

**ASSESSING THE MECHANISMS BEHIND SOUND-TASTE
CORRESPONDENCES AND THEIR IMPACT ON
MULTISENSORY FLAVOUR PERCEPTION AND EVALUATION**

QIAN (JANICE) WANG

Jesus College

University of Oxford, Oxford, UK

Supervisor: Professor CHARLES SPENCE

Dissertation submitted for the degree of Doctor of Philosophy in Experimental
Psychology at the University of Oxford, Hilary Term, 2017.

“The integration of inputs from different sensory modalities not only transforms some of their individual characteristics, but does so in ways that can enhance the quality of life.”

- B. E. Stein & M. A. Meredith, 1993, “The Merging of the Senses”, p. 11.

“No animal can live without food. Let us then pursue the corollary of this: Namely, food is about the most important influence in determining the organization of the brain and the behavior that the brain organization dictates.”

- J. Z. Young, 1968, “Influence of the mouth on the evolution of the brain”, p. 21.

Declaration

I hereby declare that this work has not been submitted previously as an exercise for a degree at this or any other university, and that it is entirely my own work.

Qian (Janice) Wang

Acknowledgments

I first embarked on this journey five years ago, for the simple reason that I ate at a very good restaurant named Alinea. It's been a difficult and often lonely road, but I wouldn't have wanted it any other way.

First of all, I want to thank my family for their unquestioning support, love, and funding, despite my often incoherent explanations of my research.

My deepest gratitude goes to my supervisor, Professor Charles Spence. Thank you for your infinite patience in tirelessly correcting all my manuscripts (in record time, no less). Your incredible work ethic, encyclopaedic knowledge, and attention to detail are truly inspirational, and I hope to be a researcher like you some day.

To my research collaborators from near and far, Klemens Knoeferle, Felipe Reinoso Carvalho, Bruno Mesz, Maria Zegna, and Steve Keller, thank you for your guidance, input, and collaboration.

To my colleagues and friends at the Crossmodal Research Laboratory and the Department of Experimental Psychology, thank you for putting up with me.

To my Crossmodal family, thank you for teaching me to play with others.

Finally, I dedicate this thesis to the Oxford University Blind Tasting Society. Two years of presidency have taught me to appreciate the beauty of wine, and reaffirmed my passion for this field of research. It has also provided me with an opportunity to meet incredible, passionate people and make lifelong friends, not to mention giving me impeccable dishwasher loading skills and a husband in the process.

The road continues.

Publications

Wang, Q. (J.), & Spence, C. (2015). Assessing the effect of musical congruency on wine tasting in a live performance setting. *i-Perception*, **6**:3.

Wang, Q. (J.), & Spence, C. (2015). Assessing the influence of the multisensory atmosphere on the taste of vodka. *Beverages*, **1**, 204-217.

Wang, Q. (J.), Woods, A., & Spence, C. (2015). “What’s your taste in music?” A comparison of the effectiveness of various soundscapes in evoking specific tastes. *i-Perception*, **6**:6.

Spence, C., & **Wang, Q. (J.)** (2015). Sonic expectations: On the sounds of opening and pouring. *Flavour*, **4**:35.

Spence, C., & **Wang, Q. (J.)** (2015). Wine & music (I): On the crossmodal matching of wine & music. *Flavour*, **4**:34.

Spence, C., & **Wang, Q. (J.)** (2015). Wine & music (II): Can you taste the music? Modulating the experience of wine through music and sound. *Flavour*, **4**:33.

Spence, C. & **Wang, Q. (J.)** (2015). Wine & music (III): So what if music influences taste? *Flavour*, **4**:36.

Wang, Q. (J.), & Spence, C. (2016). ‘Striking a sour note’: Assessing the influence of consonant and dissonant music on taste perception. *Multisensory Research*, **29**, 195-208.

Reinoso Carvalho, F., **Wang, Q. J.**, Van Ee, R., & Spence, C. (2016). The influence of soundscapes on the perception and evaluation of beers. *Food Quality and Preference*, **52**, 32-41.

Wang, Q. (J.), Wang, S., & Spence, C. (2016). “Turn up the taste”: Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and pitch. *Chemical Senses*, **41**, 345-356.

