Total	10 (15%)	8 (12%)	49 (73%)	67

## Appendix E. FDA derivatives

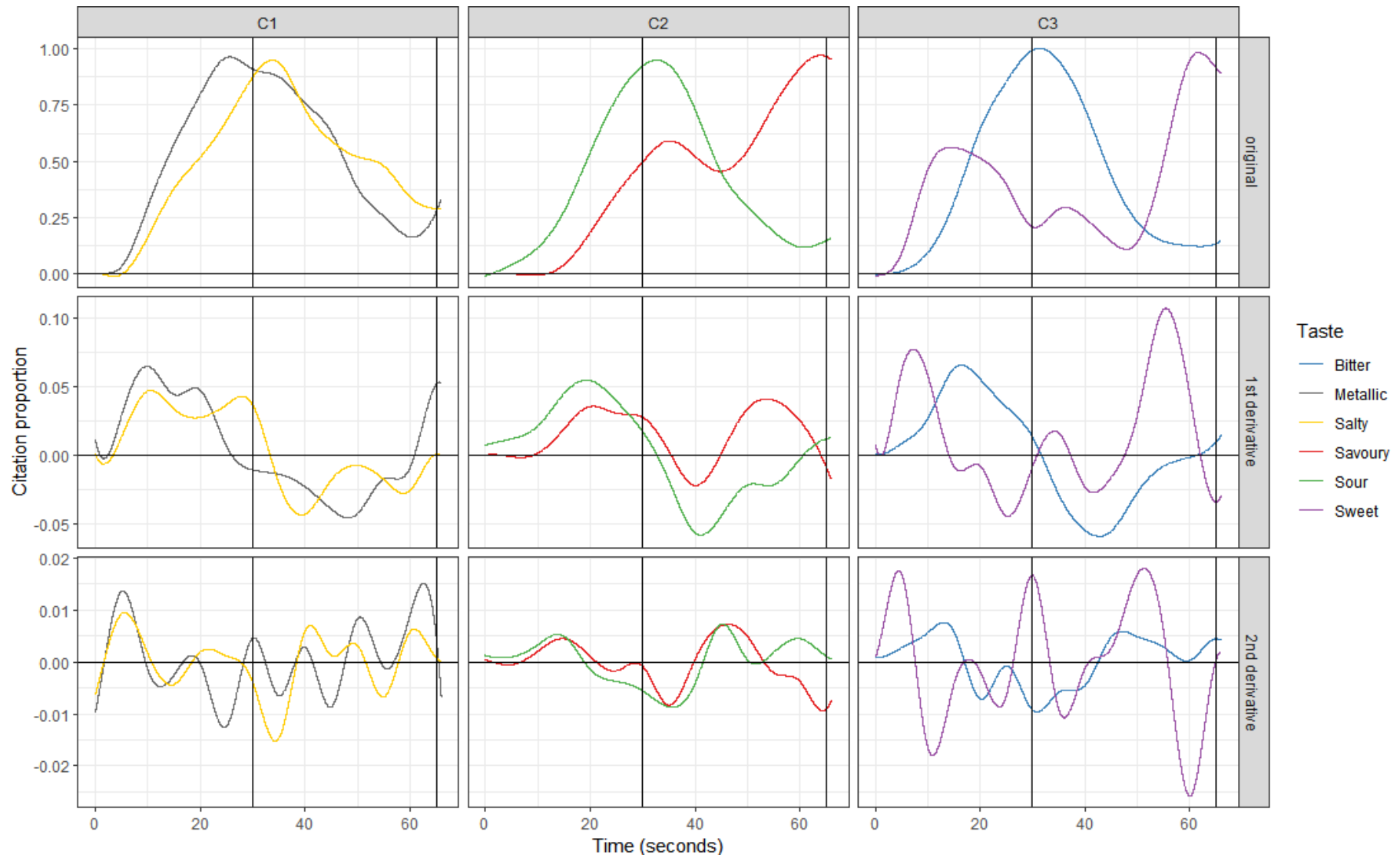


Figure E-1 Functional curves and their related first and second derivatives as obtained from fda

## **Appendix F. Mathematical details of Fatty acid detection calculation**

### **Data cleaning**

All difference from control data was rescaled to take into account response bias. The response of the blind control was set to zero, and responses to any samples scored less different from the control than the blind control was set to zero. Responses of the remaining samples were set to the difference between the response to that sample and the blind control.

### **Model fitting**

For each participant, a logistic regression model was fitted. The logistic regression model automatically calculated the intercept, which is defined as the point at which the logistic curve first increased, interpreted as the individual detection threshold. This intercept (threshold) was in the unit “level”, and was converted to oleic acid concentration by finding the point on the logarithmic curve on which the oleic acid concentrations fell corresponding to that level.

## Appendix G. Full distributions of relevant variables

Univariate plots of variables found to affect RFM prediction but not significantly different are included here.

### TT Vs TnT

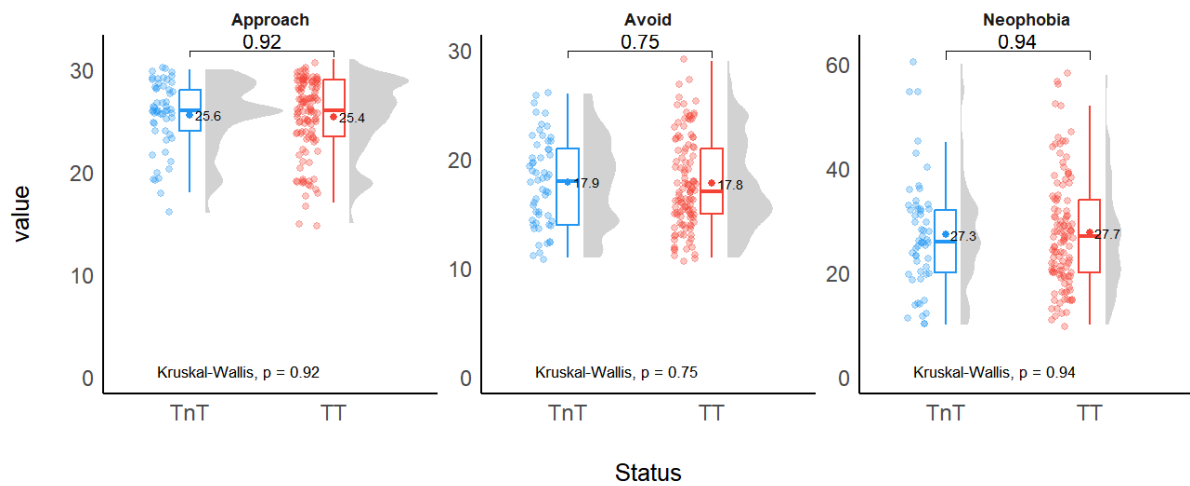


Figure G-1 Distribution of the food neophobia raw data and the Approach and Avoid dimensions of food neophobia according to TT Status, showing boxplots, means, summary statistics for the Kruskal-Wallis test.

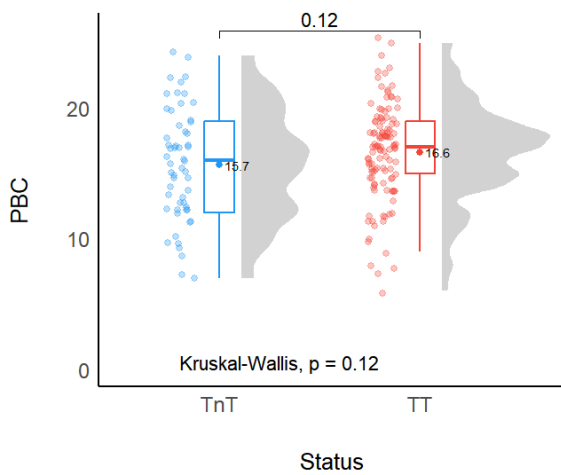


Figure G-2 Distribution of Private Body Consciousness according to TT Status, showing boxplots, means, summary statistics for the Kruskal-Wallis test.

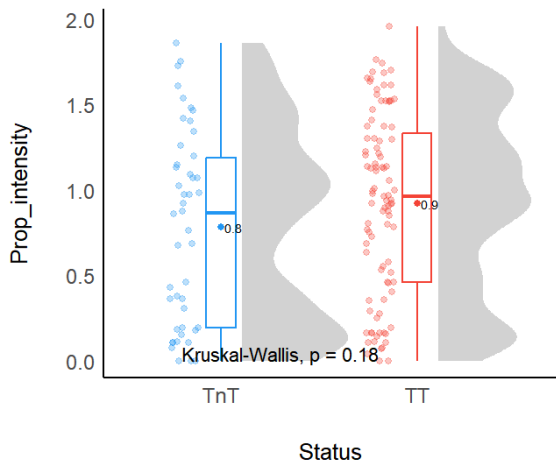


Figure F-3 Distribution PROP intensity according to TT Status, showing boxplots, means, summary statistics for the Kruskal-Wallis test.

### TT Clusters

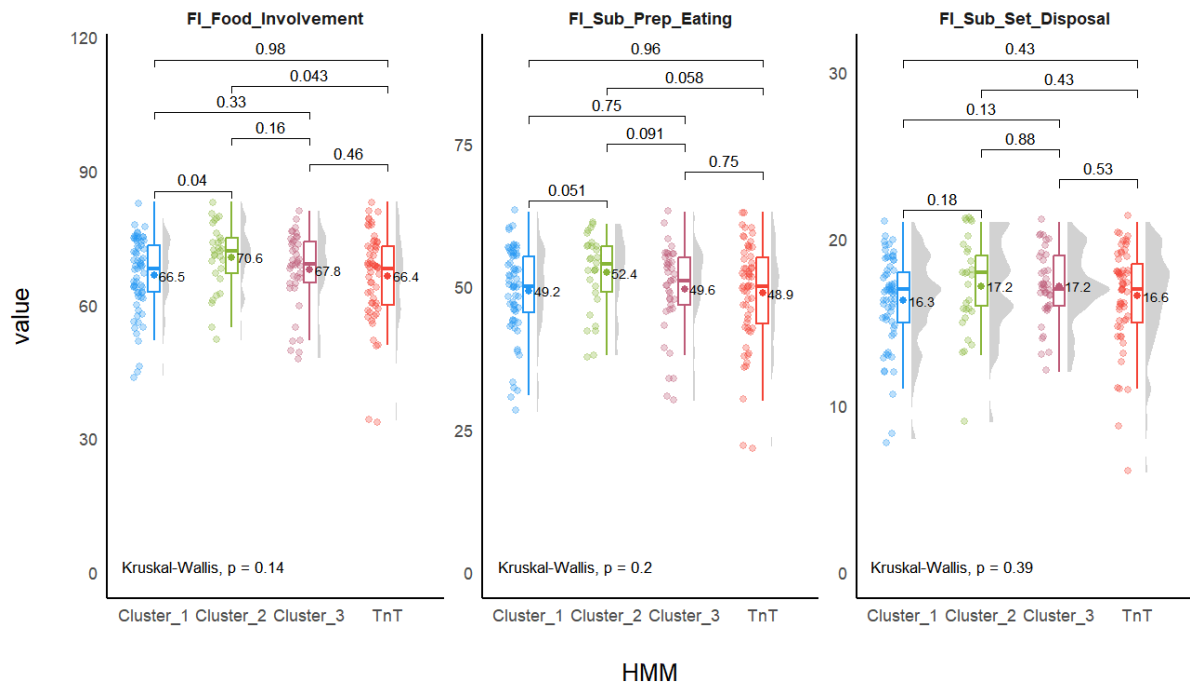


Figure G-4 Distribution of Food Involvement and its two subscales (Preparation & Eating and Set & Disposal) according to TT Cluster, showing boxplots and means, as well as summary statistics for the Kruskal-Wallis test and pairwise comparisons (results of the Wilcox-test) corrected for the false discovery rate (fdr)

## Appendix H. ALE plots binary model

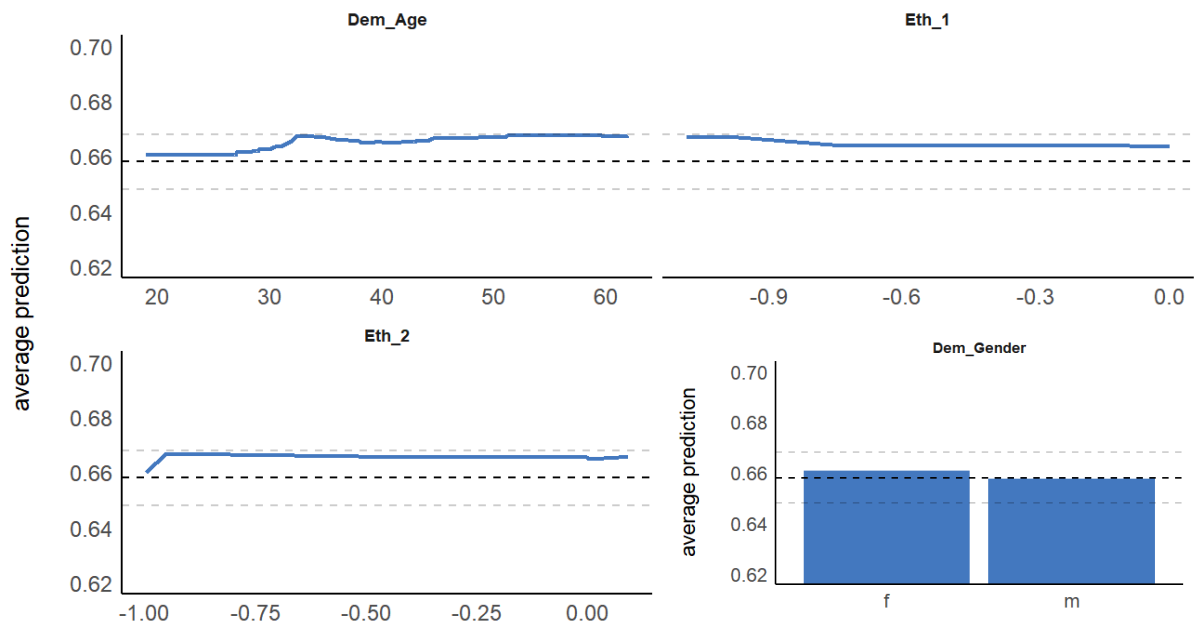


Figure H-1 Accumulated local effects (ALE) plots for the demographic variables included in the final binary (TT status) RFM, showing the , model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

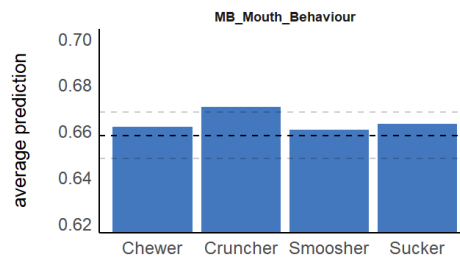


Figure H-2 Accumulated local effects (ALE) plot for oral processing behaviour included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

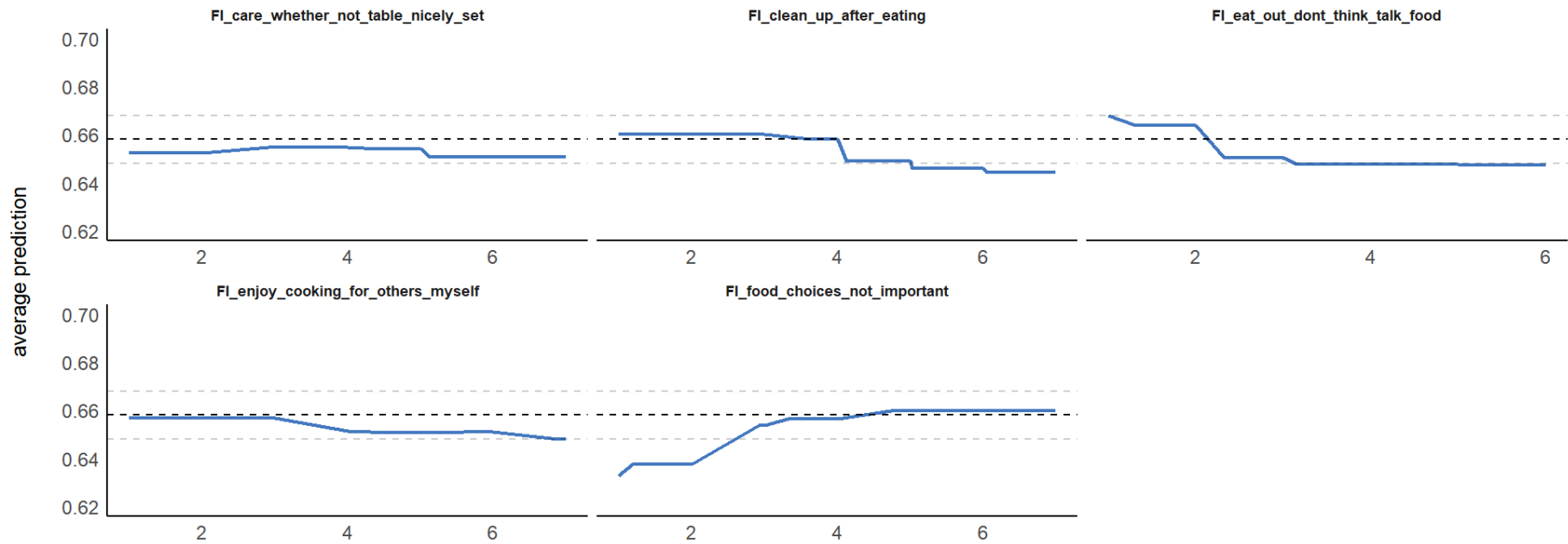


Figure H-3 Accumulated local effects (ALE) plots for the items from the food involvement scale (FIS) included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

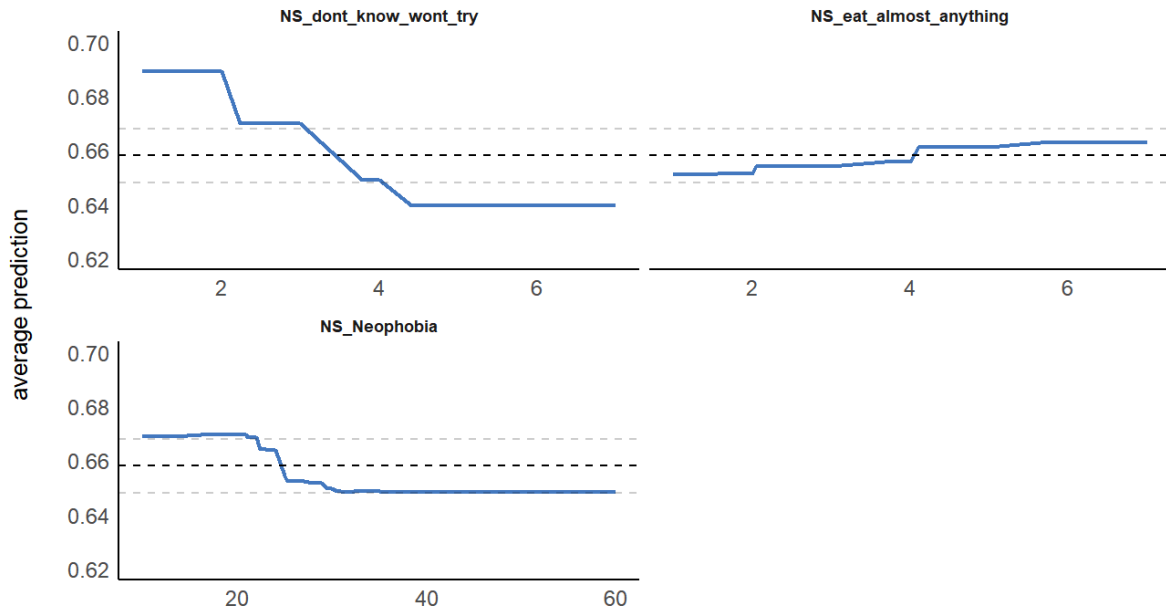


Figure H-4 Accumulated local effects (ALE) plots for calculated indexes and items from the food neophobia scale (FNS) included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

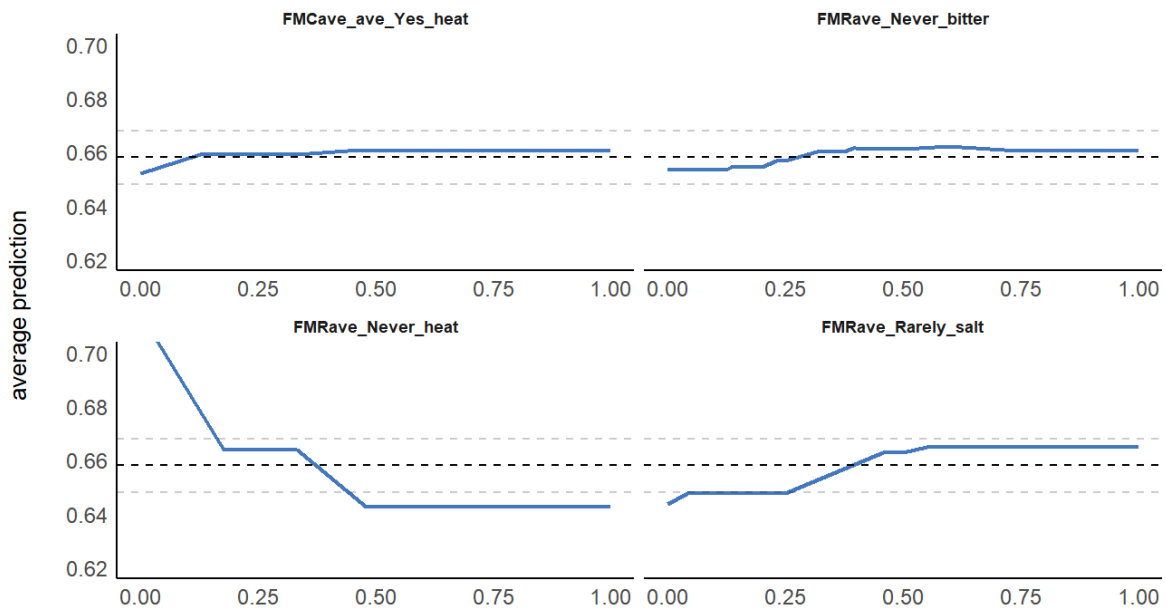


Figure H-5 Accumulated local effects (ALE) plots for calculated indexes from the food modification questionnaire in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. The prefix “FMR” indicates raw values (scale items not concatenated) and “FMC” indicates the combined/concatenated scale.

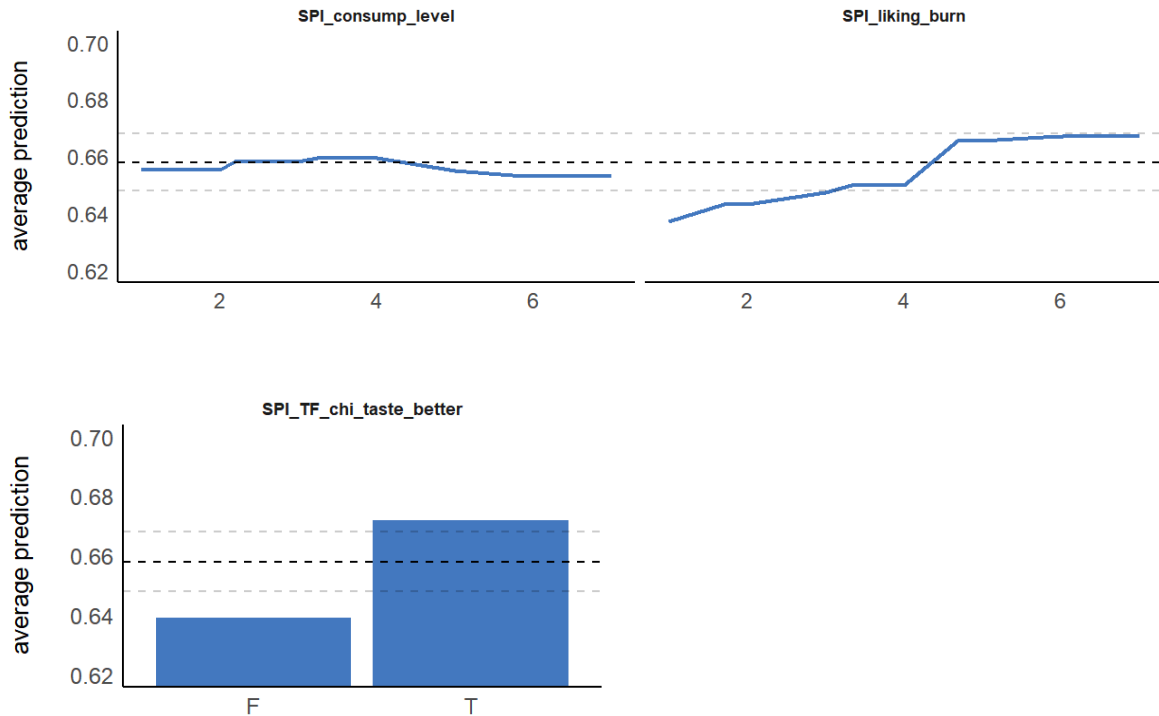


Figure H-5 Accumulated local effects (ALE) plots for the items from the spicy food questionnaire included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm 0.01$ ), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

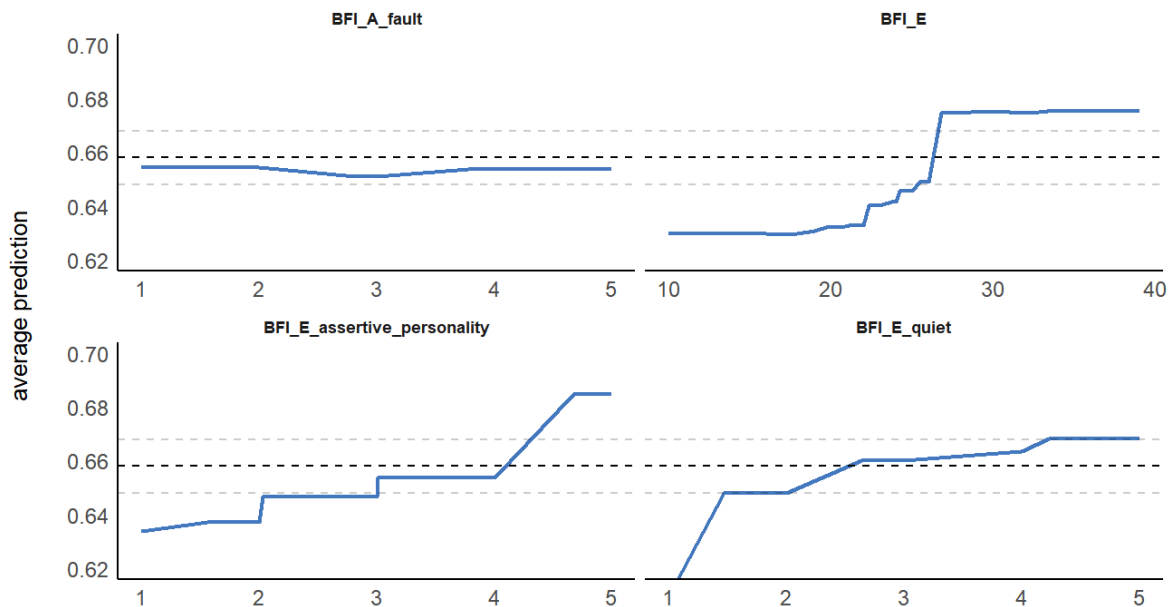


Figure H-6 Accumulated local effects (ALE) plots for personality factors and items from the Big Five Inventory (BFI) included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm 0.01$ ), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. Note that individual items are prefixed by the factors to which they contribute.

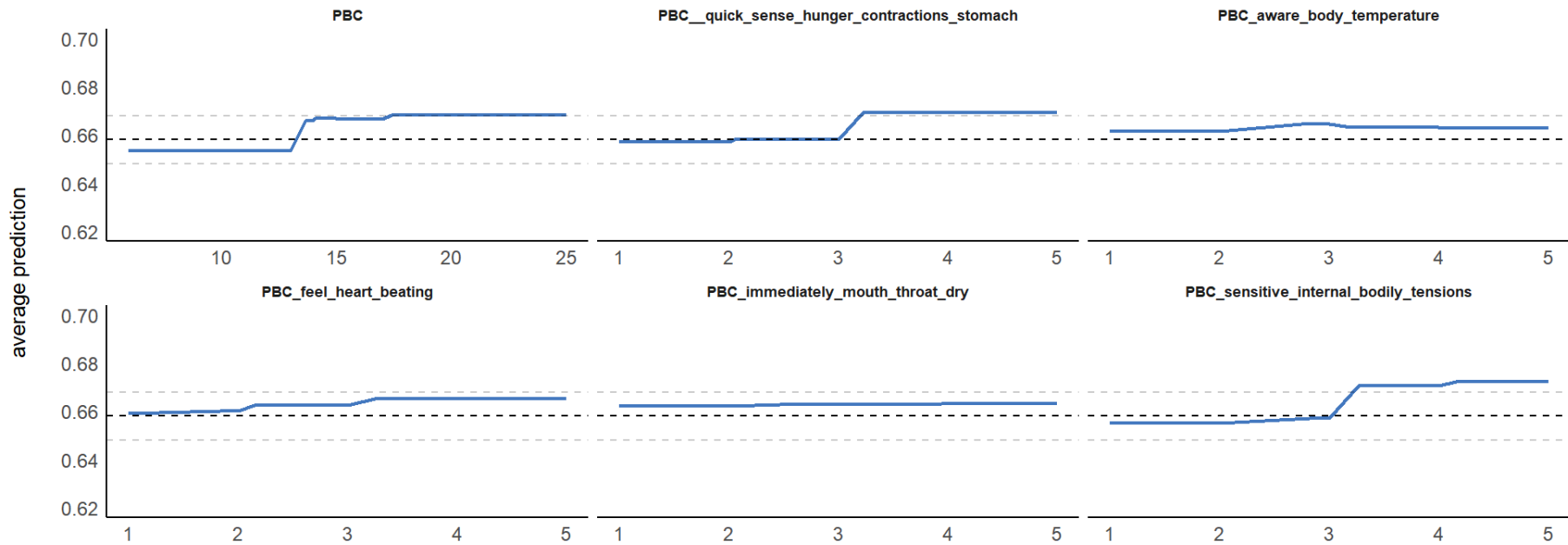


Figure H-7 Accumulated local effects (ALE) plots for Private body consciousness (PBC) and the items from the PBC questionnaire included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

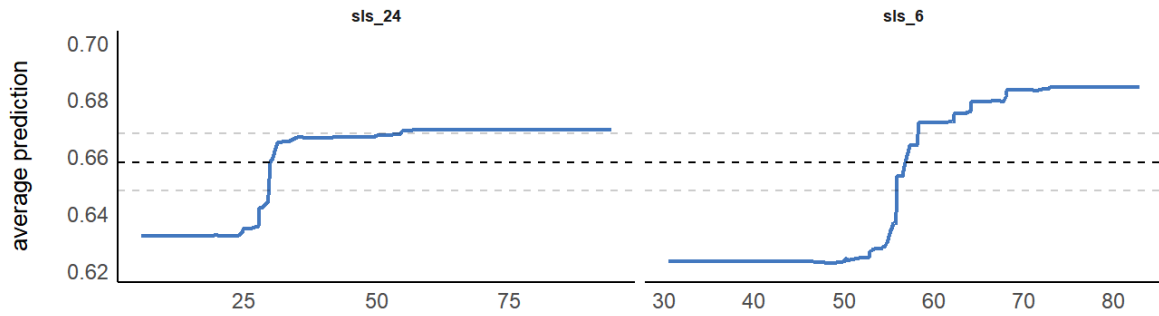


Figure H-8 Accumulated local effects (ALE) plots for sweetness response variables (liking, intensity, SLS cluster) included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability. Note that the prefix “sls” indicates liking on the LAM scale.

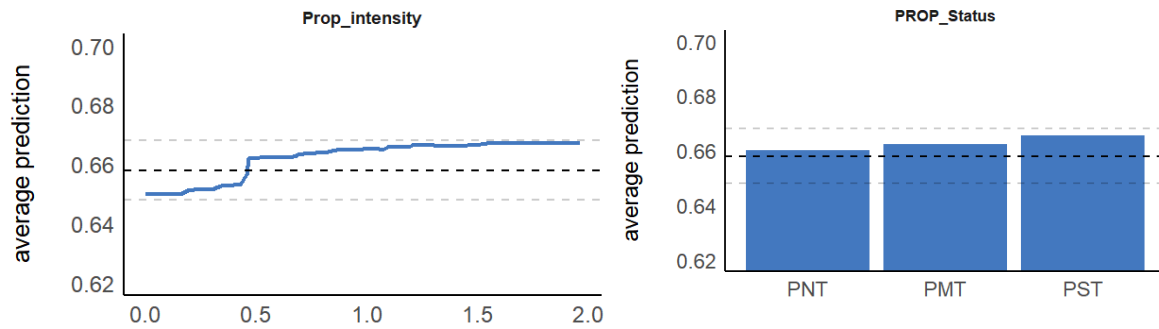


Figure H-9 Accumulated local effects (ALE) plots for PROP responsiveness variables included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

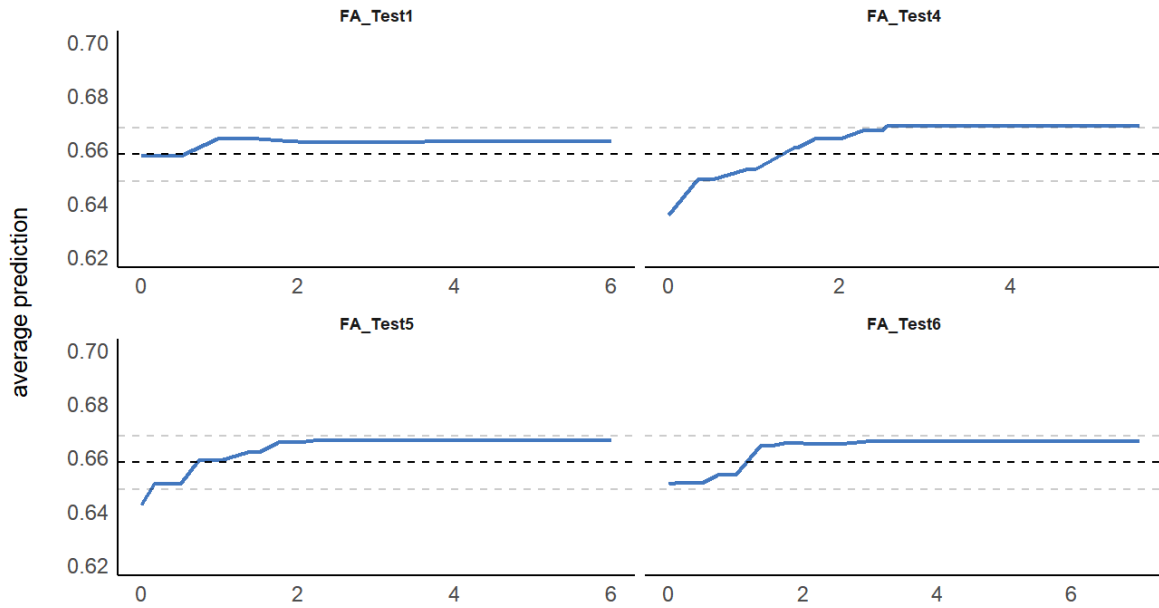


Figure H-10 Accumulated local effects (ALE) plots for Fatty acid response variables included in the final binary (TT status) RFM, showing the model intercept (0.658; TT probability based on sample size alone), two guides (intercept  $\pm$  0.01), and the effect on TT probability across the selected variables. Deviations from intercept indicate a large impact on TT probability, whereas following the intercept indicates little or no impact on TT probability.

## Appendix I. ALE plots multiclass model

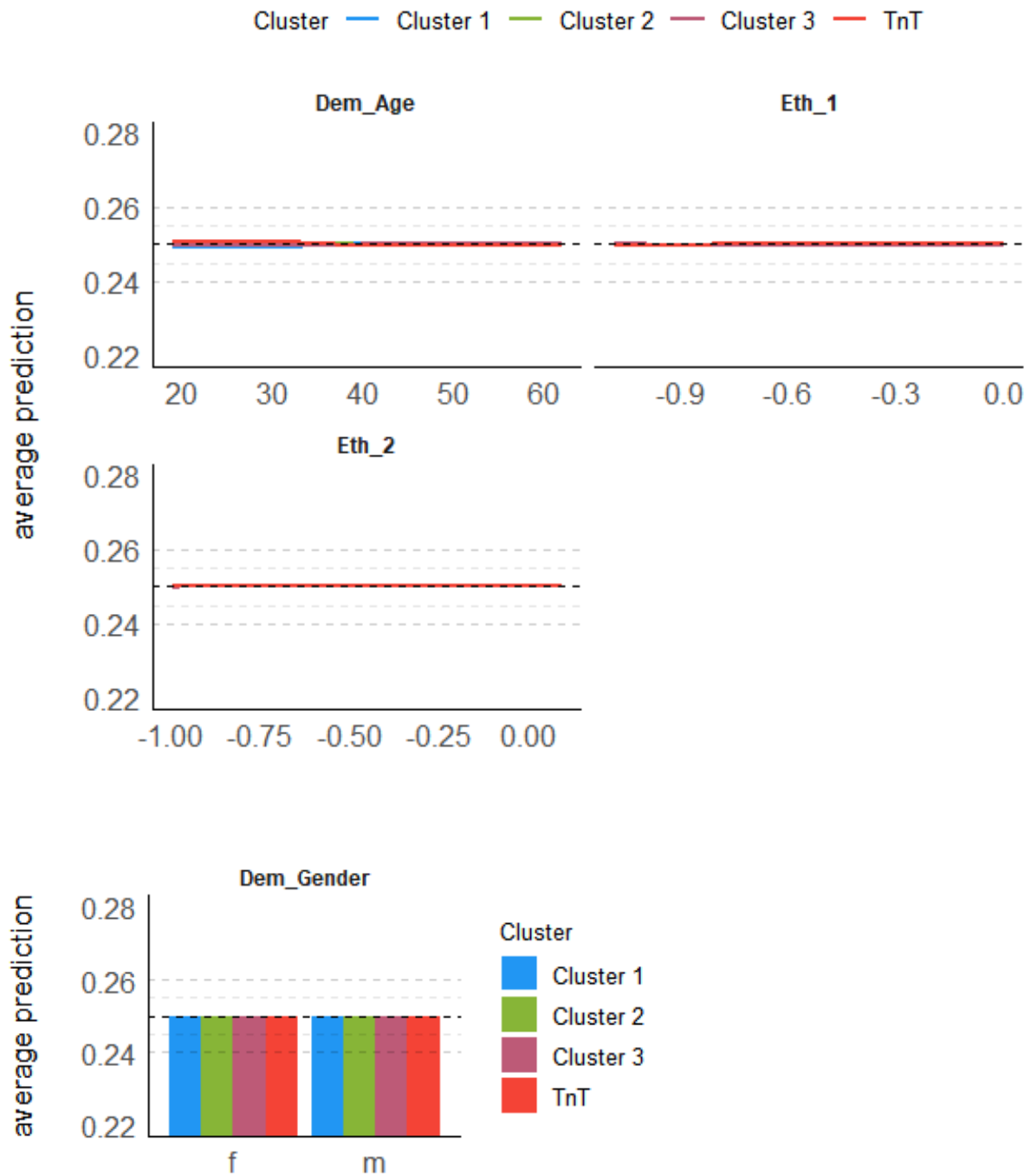


Figure I-1 Accumulated local effects (ALE) plots for the demographic variables included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the intercept indicates little or no impact on cluster probability.

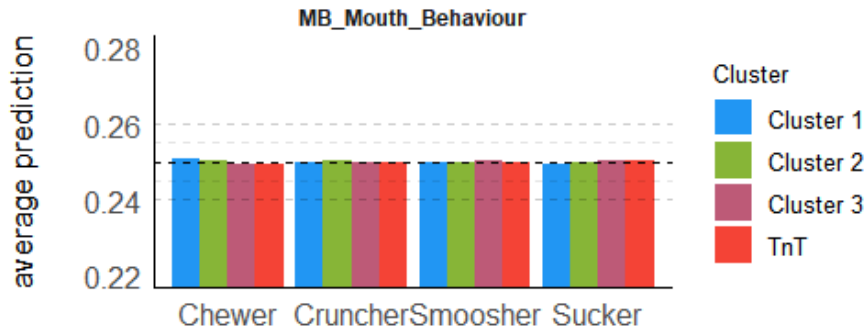


Figure I-1 Accumulated local effects (ALE) plots oral processing behaviour classification in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across different classes of oral processing behaviour. Deviation from chance indicates a large impact on cluster probability.