- Reinoso Carvalho, F., **Wang, Q. J.**, De Causmaecker, B., Steenhaut, K., Van Ee, R., & Spence, C. (2016) Tune that beer! Listening for the pitch of beer. *Beverages*, **2**:31.
- Reinoso Carvalho, F., **Wang, Q. J.**, Van Ee, R., Persoone, D., & Spence, C. (2017). “Smooth operator”: Music modulates the perceived creaminess, sweetness, and bitterness of chocolate. *Appetite*, **108**, 383-390.
- Wang, Q. (J.)**, & Spence, C. (2017). Assessing the role of emotional associations in mediating crossmodal correspondences between classical music and wine. *Beverages*, **3**:1.
- Wang, Q. (J.)**, Keller, S., & Spence, C. (2017). Sounds spicy: Enhancing the evaluation of piquancy by means of a customized Crossmodally congruent soundtrack. *Food Quality and Preference*, **58**, 1-9.
- Spence, C., **Wang, Q. J.**, & Youssef, J. (2017). Pairing flavours and the temporal order of tasting. *Flavour*, **6**:4.
- Wang, Q. (J.)**, Reinoso Carvalho, F., Persoone, D., & Spence, C. (2017). Assessing the effect of shape on expected and actual chocolate flavor. *Flavour*, **6**:2.
- Wang, Q. (J.)**, & Spence, C. (in press). The role of pitch and tempo in sound-temperature crossmodal correspondences. *Multisensory Research*.
- Wang, Q. (J.)**, Knoeferle, K., & Spence, C. (in press). Music to make your mouth water? Measuring the influence of sour music on salivation. *Frontiers in Psychology: Eating Behavior*.

Articles submitted

- Wang, Q. (J.)**, & Spence, C. (submitted). Assessing the influence of music on wine perception amongst wine professionals. *Food Science and Nutrition*.

Presentations

Wang, Q. (J.) (2015). *Multisensory psychology of wine*. Masterclass presentation. Oxford Wine Festival, 22nd August, Oxford, UK.

Wang, Q. (J.) (2016). *Exploring music-wine correspondences*. Talk given at Rethinking the Senses Flavour Workshop, 4th April, University of Oxford, UK.

Wang, Q. (J.) & Spence, C. (2016). *The role of pitch and tempo in sound-temperature crossmodal correspondences*. Poster presentation. Synaesthesia and Cross-Modal Perception Conference, 22nd April, Dublin, Ireland.

Wang, Q. (J.) (2016). *Multisensory influences on food flavour and liking*. Keynote talk. Sensory Evaluation of Food workshop, 26th July, Fraunhofer IVV, Freising, Germany.

Wang, Q. (J.), Keller, S., & Spence, C. (2016). *“The sound of spice”: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack*. Oral session at EuroSense 2016, 13th September, Dijon, France.

Wang, Q. (J.) (2016). *That sounds delicious: How hearing influences flavour perception in food and drinks*. Graduate seminar presentation, 3rd October, School of Agriculture, University of Buenos Aires, Argentina.

Wang, Q. (J.) (2016). *That sounds delicious: How hearing influences flavour perception in food and drinks*. Seminar presentation, 14th December, Institute of Acoustics, London, UK.

Wang, Q. (J.), Keller, S., & Spence, C. (2017). *“The sound of spice”: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack*. Oral session at ESCOM 2017, 31st July - 4th August, Ghent, Belgium.

Selected Press coverage

Cahill, H. (2015). I use music to change how food tastes. *Financial Times*. Accessed from <https://www.ft.com/content/5f109c1a-6d44-11e5-8171-ba1968cf791a>.

Escobedo, E. (2014). Sonic Seasoning con Deezer y Elyx. *GQ Mexico*. Accessed from <http://www.gq.com.mx/lujo-gq-elyx/articulos/maridaje-musical-sonic-seasoning-con-deezer-elyx-escondite/4286>.

Gray, R. (2016). Can music change the taste of chocolate? Take the test to find out how different sounds trick your tastebuds. *The Daily Mail Online*. Accessed from <http://www.dailymail.co.uk/sciencetech/article-3956566/Can-music-change-taste-chocolate-test-different-sounds-trick-tastebuds.html>.

Laube, N. (2015). Participar de una “cena emocional” creada por artistas. *La Nacion*. Accessed from <http://www.lanacion.com.ar/1833260-participar-de-una-cena-emocional-creada-por-artistas>.

Morley, K. (2017). Music makes curries taste 10pc spicier, scientists find. *The Telegraph*. Accessed from <http://www.telegraph.co.uk/news/2017/03/12/music-makes-curries-taste-10pc-spicier-scientists-find/>.

Reynaud, S. (2015). Le vin, une question d’oreille. *Le Figaro*. Accessed from <http://avis-vin.lefigaro.fr/magazine-vin/0118131-le-vin-une-question-d-oreille>.