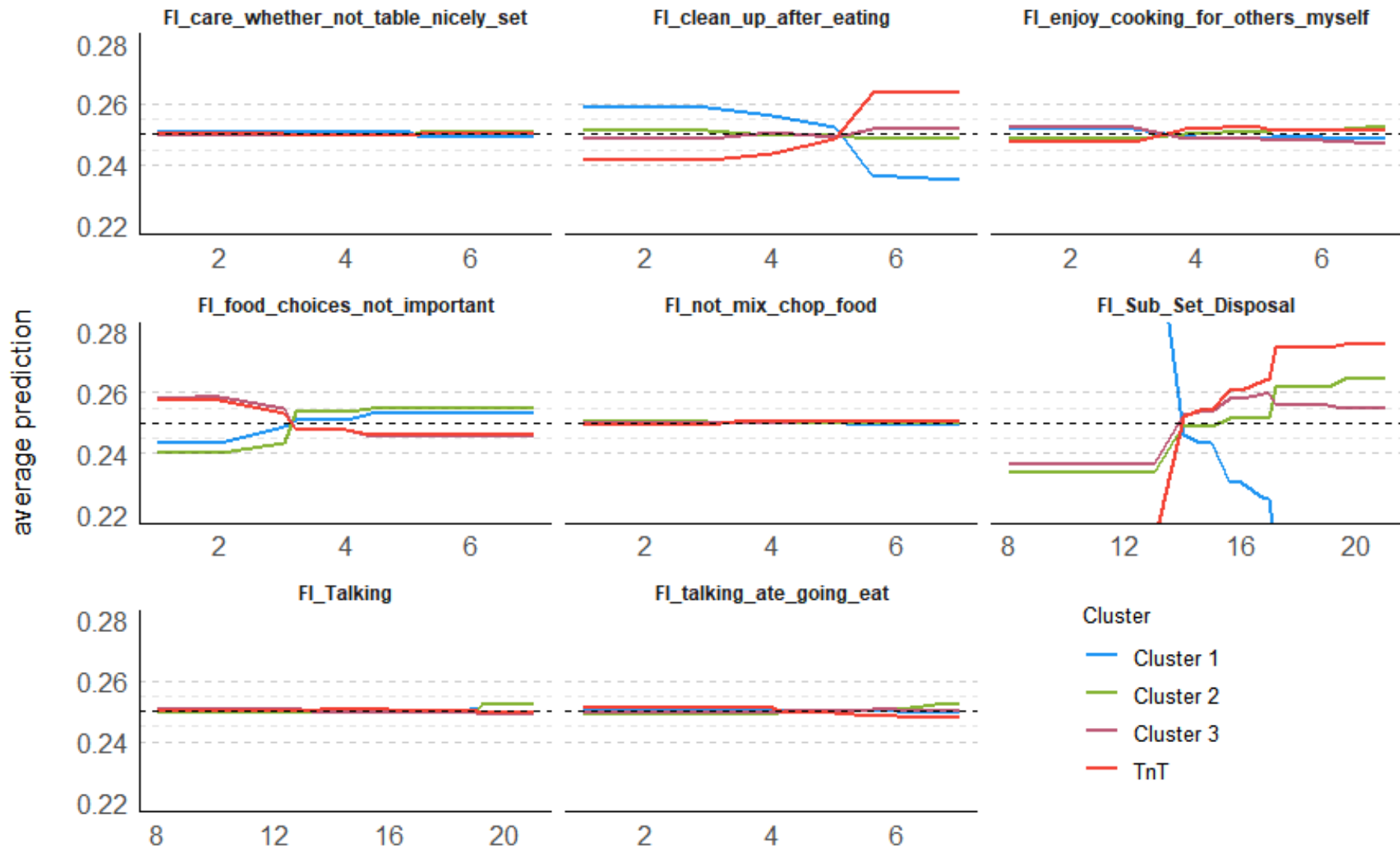


Figure I-3 Accumulated local effects (ALE) plots for items and subscales from the food involvement scale (FIS) included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

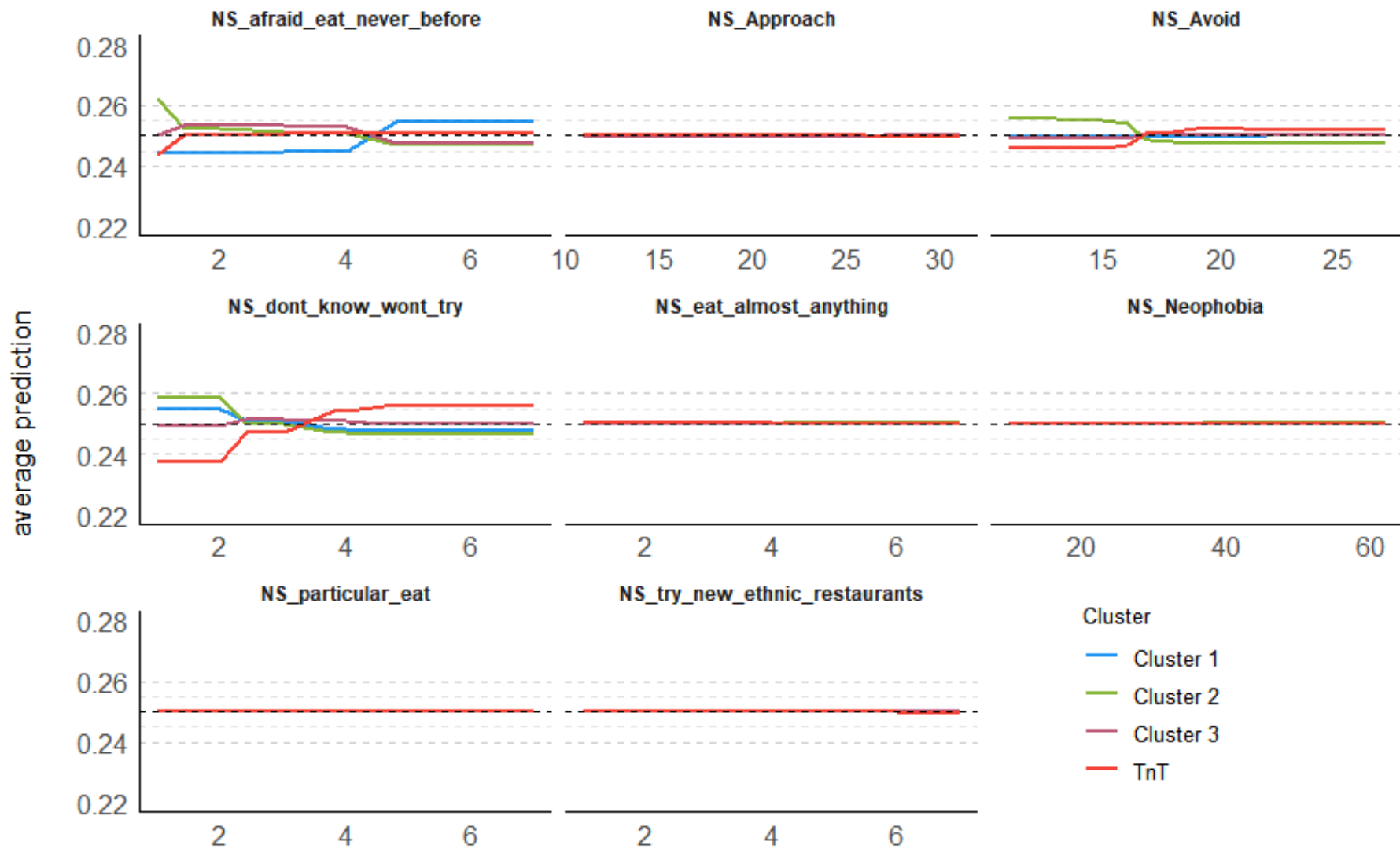


Figure I-4 Accumulated local effects (ALE) plots for items and dimensions from the food neophobia scale (FNS) included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

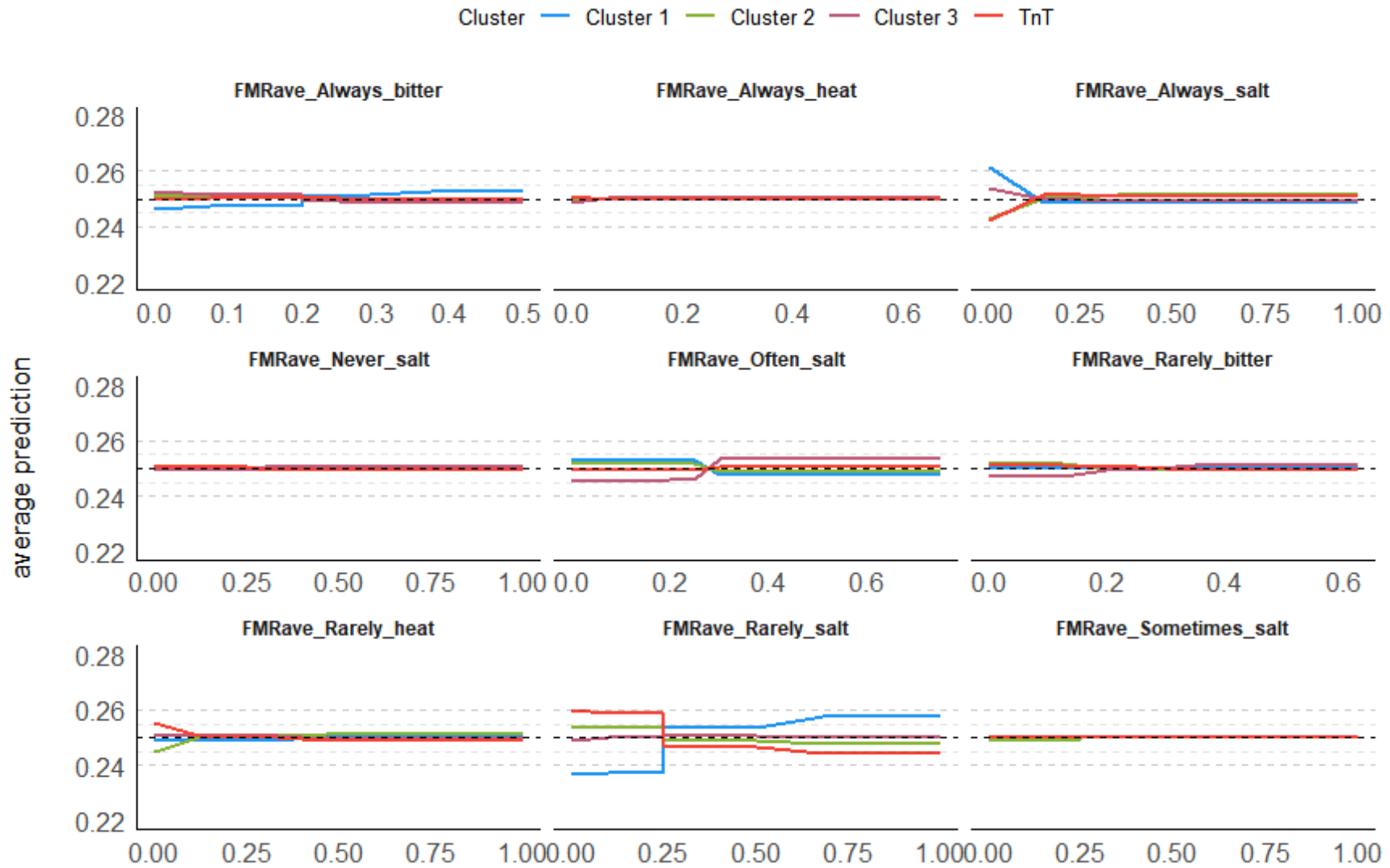


Figure I-5 Accumulated local effects (ALE) plots for calculated raw (unconcatenated) indexes from the food modification questionnaire included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance +/- 0.01 and chance +/- 0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

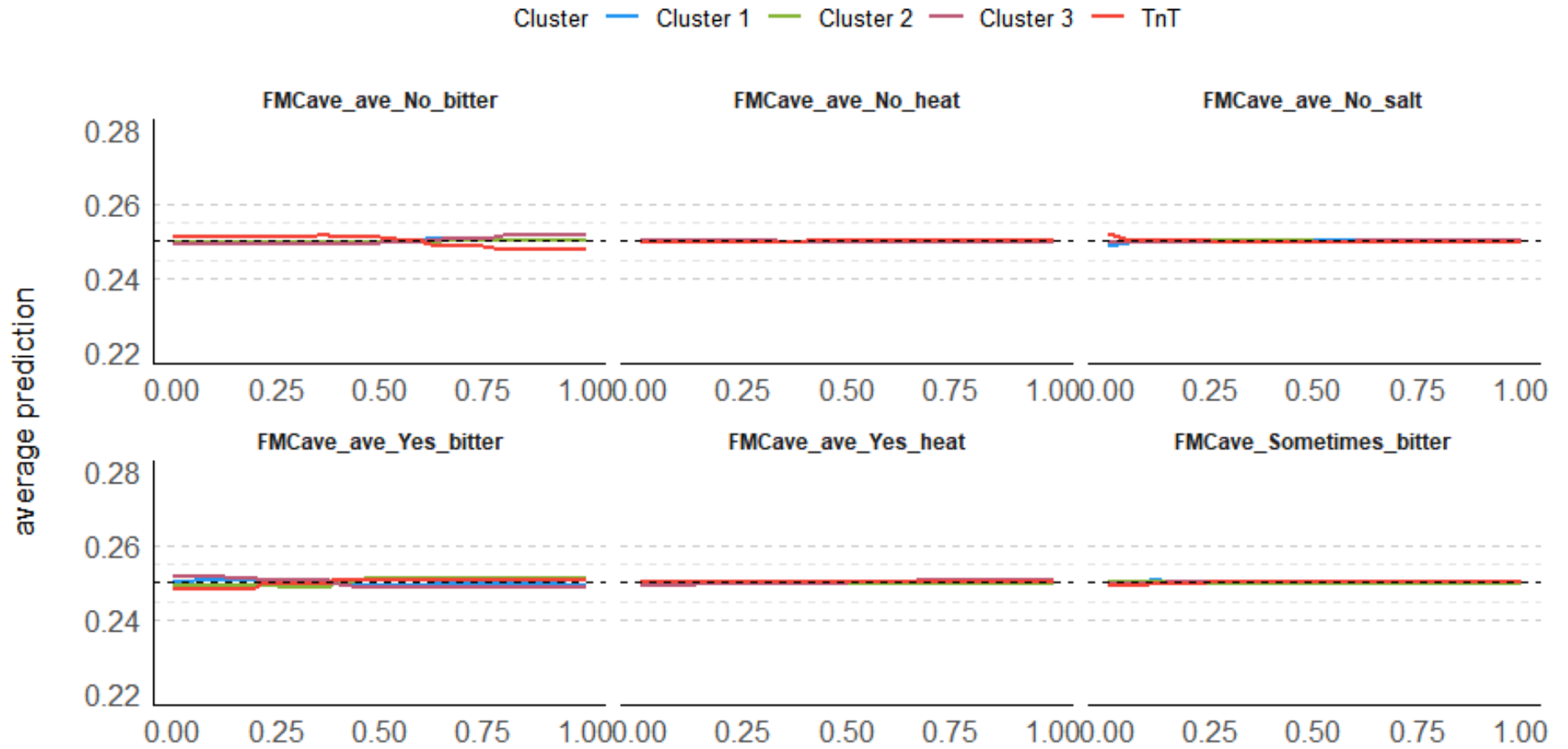


Figure I-7 Accumulated local effects (ALE) plots for combined indexes from the food modification questionnaire included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

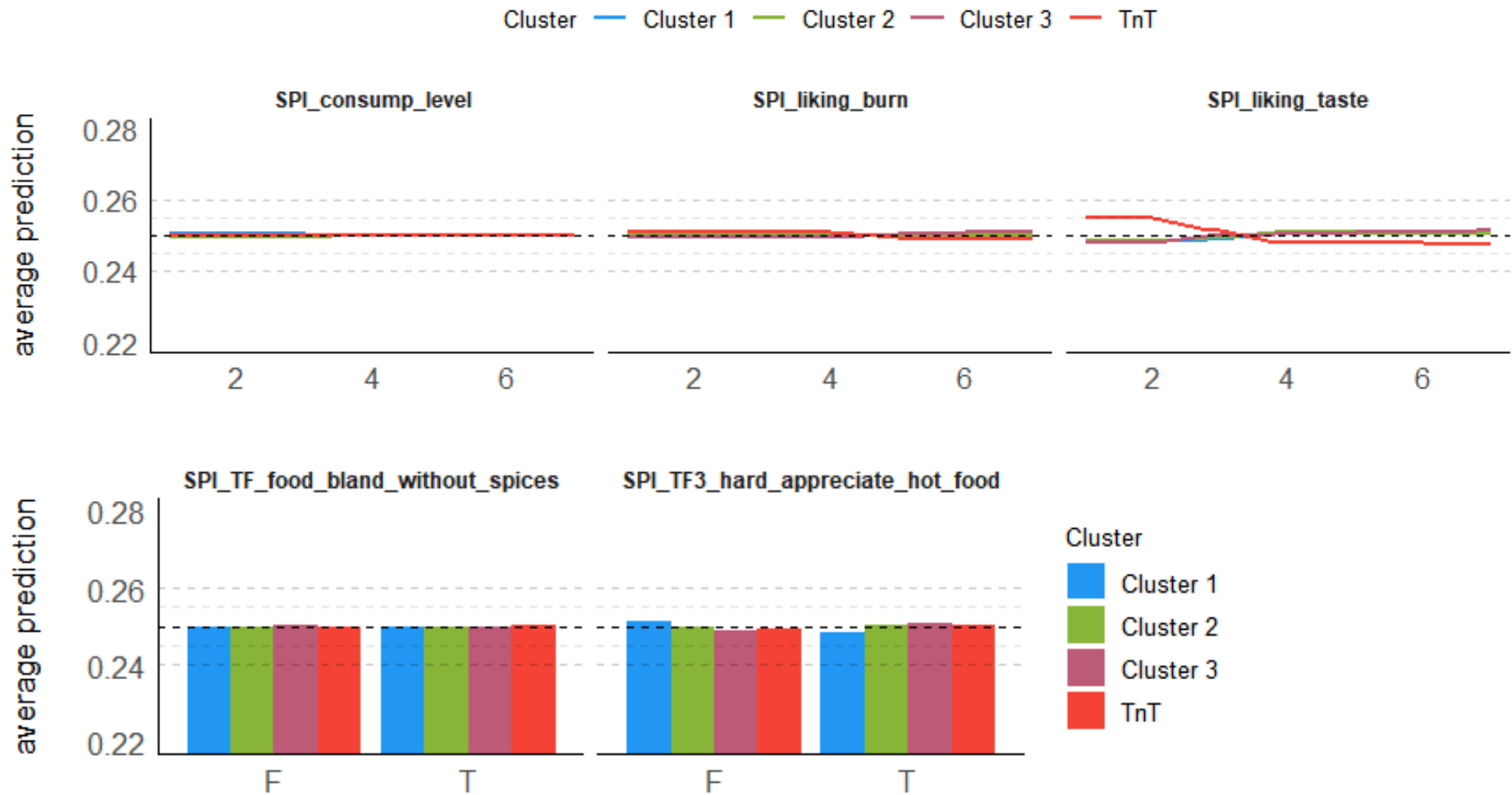


Figure I-8 Accumulated local effects (ALE) plots for items from the spicy food questionnaire included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance +/- 0.01 and chance +/- 0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

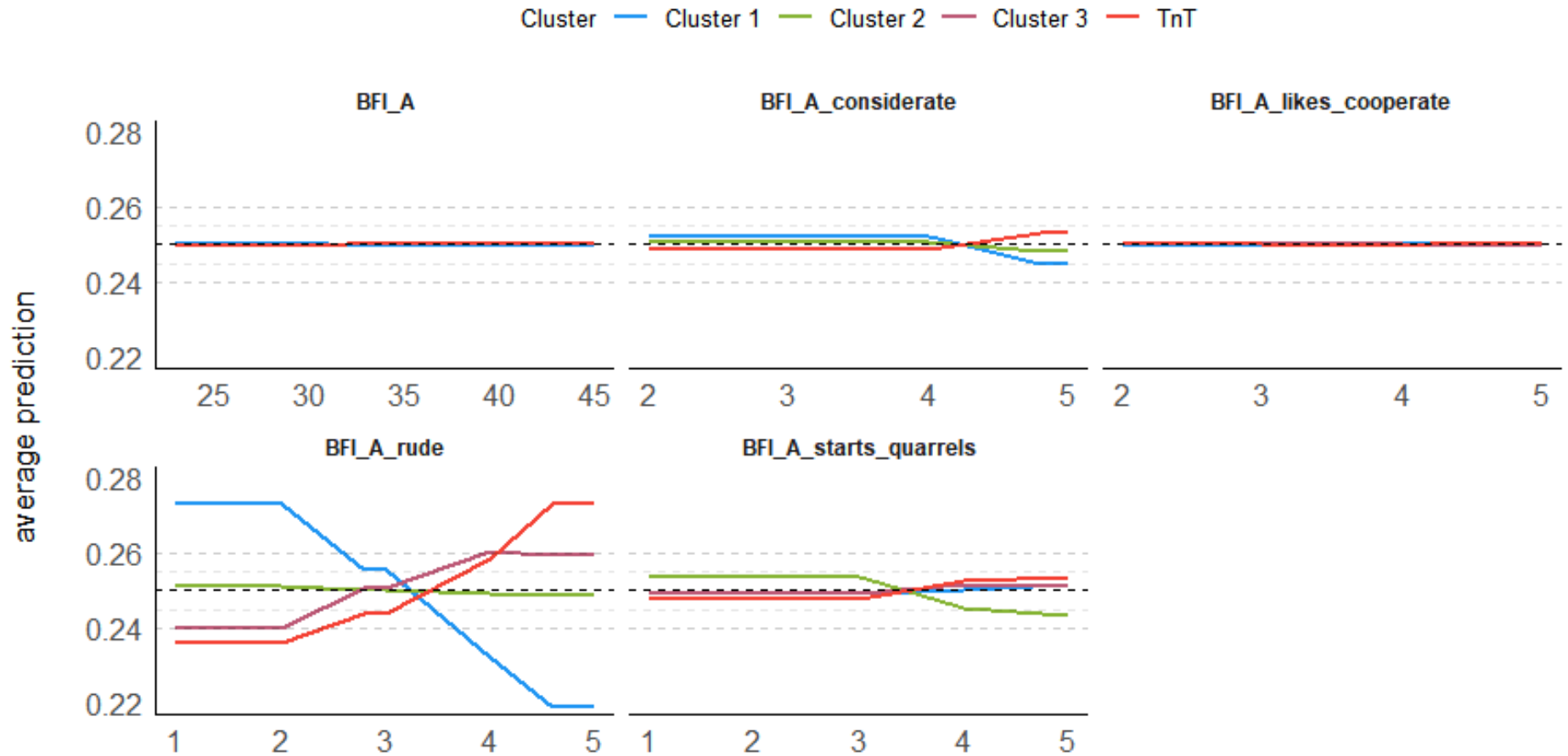


Figure I-9 Accumulated local effects (ALE) plots for personality items and factors from the Agreeableness (BFI\_A) factor of the Big Five Inventory (BFI) included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance +/- 0.01 and chance +/- 0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that individual items are prefixed by the factors to which they contribute.

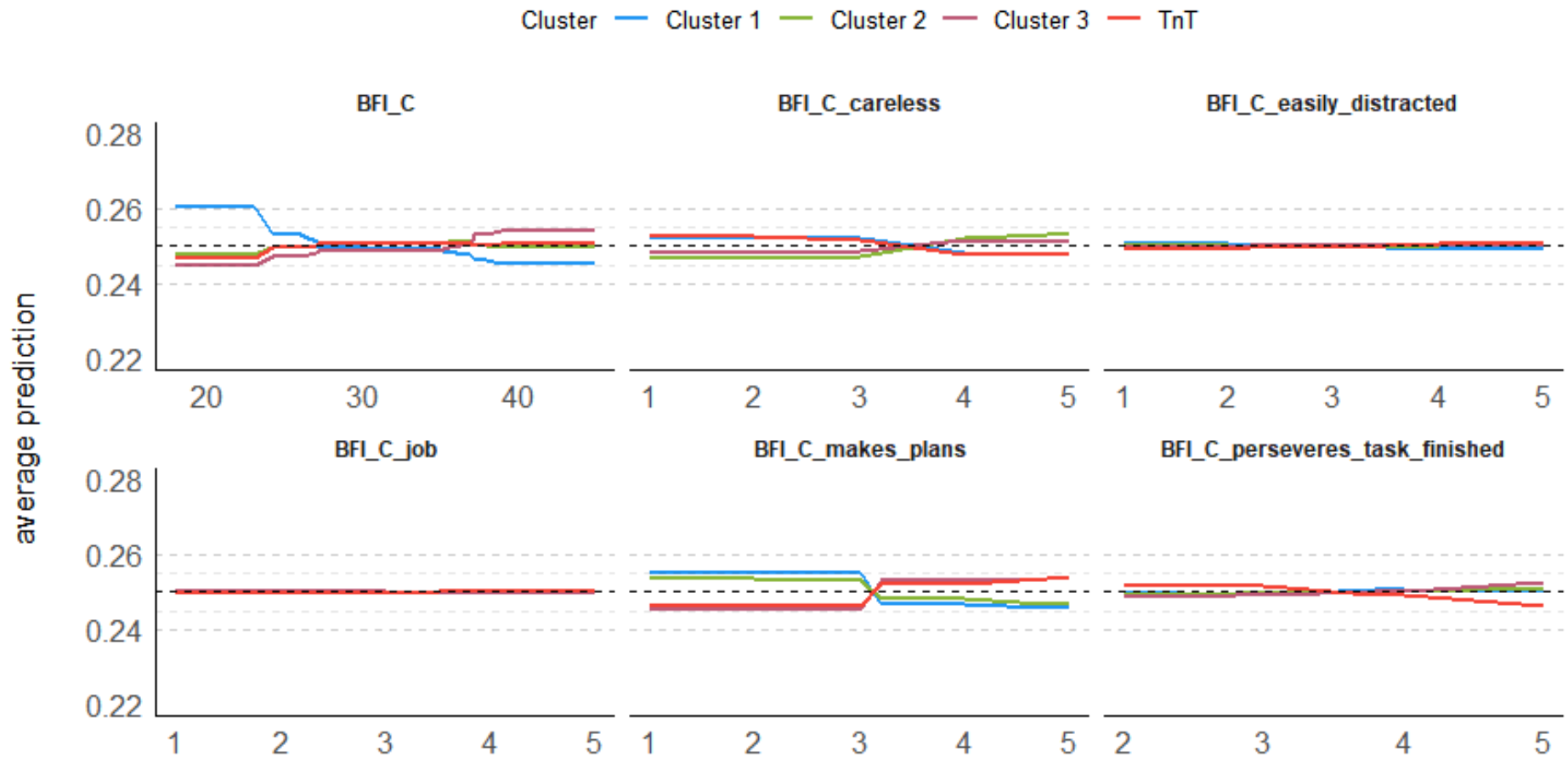


Figure I-10 Accumulated local effects (ALE) plots for personality items and factors from the Conscientiousness (BFI\_C) factor of the Big Five Inventory (BFI) included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm 0.01$  and chance  $\pm 0.005$ ), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that individual items are prefixed by the factors to which they contribute.

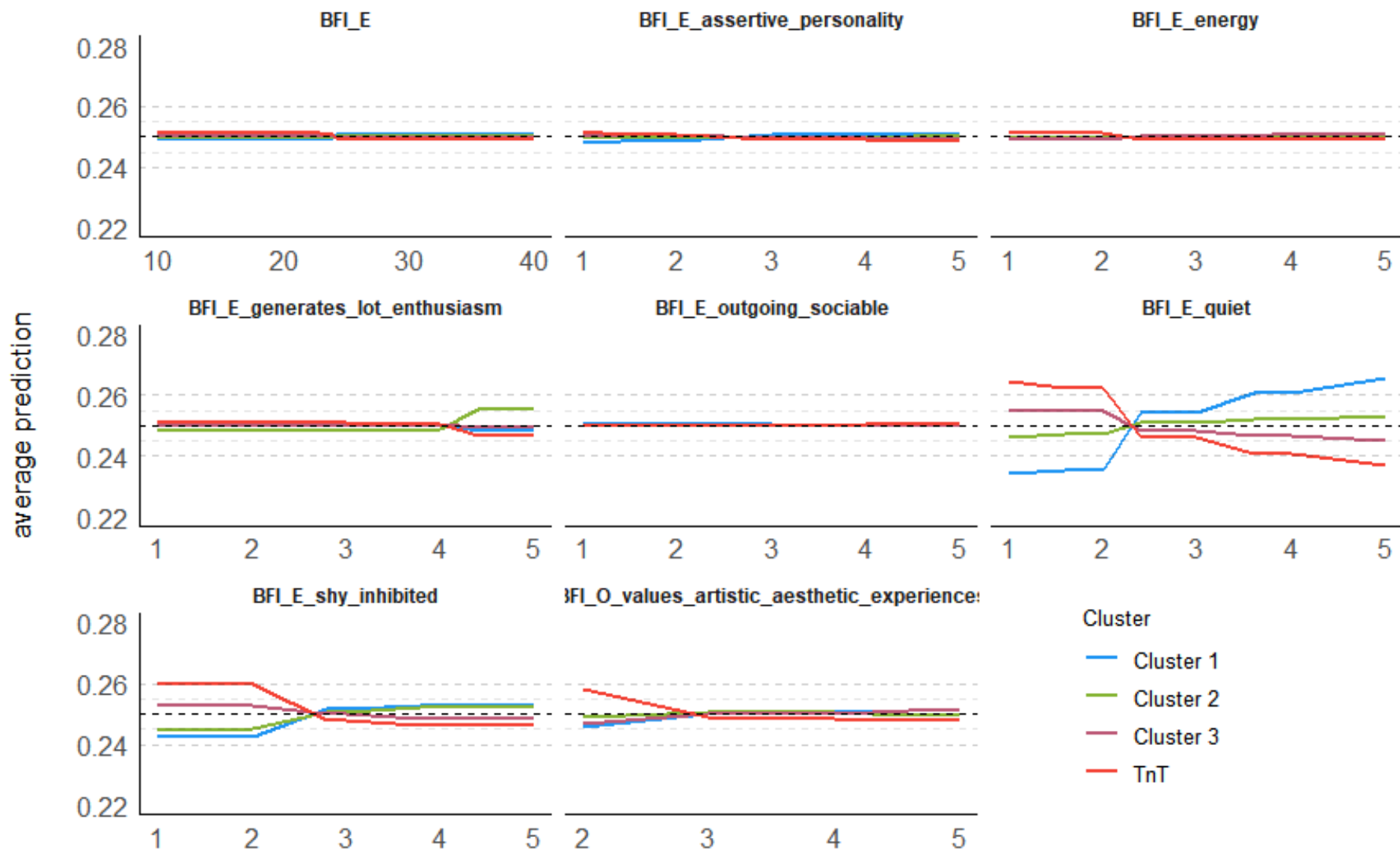


Figure I-11 Accumulated local effects (ALE) plots for personality items and factors from the Extraversion (BFI\_E) factor of the Big Five Inventory (BFI) included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that individual items are prefixed by the factors to which they contribute.

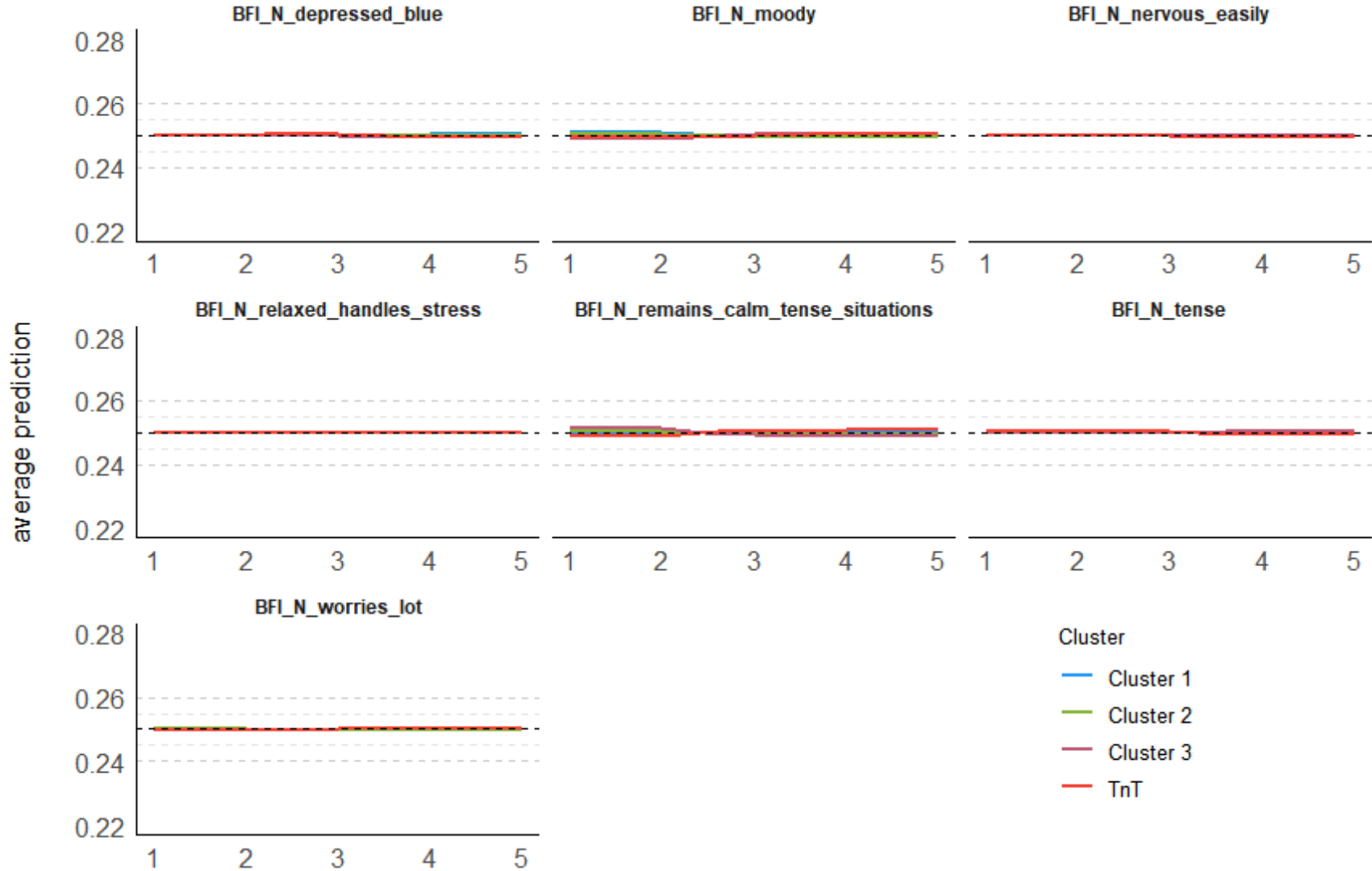


Figure I-12 Accumulated local effects (ALE) plots for personality items and factors from the Neuroticism (BFI\_N) factor of the Big Five Inventory (BFI) included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that individual items are prefixed by the factors to which they contribute.

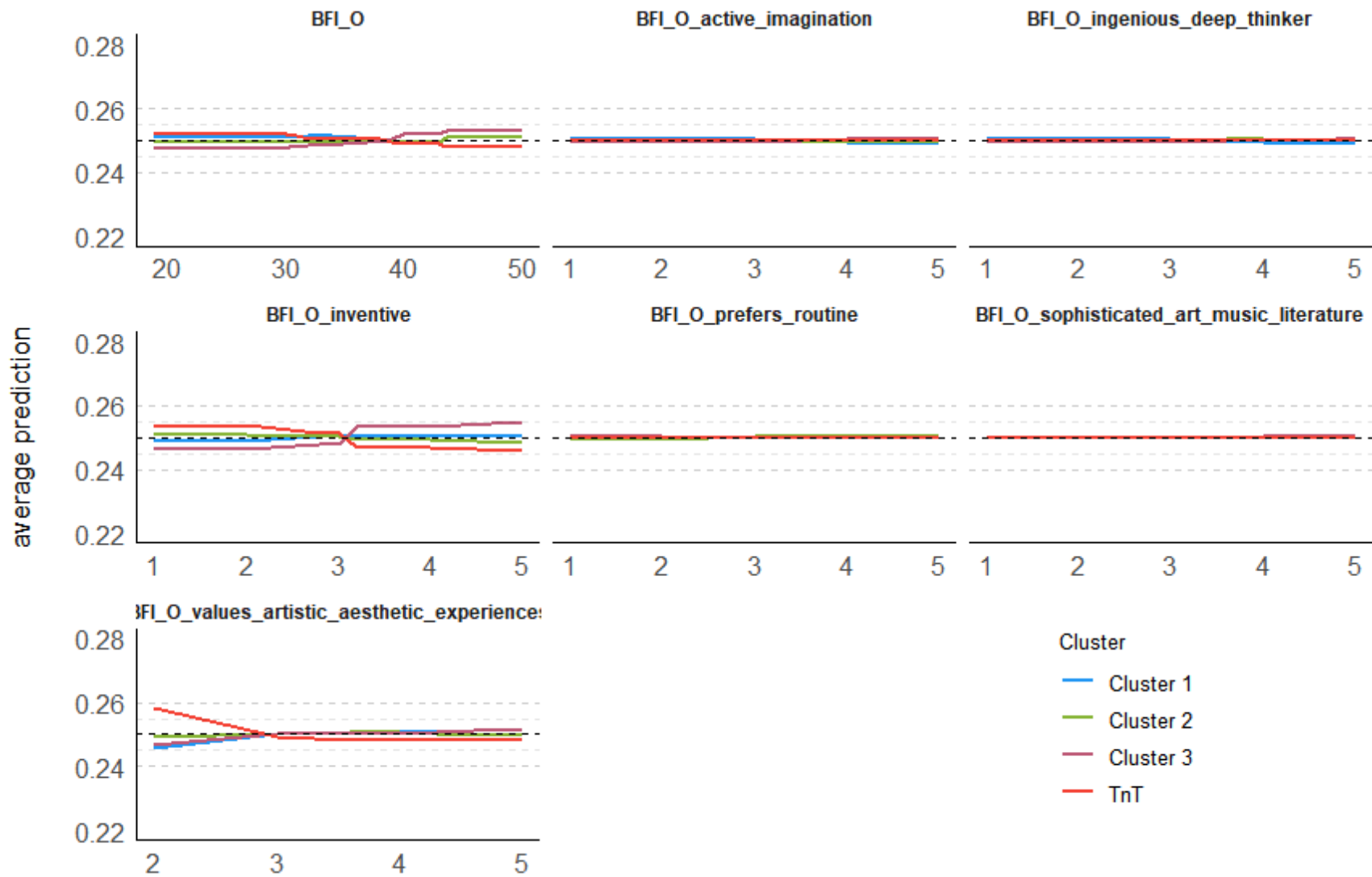


Figure I-13 Accumulated local effects (ALE) plots for personality items and factors from the Openness to Experience (BFI\_O) factor of the Big Five Inventory (BFI) included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that individual items are prefixed by the factors to which they contribute.