Knapton, S. (2016). Why the sound of New Year bells could make your beer taste sour. *The Telegraph*. Accessed from <http://www.telegraph.co.uk/science/2016/12/30/sound-new-year-bells-could-make-beer-taste-sour/>.

Note for the reader

15 experiments are presented in this thesis. 14 of which are presented in the main body of the text (21 studies total), and one further experiment is presented in the Appendix. Experiment 5 was co-led by Felipe Reinoso Carvalho (Vrije Universiteit Brussel/KU Leuven, Belgium), where we jointly worked on the experiment design, data collection, analysis, and writing. The contents of this thesis are, in part, based on preprints and accepted versions of the articles included in the Publications section, for which I give full acknowledgment.

Table of Contents

Table of Contents.....	10
Short Abstract	14
Extended Abstract.....	15
1 Background and conceptual framework.....	20
1.1 Crossmodal correspondences: What are they?	21
1.1.1 The role of crossmodal correspondences in the crossmodal binding problem ...	22
1.1.2 Examples of multisensory interactions	23
1.1.3 Influences of crossmodal correspondences on human information processing and perception.....	24
1.1.4 Some properties of crossmodal correspondences.....	27
1.1.5 Theoretical accounts of crossmodal correspondences	31
1.1.6 Unimodal correspondences?.....	37
1.1.7 Interim Summary.....	38
1.2 Associations between taste/flavour and sound attributes.....	39
1.2.1 Establishing the correspondences between taste/flavour and audition.....	39
1.2.2 Modifying the tasting experience with sound and music.....	43
1.3 Multisensory influences on flavour perception.....	50
1.3.1 A unified flavour perception.....	51
1.3.2 Neural substrates underlying flavour.....	52
1.3.3 Flavour as a multisensory perceptual system.....	53
1.4 Possible mechanisms underlying auditory modulation of taste/flavour	55
1.4.1 Sensory expectations.....	56
1.4.2 Attention.....	57
1.4.3 Neurophysiological response	57
1.4.4 Emotion Mediation / Sensation Transfer.....	59
1.4.5 Response bias	59
1.5 Aims and outline of the following chapters	60
2 How are crossmodal correspondences between sound and taste formed? The emotion mediation hypothesis.....	61
2.1 Introduction.....	61
2.1.1 Emotional mediation of crossmodal correspondences.....	61
2.1.2 Measuring emotions	62
2.1.3 Taste and emotion	63
2.1.4 Pitch and intensity.....	64
2.1.5 Beyond pitch: Comparing taste soundtracks.....	65
2.2 Experiment 1A: Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and auditory pitch	67
2.2.1 Methods.....	67
2.2.2 Results and discussion.....	70
2.3 Experiment 1B: Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and auditory pitch	82
2.3.1 Methods.....	82
2.3.2 Results and discussion.....	84
2.4 Experiment 2A: A comparison of the effectiveness of various soundscapes in evoking specific tastes and the role of emotion mediation in soundscape-taste associations	94

2.4.1	Methods.....	94
2.4.2	Results and discussion.....	101
2.4.3	Additional materials.....	109
2.5	Experiment 2B: A comparison of the effectiveness of various soundscapes in evoking specific tastes and the role of emotion mediation in soundscape-taste associations.....	110
2.5.1	Methods.....	110
2.5.2	Results and discussion.....	111
2.6	General discussion and conclusions.....	116
3	Examples of the auditory modulation of taste/flavour.....	121
3.1	Introduction.....	121
3.2	Experiment 3: Assessing the influence of consonant and dissonant music on taste perception.....	127
3.2.1	Methods.....	127
3.2.2	Results and discussion.....	130
3.3	Experiment 4: Assessing the effect of musical congruency on wine tasting in a live performance setting.....	140
3.3.1	Methods.....	140
3.3.2	Results and discussion.....	142
3.4	Experiment 5: Music modulates the perceived creaminess, sweetness, and bitterness of chocolate.....	156
3.4.1	Methods.....	156
3.4.2	Results and discussion.....	163
3.5	General discussion and conclusions.....	170
4	Response bias as a possible mechanism underlying the auditory modulation of taste/flavour.....	172
4.1	Introduction.....	172
4.2	Experiment 6: The effect of auditory stimuli timing on sound-taste interactions.....	174
4.2.1	Methods.....	174
4.2.2	Results and discussion.....	176
5	Sensory expectation as a possible mechanism underlying the auditory modulation of taste/flavour.....	183
5.1	Introduction.....	183
5.2	Experiment 7A: Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack.....	185
5.2.1	Methods.....	185
5.2.2	Results and discussion.....	187
5.3	Experiment 7B: Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack.....	192
5.3.1	Methods.....	192
5.3.2	Results and discussion.....	196
5.4	Experiment 7C: Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack.....	200
5.4.1	Methods.....	200
5.4.2	Results and discussion.....	202
5.5	Experiment 7D: Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack.....	204
5.5.1	Methods.....	204
5.5.2	Results and discussion.....	206
5.6	Experiment 8: Assessing the role of expectations on auditory taste modulation of chocolate.....	210