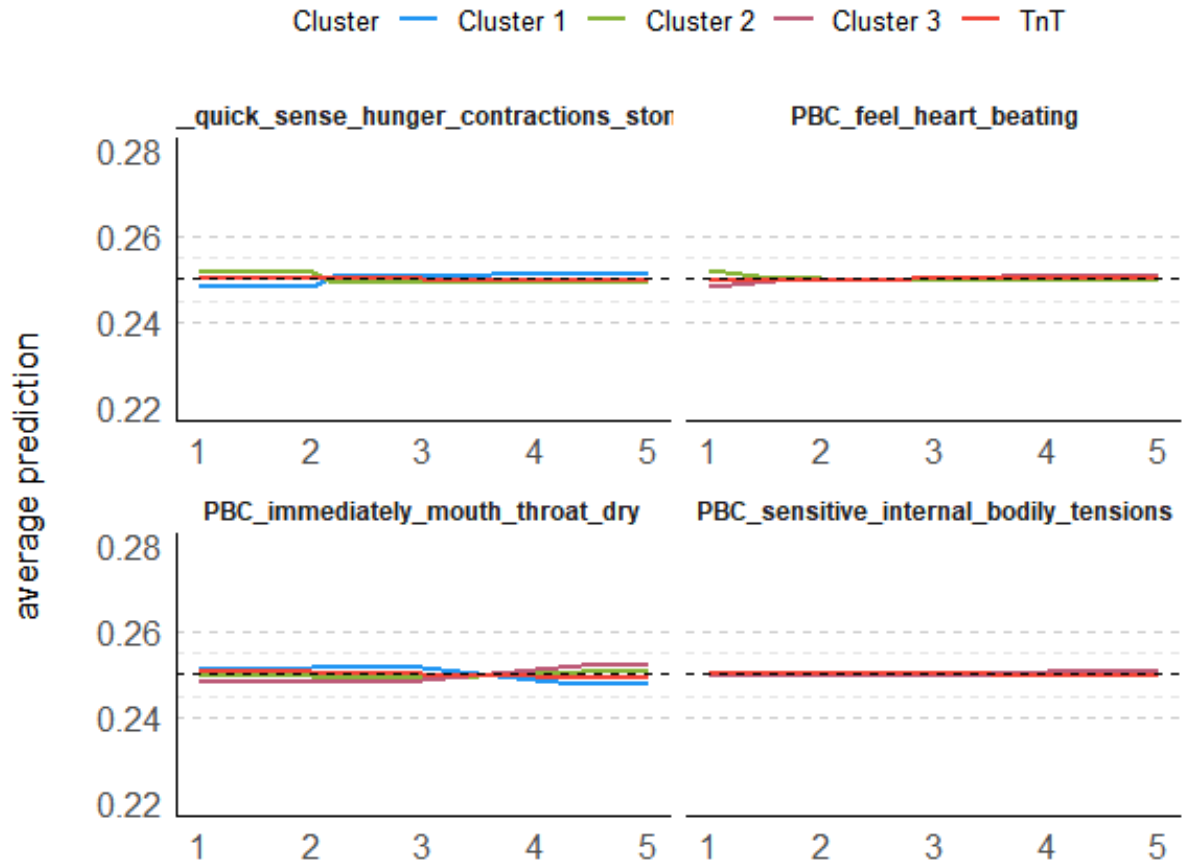


Figure I-14 Accumulated local effects (ALE) plots for items from the Private body consciousness (PBC) questionnaire included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability

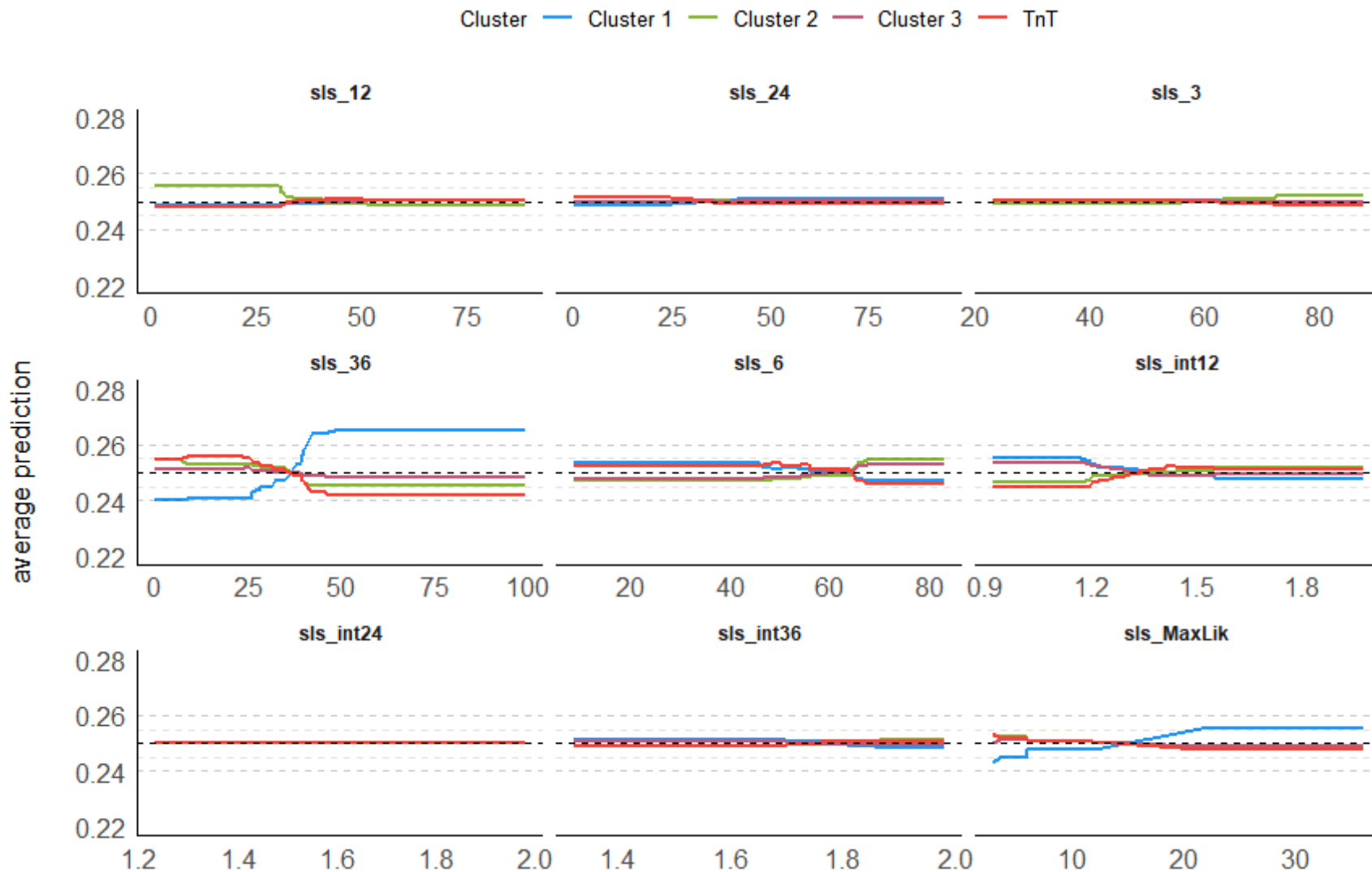


Figure I-15 Accumulated local effects (ALE) plots for sweetness response variables (liking, intensity) included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability. Note that the prefix “sls” indicates liking on the LAM scale, and the prefix “sls\_int” indicates the intensity rating of each sample on the gLMS.

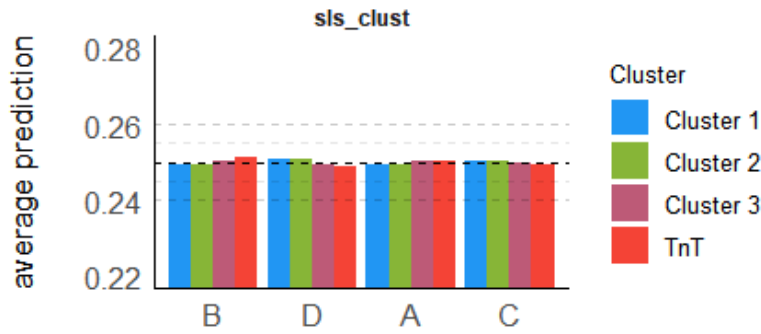


Figure I-16 Accumulated local effects (ALE) plots for SLS cluster included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm 0.01$  and chance  $\pm 0.005$ ), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

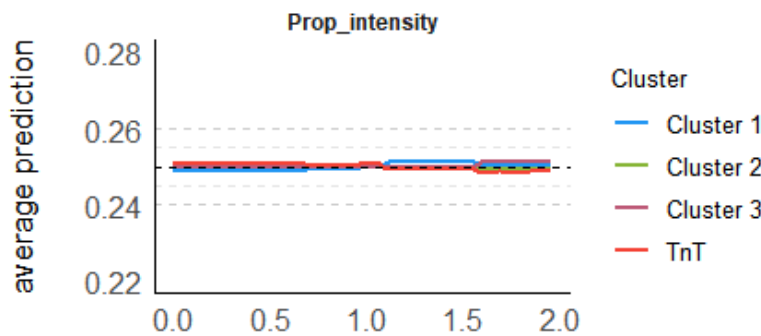
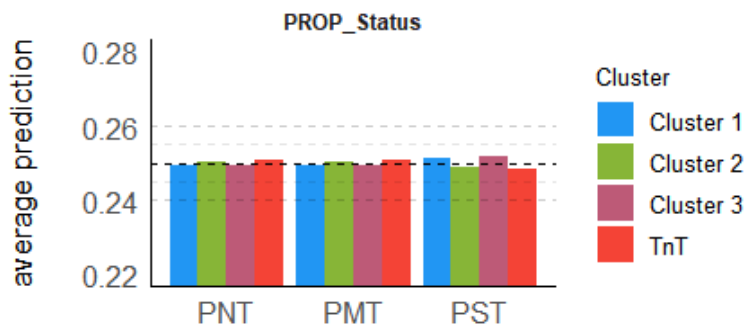


Figure I-17 Accumulated local effects (ALE) plots for PROP responsiveness variables included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm 0.01$  and chance  $\pm 0.005$ ), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

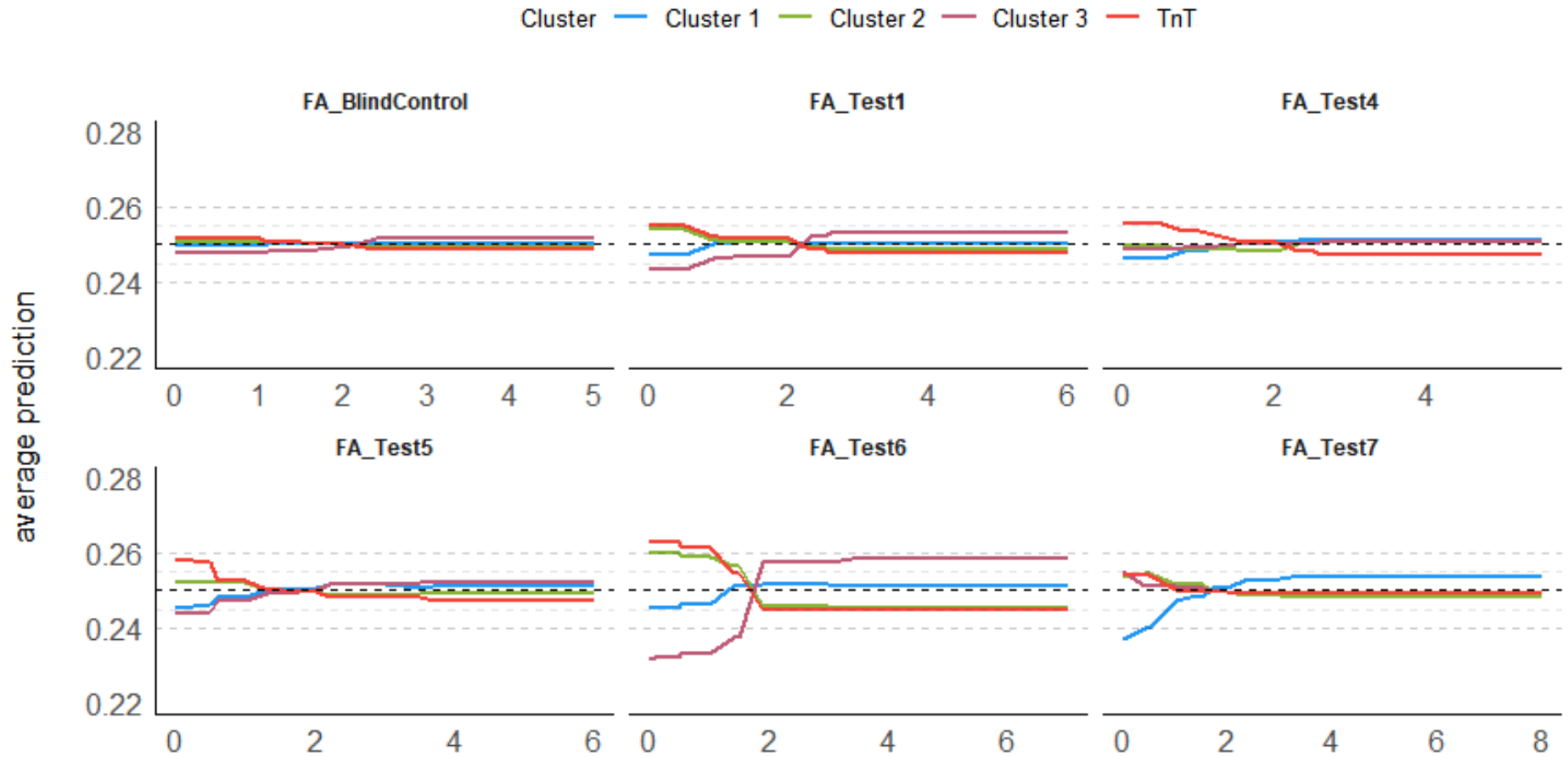


Figure I-18 Accumulated local effects (ALE) plots for Fatty acid response variables included in the final multiclass (TT cluster) RFM, showing chance probability (0.25; baseline cluster probability), four guides (chance  $\pm$  0.01 and chance  $\pm$  0.005), and the effect on cluster probability across the selected variables. Deviation from chance indicates a large impact on cluster probability, whereas following the chance line indicates little or no impact on cluster probability.

## Appendix J. Changes in explained variance ROSA model

Table J-1 All items showing a reduction in explained variance (prediction) during fitting of distinct component ROSA model when the first component resulting in a reduction was fitted

Item	Category	Comps	Explained variance (%)	Change in explained variance
Cream based dairy	Dairy	21	22.22	-1.07x10 <sup>-14</sup>
Wholemeal bread	Cereals	21	18.42	-1.07x10 <sup>-14</sup>
Biscuits	SweetsBakedGoodsMiscellaneous	21	19.54	-4.62x10 <sup>-14</sup>
Banana	Fruit	21	34.17	-4.97x10 <sup>-14</sup>
Cake/Pastries	SweetsBakedGoodsMiscellaneous	21	25.52	-5.68x10 <sup>-14</sup>
Onion, Leeks	Vegetables	21	15.20	-5.68x10 <sup>-14</sup>
Jam/Syrup	SweetsBakedGoodsMiscellaneous	21	27.79	-8.88x10 <sup>-14</sup>
Low fat cheese	Dairy	21	33.88	-1.21x10 <sup>-13</sup>
Brown Rice/Wholemeal Pasta	Cereals	21	15.81	-1.33x10 <sup>-13</sup>
Other Greens	Vegetables	21	25.65	-1.67x10 <sup>-13</sup>
White rice / Couscous	Cereals	21	29.59	-1.99x10 <sup>-13</sup>
Fats	SweetsBakedGoodsMiscellaneous	21	40.56	-2.13x10 <sup>-13</sup>
Tomatoes	Vegetables	21	30.17	-2.13x10 <sup>-13</sup>
Dried fruit	Fruit	21	22.95	-2.20x10 <sup>-13</sup>
Stone fruit	Fruit	21	40.77	-2.56x10 <sup>-13</sup>
Other Vegetables	Vegetables	21	21.02	-3.45x10 <sup>-13</sup>
Sweets, Lollies	SweetsBakedGoodsMiscellaneous	21	33.68	-4.33x10 <sup>-13</sup>
White Bread	Cereals	21	29.48	-4.33x10 <sup>-13</sup>
Apples/Pears	Fruit	21	33.33	-4.41x10 <sup>-13</sup>
Citrus fruit	Fruit	21	20.81	-4.87x10 <sup>-13</sup>
Sausages, Hotdogs	Protein	21	29.13	-4.87x10 <sup>-13</sup>
Pizza/Hamburgers	FastFood	21	26.33	-4.97x10 <sup>-13</sup>
Fried fish	Protein	21	28.58	-5.12x10 <sup>-13</sup>
Poultry	Protein	21	27.16	-7.78x10 <sup>-13</sup>
Other Cold Breakfast Cereal	Cereals	21	21.37	-9.66x10 <sup>-13</sup>
Oils	SweetsBakedGoodsMiscellaneous	21	22.04	-1.02x10 <sup>-12</sup>
Nuts	SweetsBakedGoodsMiscellaneous	21	22.47	-2.27x10 <sup>-12</sup>

## Appendix K. Contributions of datablocks distinct model

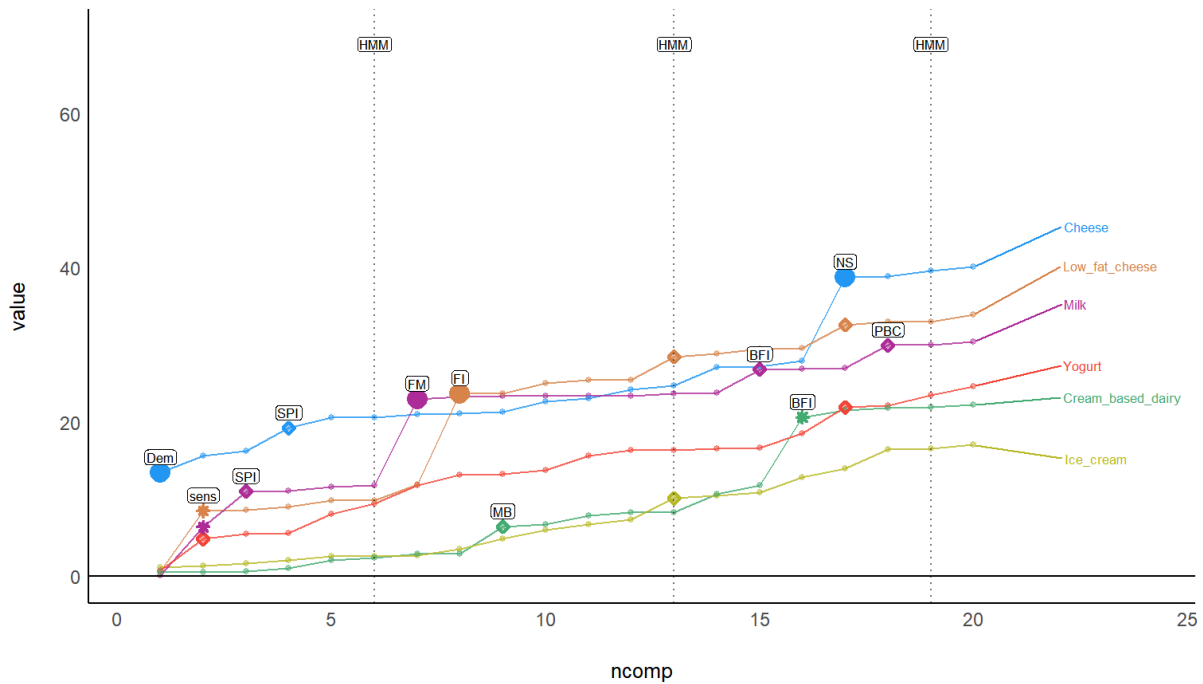


Figure K-1 Contributions of data blocks to the explained variance of items in the Dairy category using the Distinct component model. Component  $\times$  item combinations explaining a large amount of variation are labelled as follows: Greater than 5%  $\diamond$ ; Greater than 10%  $*$ ; Greater than 15%  $\bullet$ . Abbreviations: Dem: Demographics; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

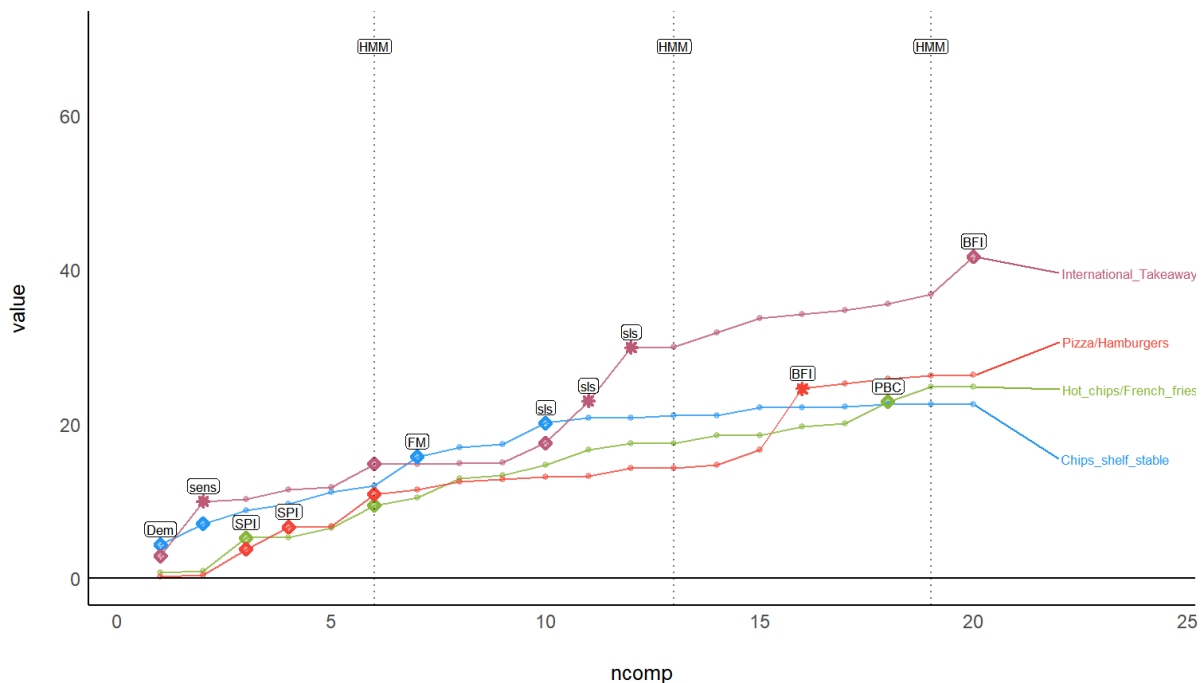


Figure K-2 Contributions of data blocks to the explained variance of items in the Fast Food category using the Distinct component model. Component  $\times$  item combinations explaining a large amount of variation are labelled as follows: Greater than 5%  $\diamond$ ; Greater than 10%  $*$ ; Greater than 15%  $\bullet$ . Abbreviations: Dem: Demographics; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

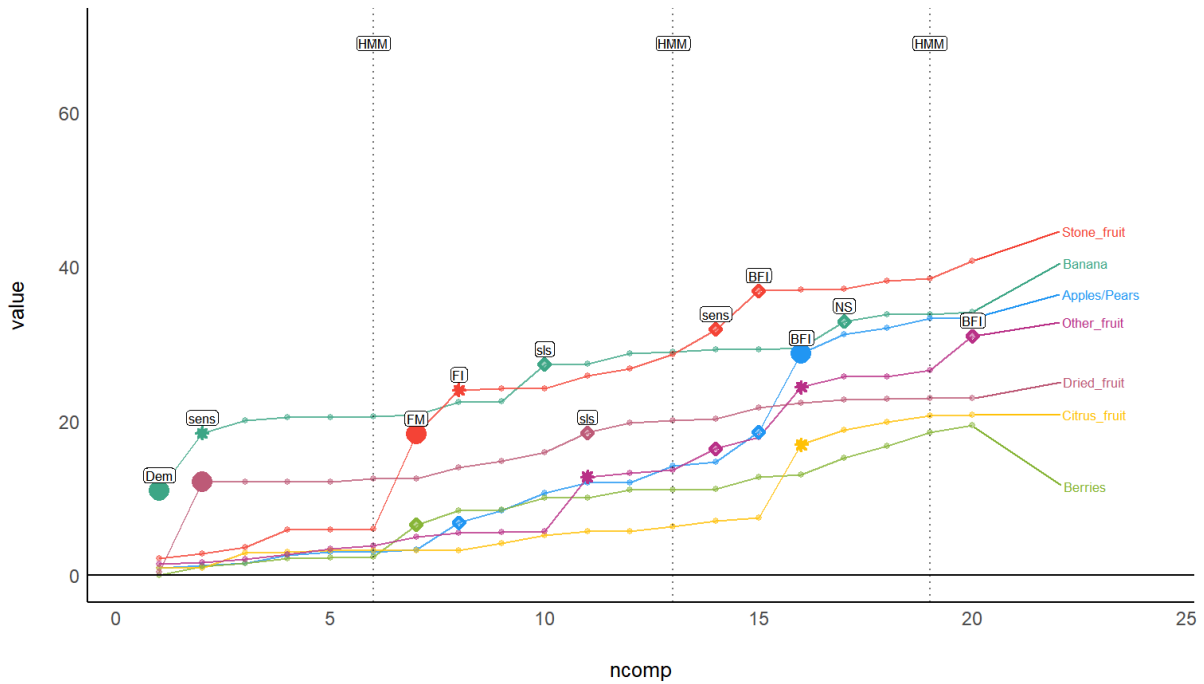


Figure K-3 Contributions of data blocks to the explained variance of items in the Fruits category using the Distinct component model. Component x item combinations explaining a large amount of variation are labelled as follows: Greater than 5%◊, Greater than 10%\*, Greater than 15%● Abbreviations Dem: Demographics; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

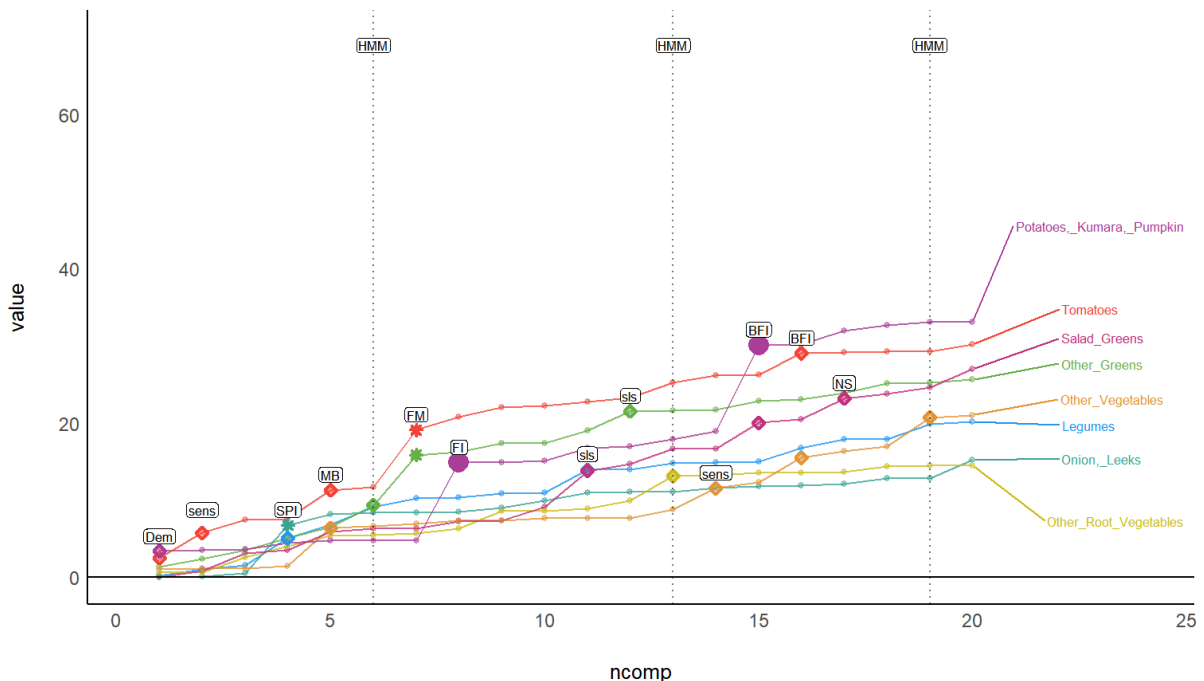


Figure K-4 Contributions of data blocks to the explained variance of items in the Vegetables category using the Distinct component model. Component x item combinations explaining a large amount of variation are labelled as follows: Greater than 5%◊, Greater than 10%\*, Greater than 15%● Abbreviations Dem: Demographics; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

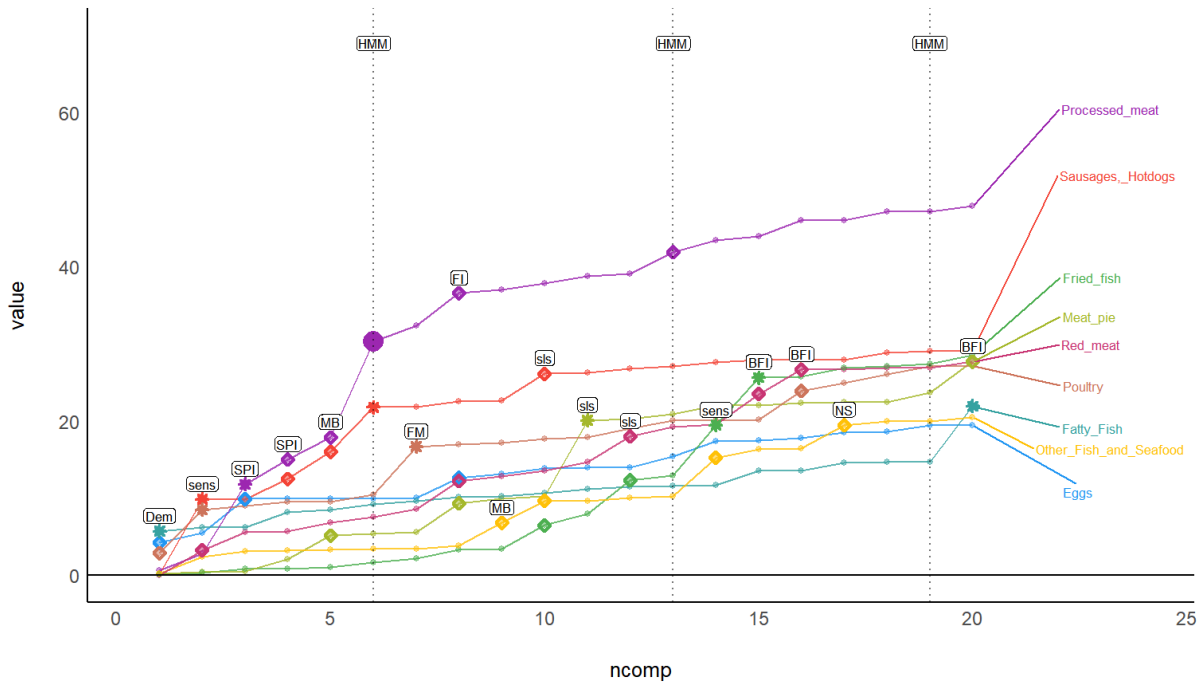


Figure K-5 Contributions of data blocks to the explained variance of items in the Proteins category using the Distinct component model. Component x item combinations explaining a large amount of variation are labelled as follows: Greater than 5%◊, Greater than 10%\*, Greater than 15%• Abbreviations: Dem: Demographics; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

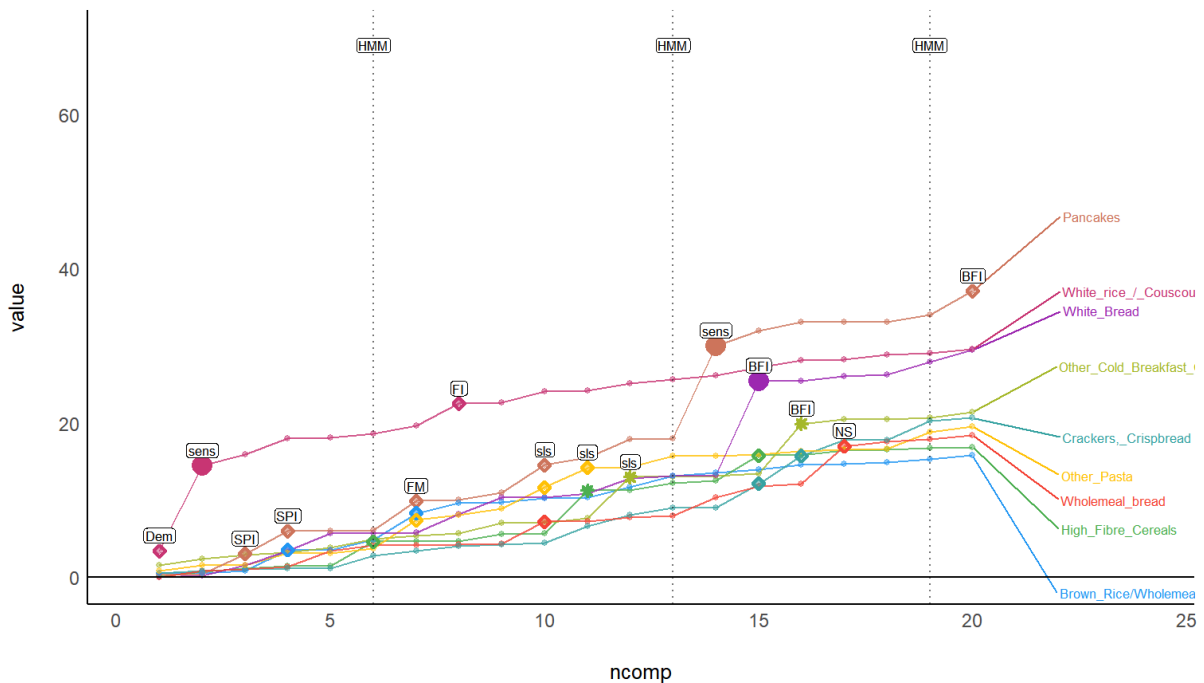


Figure K-6 Contributions of data blocks to the explained variance of items in the Cereals category using the Distinct component model. Component x item combinations explaining a large amount of variation are labelled as follows: Greater than 5%◊, Greater than 10%\*, Greater than 15%• Abbreviations: Dem: Demographics; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

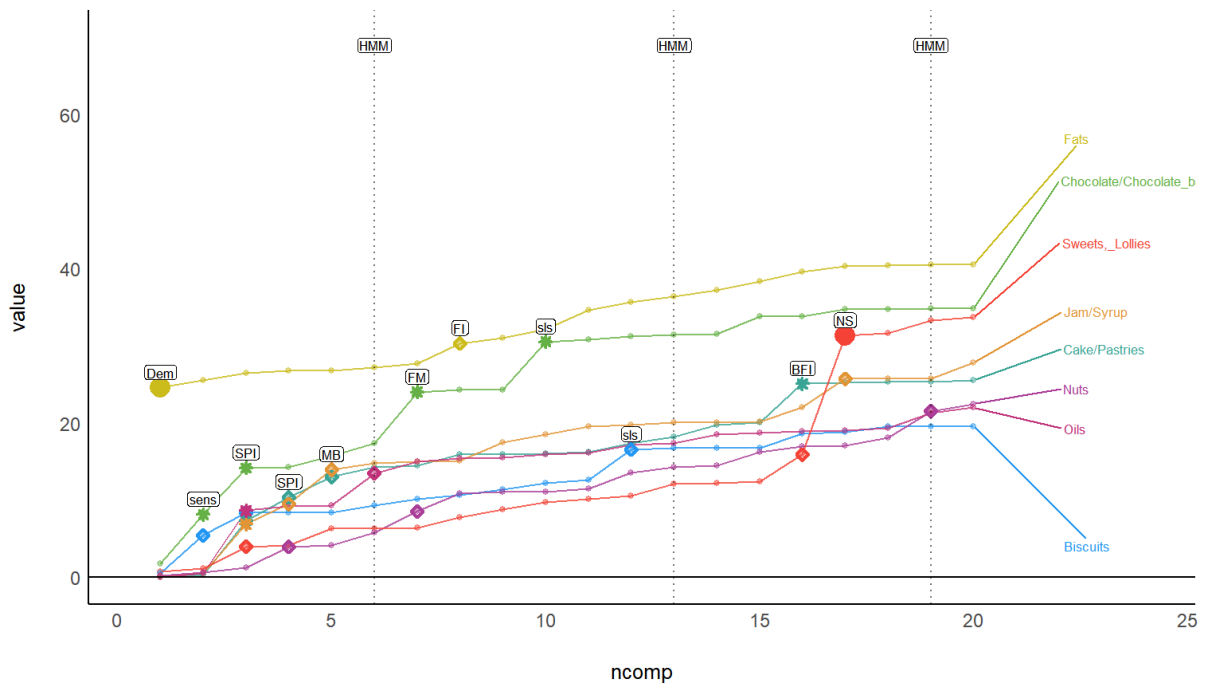


Figure K-7 Contributions of data blocks to the explained variance of items in the Sweets, baked goods, Misc category using the Distinct component model. Component  $\times$  item combinations explaining a large amount of variation are labelled as follows: Greater than 5%:  $\diamond$ , Greater than 10%:  $*$ , Greater than 15%:  $\bullet$  Abbreviations: Dem: Demographics; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

## Appendix L. Contributions of data blocks common component model

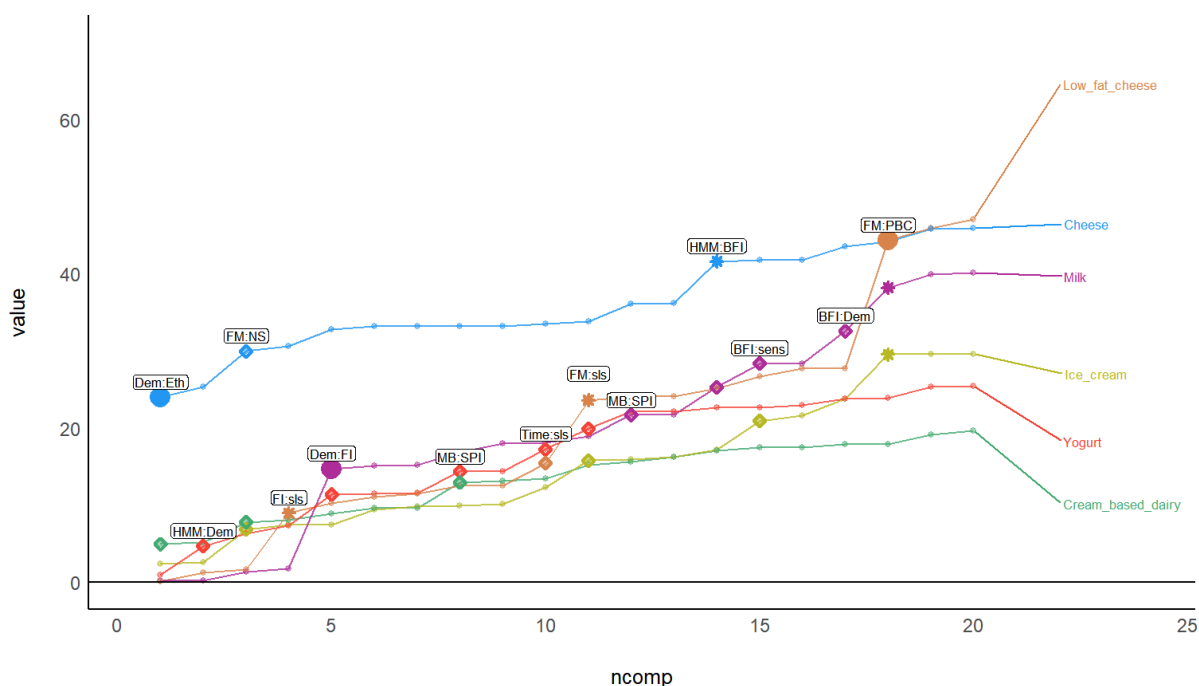


Figure L-1 Contributions of data blocks to the explained variance of items in the Dairy category using the Common component model. Component  $\times$  item combinations explaining a large amount of variation are labelled as follows: Greater than 5%  $\diamond$ ; Greater than 10%  $*$ ; Greater than 15%  $\bullet$ . Abbreviations: Dem: Demographics; Eth: Ethnicity; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

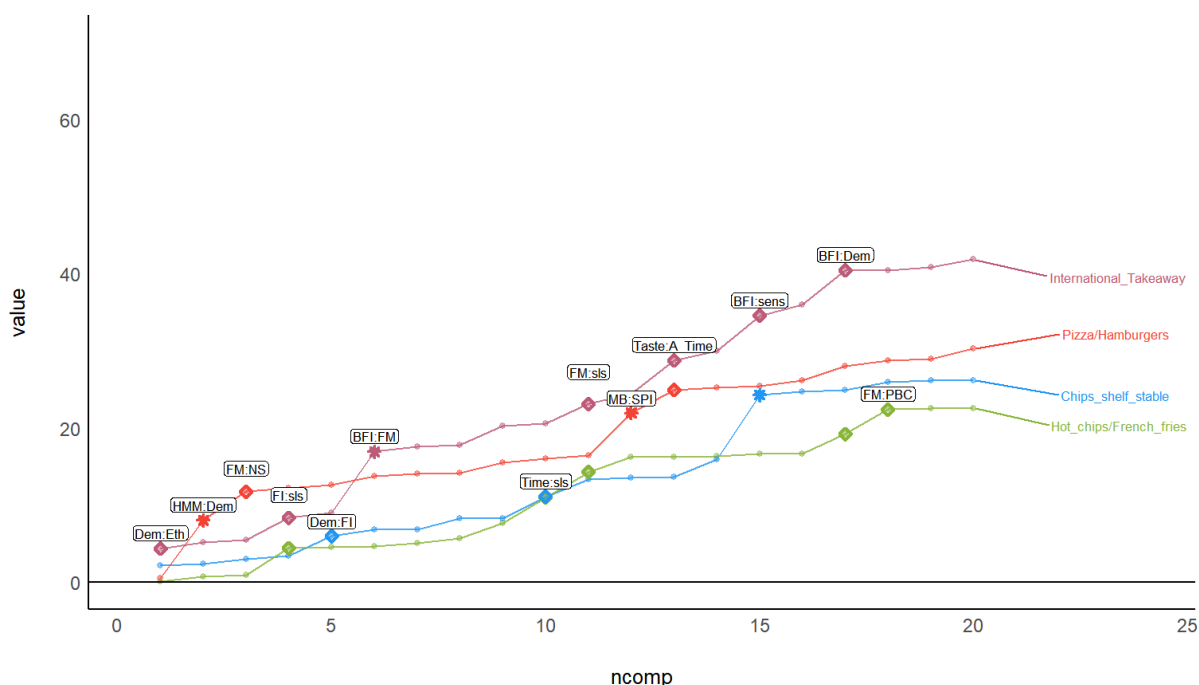


Figure L-2 Contributions of data blocks to the explained variance of items in the Fast Food category using the Common component model. Component  $\times$  item combinations explaining a large amount of variation are labelled as follows: Greater than 5%  $\diamond$ ; Greater than 10%  $*$ ; Greater than 15%  $\bullet$ . Abbreviations: Dem: Demographics; Eth: Ethnicity; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

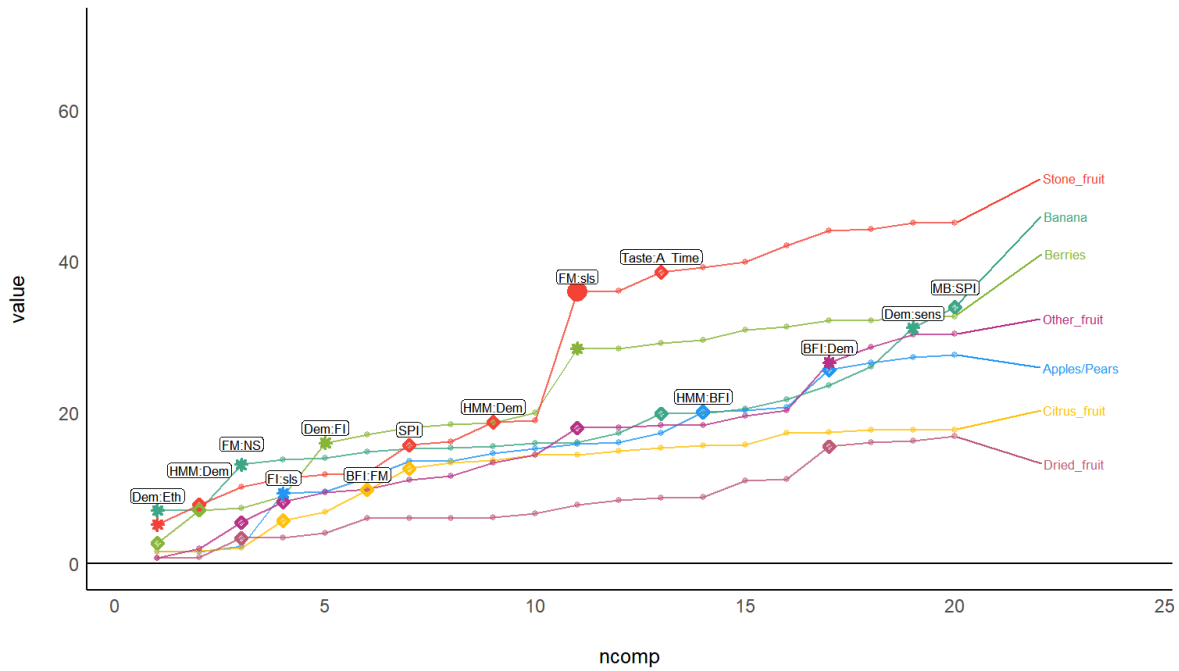


Figure L-3 Contributions of data backs to the explained variance of items in the Fruits category using the Common component model. Component x item combinations explaining a large amount of variation are labelled as follows: Greater than 5%◊, Greater than 10%\*, Greater than 15%● Abbreviations: Dem: Demographics; Eth: Ethnicity; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

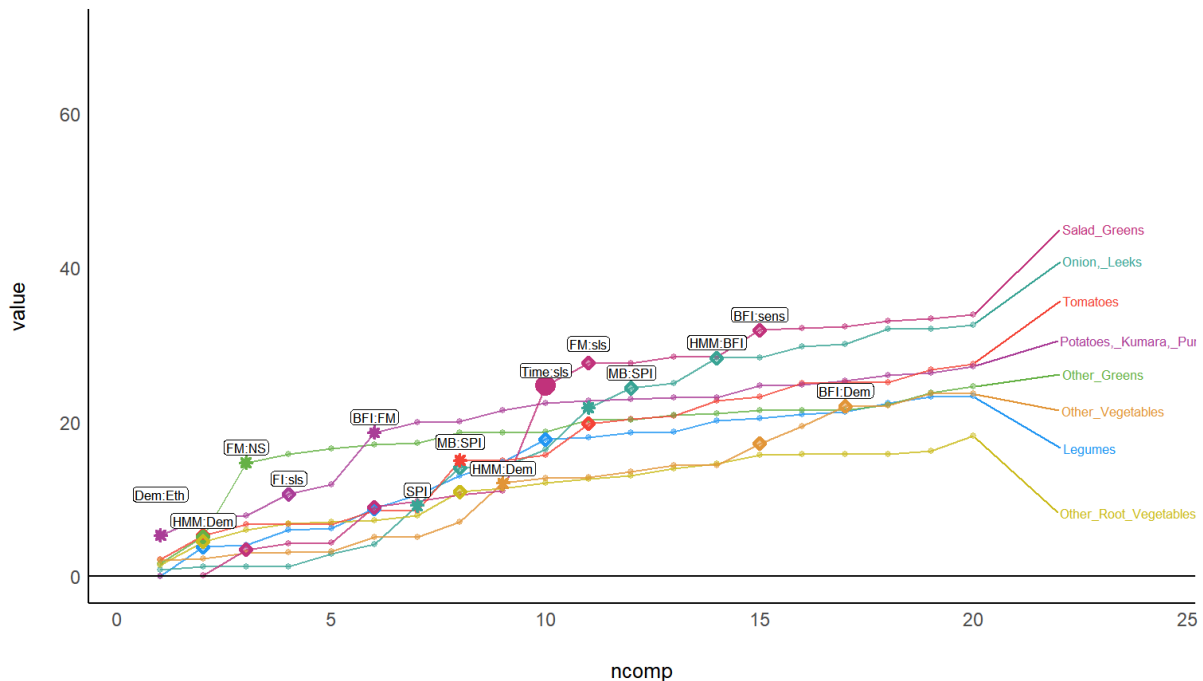


Figure L-4 Contributions of data backs to the explained variance of items in the Vegetables category using the Common component model. Component x item combinations explaining a large amount of variation are labelled as follows: Greater than 5%◊, Greater than 10%\*, Greater than 15%● Abbreviations: Dem: Demographics; Eth: Ethnicity; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking.

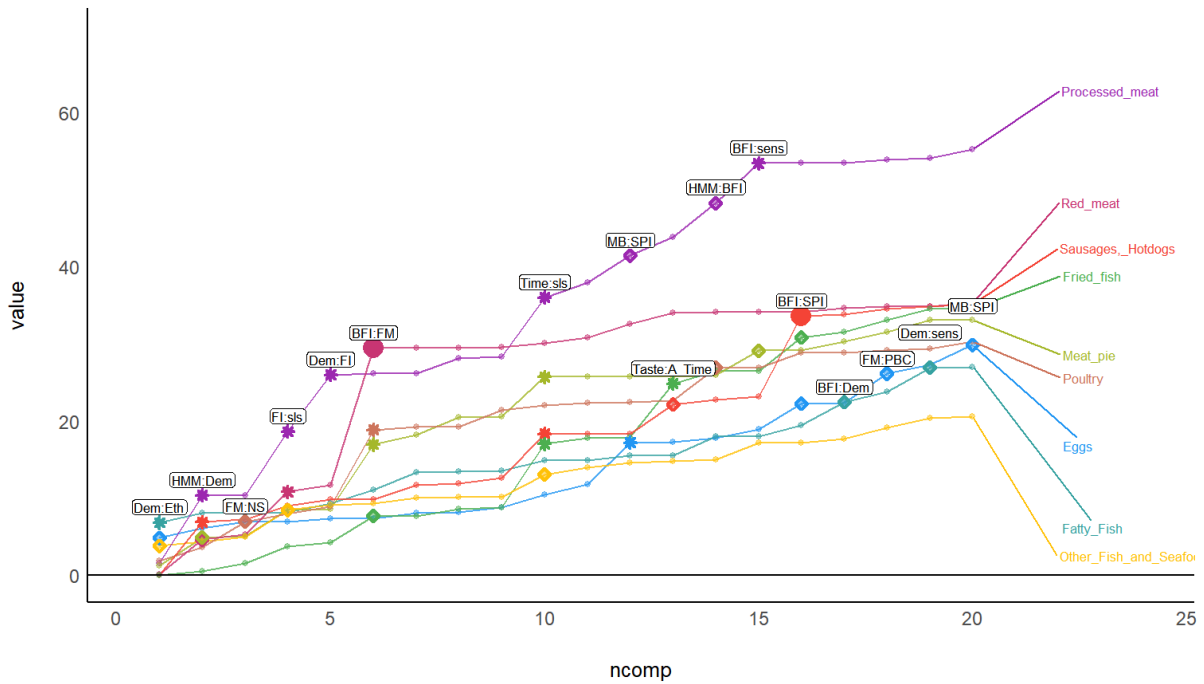


Figure L-5 Contributions of data blocks to the explained variance of items in the Proteins category using the Common component model. Component  $x$  item combinations explaining a large amount of variation are labelled as follows: Greater than 5%◊, Greater than 10%\*, Greater than 15%● Abbreviations: Dem: Demographics; Eth: Ethnicity; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking.

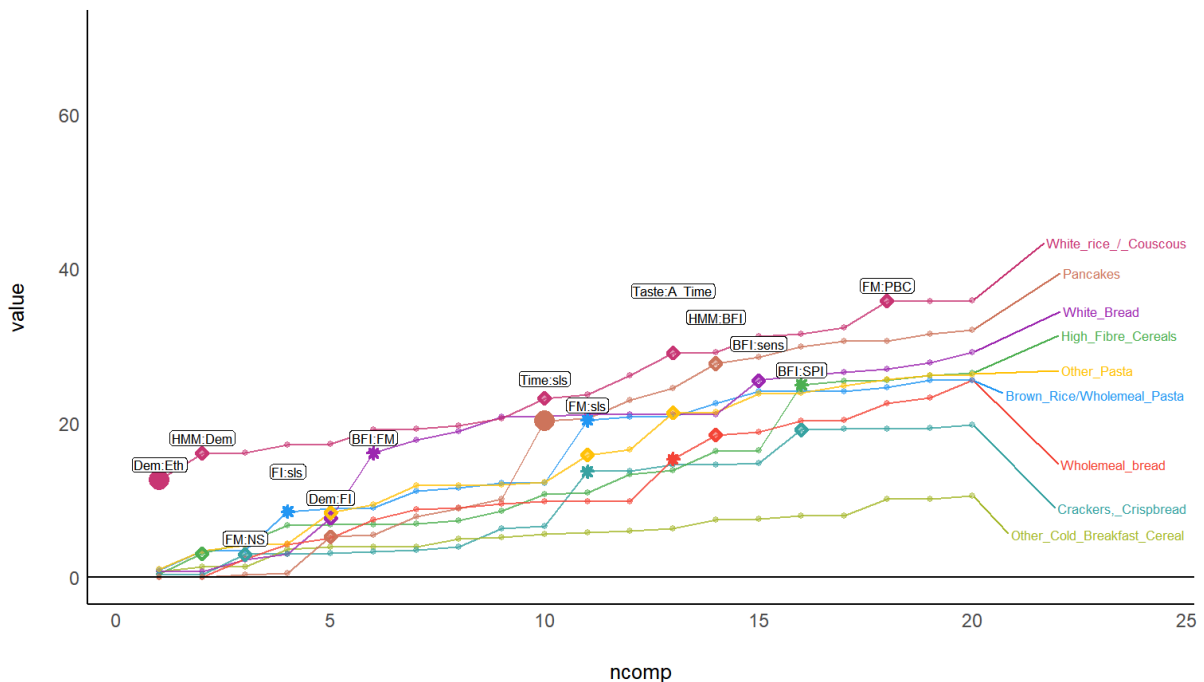


Figure L-6 Contributions of data blocks to the explained variance of items in the Cereals category using the Common component model. Component  $x$  item combinations explaining a large amount of variation are labelled as follows: Greater than 5%◊, Greater than 10%\*, Greater than 15%● Abbreviations: Dem: Demographics; Eth: Ethnicity; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking.

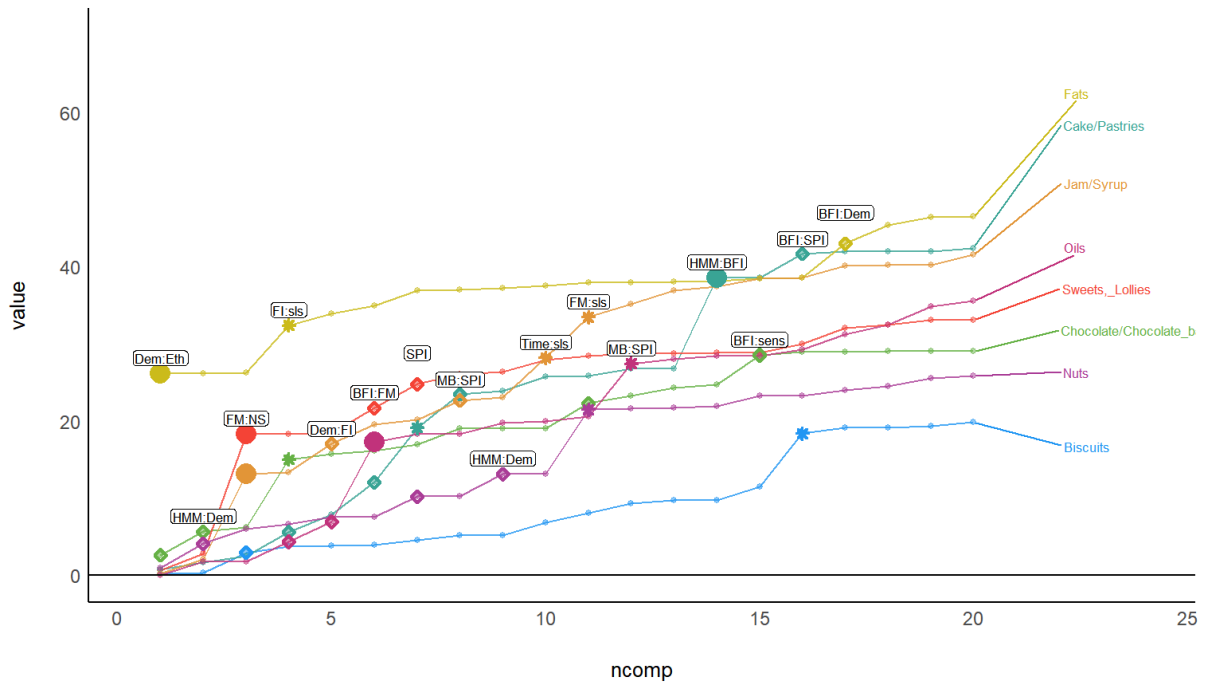


Figure L-7 Contributions of data blocks to the explained variance of items in the Sweets, baked goods, miscellaneous category using the Common component model. Component  $\times$  item combinations explaining a large amount of variation are labelled as follows: Greater than 5%:  $\diamond$ , Greater than 10%: \*, Greater than 15%:  $\bullet$ . Abbreviations: : Dem: Demographics; Eth: Ethnicity; FI: Food involvement Scale; FM: Food Modification; MB: Oral Processing; NS: Food Neophobia Scale; SPI: Spicy food liking; BFI: Big Five Inventory; PBC: Private Body Consciousness; sens: Orosensory response; Fatty Acid response; sls: Sweet Liking

## Appendix M. Variable name abbreviations

Table M-1 Variable names and abbreviations used in the ROSA model

Category	Data Block	Variable	Definition	
Demographics	Demographics	Dem_Age	Age	
		Dem_Gender_m	Gender (male)	
	Ethnicity	Eth_Asian	Asian Ethnicity	
		Eth_European	European Ethnicity	
		Eth_Maori	Maori Ethnicity	
		Eth_Middle_Eastern_Latin_American_African	Eth Middle Eastern Latin American African	
Food Behaviour	Food Involvement	FI_care_whether_not_table_nicely_set	I care whether or not a table is nicely set	
		FI_clean_up_after_eating	I do most or all of the clean up after eating	
		FI_cooking_barbequing_not_fun	Cooking or barbequing is not much fun	
		FI_dont_think_food	I don't think much about food each day	
		FI_eat_out_dont_think_talk_food	When I eat out, I don't think or talk much about how the food tastes	
		FI_enjoy_cooking_for_others_myself	I enjoy cooking for others and myself	
		FI_food_choices_not_important	Compared with other daily decisions, my food choices are not very important	
		FI_not_mix_chop_food	I do not like to mix or chop food	
		FI_not_wash_dishes_clean_table	I do not wash dishes or clean the table	
		FI_own_food_shopping	I do most or all of my own food shopping	
		FI_talking_ate_going_eat	Talking about what I ate or am going to eat is something I like to do	
		FI_travel_anticipate_eating_food	When I travel, one of the things I anticipate most is eating the food there	
		Food Modification	FM_aromatic_salt_(mixture_of_salt_and_seasoning)^meal^eating	Adding aromatic salt (mixture of salt and seasonings) to a meal while eating
			FM_chilli^meal^preparing	Adding chilli to a meal while preparing it
	FM_cream^coffee^drinking		Adding cream to coffee	
	FM_hot_chilli_sauce^meal^eating		Adding hot chilli sauce to a meal while eating	
	FM_ketchup^meal^eating		Adding ketchup to a meal while eating	
	FM_low-calorie_or_non-nutritive_sweetener^coffee^drinking		Adding low -calorie or non-nutritive sweetener to tea	
	FM_low-calorie_or_non-nutritive_sweetener^tea^drinking		Adding low -calorie or non-nutritive sweetener to tea	
	FM_milk^coffee^drinking		Adding milk to coffee	
	FM_milk^tea^drinking		Adding milk to tea	
	FM_pepper^meal^eating		Adding pepper to a meal while eating	
	FM_salt^meal^eating		Adding salt to a meal while eating	
	FM_salt^water^cooking_vegetables		Adding salt to water while cooking vegetables	
	FM_soy_sauce^meal^eating		Adding soy sauce to a meal while eating	
	Neophobia	FM_sugar^coffee^drinking	Adding sugar to coffee	
		FM_sugar_or_honey^tea^drinking	Adding sugar or honey to tea	
NS_afraid_eat_never_before		I am afraid to eat things I have never had before		
NS_at_dinner_parties_try_new		At dinner parties, I will try a new food		

Category	Data Block	Variable	Definition
Food Behaviour	Neophobia	NS_constantly_sampling_new_different	I am constantly sampling new and different foods
		NS_different_countries	I like foods from different countries
		NS_dont_know_wont_try	If I don't know what is in a food, I won't try it
		NS_dont_trust_new	I don't trust new foods
		NS_eat_almost_anything	I will eat almost anything
		NS_ethnic_looks_too_weird_eat	Ethnic food looks too weird to eat
		NS_particular_eat	I am very particular about the foods I will eat
	Oral Processing	NS_try_new_ethnic_restaurants	I like to try new ethnic restaurants.
		MB_Chewer	Chewer
		MB_Cruncher	Cruncher
		MB_Smoosher	Smoosher
	Spicy food liking	MB_Sucker	Sucker
		SPI_consump_level	Spice consumption frequency
		SPI_liking_burn	Hedonic response to burn of spices
		SPI_liking_taste	Hedonic response to taste of spices
		SPI_TF_chi_taste_better_T	I think spices make food taste better
		SPI_TF_food_bland_without_spices_T	Without hot spices, I find that food tastes bland
	Sweet liking	SPI_TF3_hard_appreciate_hot_food_T	I find it hard to appreciate the flavours of food when the food contains hot spices
		sls_12	Hedonic score 12% w/v sucrose solution
		sls_24	Hedonic score 24% w/v sucrose solution
		sls_3	Hedonic score 3% w/v sucrose solution
		sls_36	Hedonic score 36% w/v sucrose solution
		sls_6	Hedonic score 6% w/v sucrose solution
		sls_clust_A	Medium Sweet Likers
		sls_clust_C	High Sweet Likers
		sls_clust_D	Sweet liking cluster Unclassified
		sls_MaxLik	Sucrose concentration with highest liking score
Orosensory Sensitivity	Orosensory Sensitivity	Oro_FA_BlindControl	Response to 0 mM/L oleic acid (blank)
		Oro_FA_est_threshold	Estimated threshold for Oleic acid
		Oro_FA_Test1	Response to 0.5mM/L oleic acid
		Oro_FA_Test2	Response to 1mM/L oleic acid
		Oro_FA_Test3	Response to 2mM/L oleic acid
		Oro_FA_Test4	Response to 4mM/L oleic acid
		Oro_FA_Test5	Response to 8mM/L oleic acid
		Oro_FA_Test6	Response to 16mM/L oleic acid
		Oro_FA_Test7	Response to 32mM/L oleic acid
		Oro_PROP_Status	PROP Status
		Oro_slsint_12	Sweetness intensity 12%(w/v) sucrose

Category	Data Block	Variable	Definition
Orosensory Sensitivity	Orosensory Sensitivity	Oro_slsint_24	Sweetness intensity 12%(w/v) sucrose
		Oro_slsint_3	Sweetness intensity 3%(w/v) sucrose
		Oro_slsint_36	Sweetness intensity 24%(w/v) sucrose
		Oro_slsint_6	Sweetness intensity 6%(w/v) sucrose
Personality	Big Five Inventory	BFI_A_cold_aloof	Can be cold and aloof
		BFI_A_considerate	Is considerate and kind to almost everyone.
		BFI_A_fault	Tends to find fault with others.
		BFI_A_forgiving_nature	Has a forgiving nature
		BFI_A_helpful_unselfish	Is helpful and unselfish with others.
		BFI_A_likes_cooperate	Likes to cooperate with others.
		BFI_A_rude	Is sometimes rude to others.
		BFI_A_starts_quarrels	Starts quarrels with others.
		BFI_A_trusting	Is generally trusting.
		BFI_C_careless	Can be somewhat careless.
		BFI_C_disorganised	Tends to be disorganised
		BFI_C_easily_distracted	Is easily distracted.
		BFI_C_efficiently	Does things efficiently.
		BFI_C_job	Does a thorough job.
		BFI_C_lazy	Tends to be lazy.
		BFI_C_makes_plans	Makes plans and follows through with them.
		BFI_C_perseveres_task_finished	Perseveres until the task is finished.
		BFI_C_reliable_worker	Is a reliable worker.
		BFI_E_assertive_personality	Has an assertive personality.
		BFI_E_energy	Is full of energy.
		BFI_E_generates_lot_enthusiasm	Generates a lot of enthusiasm
		BFI_E_outgoing_sociable	Is outgoing, sociable.
		BFI_E_quiet	Tends to be quiet.
		BFI_E_reserved	Is reserved
BFI_E_shy_inhibited	Is sometimes shy, inhibited		
BFI_E_talkative	Is talkative.		
BFI_N_depressed_blue	Is depressed, blue.		
BFI_N_emotionally_stable_easily_upset	Is emotionally stable, not easily upset.		
BFI_N_moody	Can be moody.		
BFI_N_nervous_easily	Gets nervous easily.		
BFI_N_relaxed_handles_stress	Is relaxed, handles stress well.		
BFI_N_remains_calm_tense_situations	Remains calm in tense situations.		
BFI_N_tense	Can be tense.		
BFI_N_worries_lot	Worries a lot		
BFI_O_active_imagination	Has an active imagination		

Category	Data Block	Variable	Definition	
Personality	Big Five Inventory	BFI_O_artistic	Have few artistic interests.	
		BFI_O_curious	Is curious about many different things.	
		BFI_O_ingenious_deep_thinker	Is ingenious, a deep thinker.	
		BFI_O_inventive	Is inventive.	
		BFI_O_likes_reflect_play_ideas	Likes to reflect, play with ideas.	
		BFI_O_original_ideas	Is original, comes up with new ideas.	
		BFI_O_prefers_routine	Prefers work that is routine.	
		BFI_O_sophisticated_art_music_literature	Is sophisticated in art, music, or literature.	
	Private body consciousness	BFI_O_values_artistic_aesthetic_experiences	Values artistic, aesthetic experiences.	
		PBC__quick_sense_hunger_contractions_stomach	I am quick to sense the hunger contractions of my stomach.	
		PBC_aware_body_temperature	I'm very aware of changes in my body temperature.	
		PBC_feel_heart_beating	I can often feel my heart beating.	
		PBC_immediately_mouth_throat_dry	I know immediately when my mouth or throat gets dry.	
Thermal Taste	TT Temperature	PBC_sensitive_internal_bodily_tensions	I am sensitive to internal bodily tensions.	
		Temp_10C	Thermal taste reported at 10°C during the cooling trail portion	
		Temp_10W	Thermal taste reported at 10°C during the warming trail portion	
		Temp_15C	Thermal taste reported at 15°C during the cooling trail portion	
		Temp_15W	Thermal taste reported at 15°C during the warming trail portion	
		Temp_20C	Thermal taste reported at 20°C during the cooling trail portion	
		Temp_20W	Thermal taste reported at 20°C during the warming trail portion	
		Temp_25C	Thermal taste reported at 25°C during the cooling trail portion	
		Temp_25W	Thermal taste reported at 25°C during the warming trail portion	
		Temp_30C	Thermal taste reported at 30°C during the cooling trail portion	
		Temp_30W	Thermal taste reported at 30°C during the warming trail portion	
		Temp_35W	Thermal taste reported at 35°C during the warming trail portion	
		Temp_5C	Thermal taste reported at 5°C during the cooling trail portion	
		TT Taste	Taste_Bitter	Bitter Thermal Taste reported
			Taste_Metallic	Metallic Thermal Taste reported
	Taste_Salty		Salty Thermal Taste reported	
	Taste_Savoury		Savoury Thermal Taste reported	
	Taste_Sour		Sour Thermal Taste reported	
	Taste_Sweet		Sweet Thermal Taste reported	
	TT Cluster		HMM_Cluster_1	Thermal taste Cluster 1
			HMM_Cluster_2	Thermal taste Cluster 2
		HMM_Cluster_3	Thermal taste Cluster 3	

## Appendix N. Variable importance values

Table N-1 VIP values of all variables in the first ten selected data blocks exceeding the VIP threshold of 1, arranged by data block. Counts refer to the number of variables from a specific data block exceeding the threshold

Data block	Count	Variable	VIP value
<b>TT Cluster</b>	<b>2</b>		
		HMM_Cluster_2	19.59
		HMM_Cluster_1	1.07
<b>Demographics</b>	<b>1</b>		
		Dem_Age	8.50
<b>Orosensory response</b>	<b>12</b>		
		Oro_slsint_12	17.82
		Oro_FA_BlindControl	4.74
		Oro_slsint_36	4.60
		Oro_slsint_6	3.71
		Oro_PROP_Status	2.98
		Oro_FA_Test1	2.94
		Oro_FA_Test7	2.58
		Oro_slsint_3	2.41
		Oro_FA_Test5	2.13
		Oro_slsint_24	1.87
		Oro_FA_Test6	1.24
		Oro_FA_Test2	1.13
<b>Spicy food liking</b>	<b>6</b>		
		SPI_TF3_hard_appreciate_hot_food_T	102.00
		SPI_TF_chi_taste_better_T	11.74
		SPI_TF_food_bland_without_spices_T	5.81
		SPI_consump_level	5.48
		SPI_liking_burn	5.46
		SPI_liking_taste	1.46
<b>Oral Processing</b>	<b>3</b>		
		MB_Mouth_Behaviour_Smoosher	75.07
		MB_Mouth_Behaviour_Chewer	8.97
		MB_Mouth_Behaviour_Sucker	8.09
<b>Food Modification</b>	<b>7</b>		
		FM_pepper^meal^eating	5.30
		FM_low-calorie_or_non-nutritive_sweetener^tea^drinking	3.23
		FM_chilli^meal^preparing	2.04
		FM_low-calorie_or_non-nutritive_sweetener^coffee^drinking	1.95
		FM_milk^coffee^drinking	1.21
		FM_sugar^coffee^drinking	1.10
		FM_soy_sauce^meal^eating	1.05
<b>Food Involvement</b>	<b>5</b>		
		FI_food_choices_not_important	6.65
		FI_travel_anticipate_eating_food	4.36
		FI_eat_out_dont_think_talk_food	3.46
		FI_enjoy_cooking_for_others_myself	1.43
		FI_dont_think_food	1.02
<b>Sweet Liking</b>	<b>5</b>		
		sls_clust_C	6.10
		sls_clust_D	3.28
		sls_clust_A	2.86
		sls_36	2.25
		sls_MaxLik	1.54

Table N- 1 (cont) VIP values of all variables in data blocks not selected in the first ten exceeding the VIP threshold of 1, arranged by data block. Counts refer to the number of variables from a specific data block exceeding the threshold

<b>Data block</b>	<b>Count</b>	<b>Variable</b>	<b>VIP value</b>
<b>Big Five Inventory</b>	<b>19</b>		
		BFI_E_talkative	16.09
		BFI_N_depressed_blue	7.38
		BFI_A_rude	4.10
		BFI_C_job	3.80
		BFI_O_original_ideas	3.61
		BFI_E_energy	3.06
		BFI_O_values_artistic_aesthetic_experiences	2.96
		BFI_O_active_imagination	2.89
		BFI_E_reserved	2.43
		BFI_C_disorganised	2.14
		BFI_C_perseveres_task_finished	2.05
		BFI_A_cold_aloof	2.02
		BFI_E_shy_inhibited	1.70
		BFI_N_worries_lot	1.65
		BFI_O_likes_reflect_play_ideas	1.39
		BFI_C_lazy	1.36
		BFI_A_likes_cooperate	1.20
		BFI_A_fault	1.16
		BFI_E_quiet	1.12
<b>Neophobia</b>	<b>5</b>		
		NS_dont_trust_new	3.05
		NS_at_dinner_parties_try_new	2.59
		NS_try_new_ethnic_restaurants	1.66
		NS_dont_know_wont_try	1.52
		NS_different_countries	1.16
<b>Private Body Consciousness</b>	<b>2</b>		
		PBC_sensitive_internal_bodily_tensions	4.68
		PBC_quick_sense_hunger_contractions_stomach	1.55

## Appendix O. Additional x-loadings plots

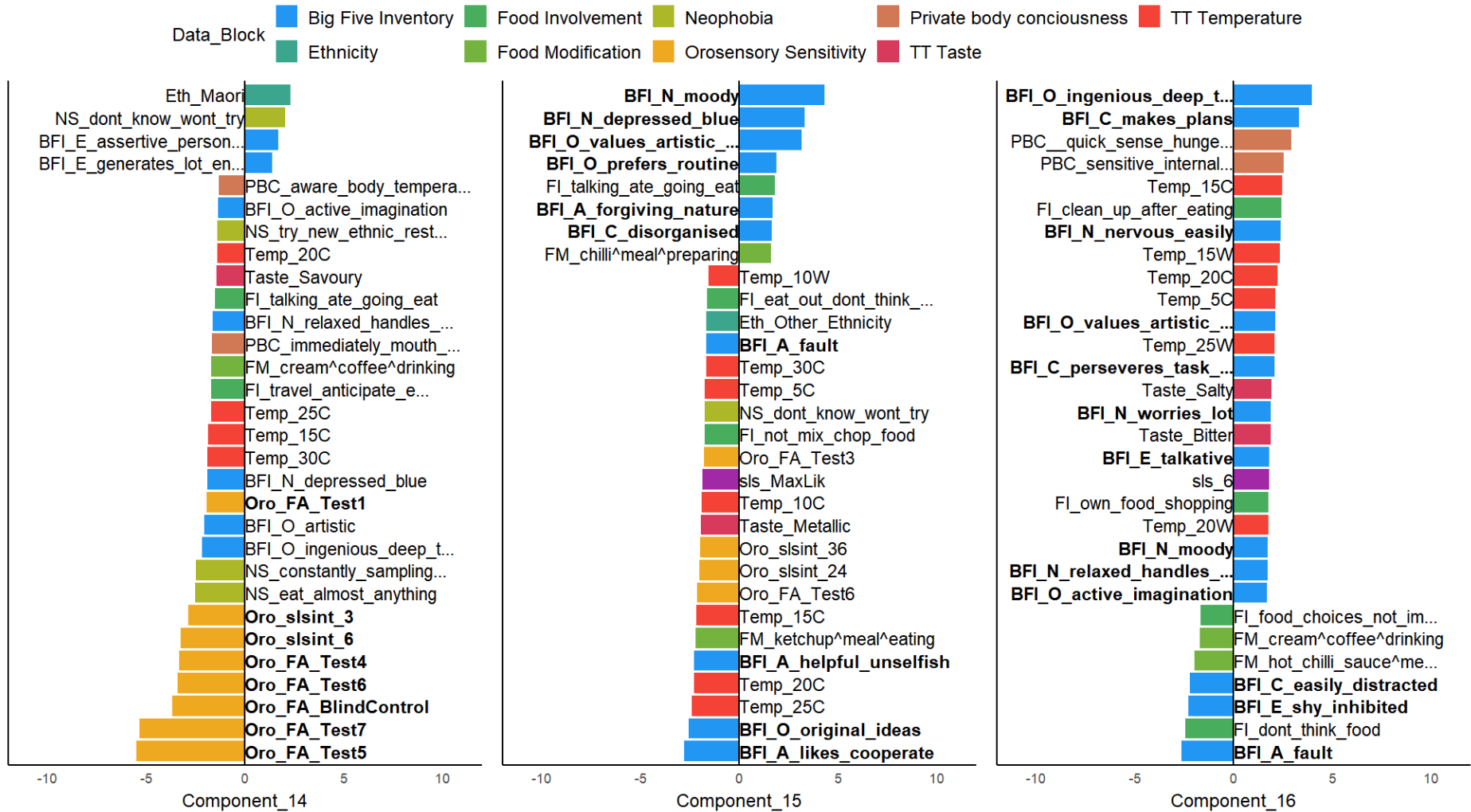


Figure O-1 Plot showing x-loadings for Component 14 (Orosensory Response), Component 15 (Big Five Inventory) and Component 16 (Big Five Inventory). Loadings with the 30 highest absolute values are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M

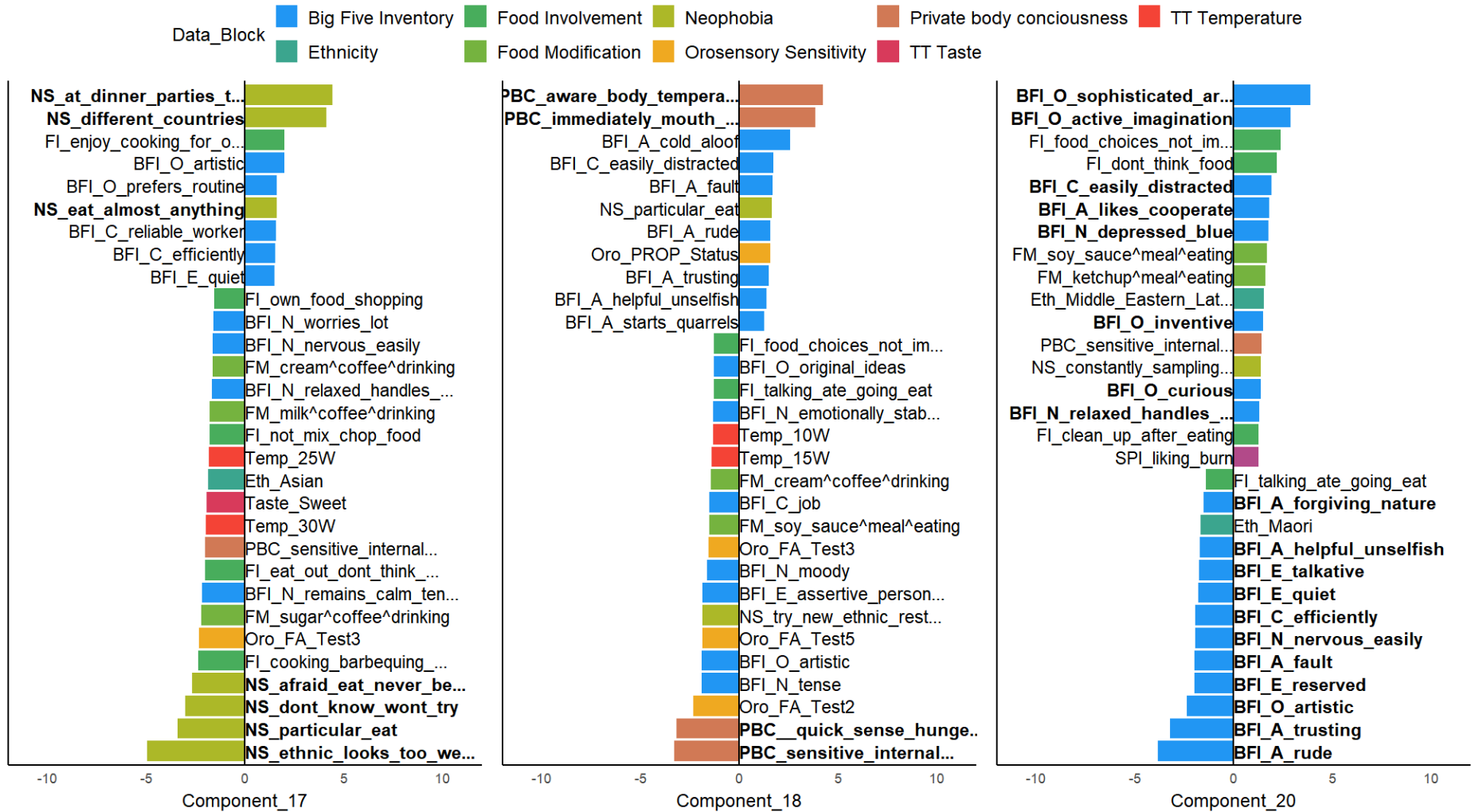


Figure O-2 Plot showing x-loadings for Component 17 (Neophobia), Component 18 (Private Body Consciousness) and Component 20 (Big Five Inventory). Loadings with the 30 highest absolute values are shown. Variables in **bold** relate to the data block from which the component was extracted. For full abbreviations of the variable names, refer to Appendix M