5.6.1	Methods.....	210
5.6.2	Results and discussion.....	212
5.7	General Discussion and conclusions.....	216
6	Attention as a possible mechanism underlying the auditory modulation of taste/flavour	218
6.1	Introduction.....	218
6.2	Experiment 9: The role of task load on perceptual effects of sound-taste correspondences.....	222
6.2.1	Methods.....	222
6.2.2	Results and discussion.....	224
6.3	Experiment 10A: Analysing the impact of music on wine perception via time-based methodologies: Temporal Dominance of Sensations (TDS).....	227
6.3.1	Methods.....	227
6.3.2	Results and discussion.....	229
6.4	Experiment 10B: Analysing the impact of music on wine perception via time-based methodologies: Time Intensity (TI)	235
6.4.1	Methods.....	235
6.4.2	Results and discussion.....	236
6.5	General discussion and conclusions	242
7	Emotion as a possible mechanism underlying the auditory modulation of taste/flavour	244
7.1	Introduction.....	244
7.2	Experiment 11: Assessing the role of valence on the influence of visual and auditory stimuli on taste perception.....	246
7.2.1	Methods.....	246
7.2.2	Results and discussion.....	249
8	Physiological response as a possible mechanism underlying the auditory modulation of taste/flavour	256
8.1	Introduction.....	256
8.2	Experiment 12: Music to make your mouth water? Measuring the influence of sour music on salivation.....	258
8.2.1	Methods.....	258
8.2.2	Results and discussion.....	260
9	Individual differences in sound-taste/flavour interactions	264
9.1	Introduction.....	264
9.1.1	Tasting expertise.....	264
9.1.2	Taster status.....	266
9.2	Experiment 13A: Assessing the influence of music on wine perception amongst wine professionals.....	268
9.2.1	Methods.....	268
9.2.2	Results and discussion.....	270
9.3	Experiment 13B: Assessing the influence of music on wine perception amongst wine professionals.....	274
9.3.1	Methods.....	274
9.3.2	Results and discussion.....	276
9.4	Experiment 14: The influence of chocolate cacao content and individual differences on perceptual effects of sound-taste correspondence	278
9.4.1	Methods.....	278
9.4.2	Results and discussion.....	280
9.5	General discussion and conclusions	285

10 General discussion: Conclusions, limitations, and directions for future research.....	287
10.1 Summary of findings	287
10.2 A framework for how crossmodal correspondences influence multisensory flavour perception	291
10.3 Limitations and directions for future research	296
10.3.1 Limitations	296
10.3.2 Case study 1: Taste-congruent music in restaurants	306
10.3.3 Case study 2: Promoting healthier eating	310
10.4 General conclusions	313
REFERENCES.....	315
APPENDIX A The role of pitch and tempo in sound-temperature crossmodal correspondences	364
A.1 Introduction.....	364
A.2 Experiment 15A: Pre-study	365
A.2.1 Methods	365
A.2.2 Results and discussion.....	368
A.3 Experiment 15B: Main experiment.....	371
A.3.1 Methods.....	371
A.3.2 Results and discussion.....	373
A.4 General discussion and conclusions	376
A.5 References	380

Short Abstract

Thesis title: Assessing the mechanisms behind sound-taste correspondences and their impact on multisensory flavour perception and evaluation

By Qian (Janice) Wang, Jesus College, D.Phil. Student of Experimental Psychology, Hilary 2017.

Recent research has shown that food and beverage perception can be influenced by means of aromas, shapes, colours, and even sounds (e.g. Auvray & Spence, 2008; Spence & Piqueras-Fiszman, 2014). Over-and-above any sounds involved in the eating process, various studies have begun to demonstrate the relationship between auditory attributes and taste/flavour perception. This thesis examines crossmodal correspondences – defined as the often-surprising associations between basic attributes of different sensory modalities – between sounds and tastes/flavours. The results reported here demonstrate that participants can make non-random matches between sound and a range of taste/flavour attributes including basic tastes, creaminess, spiciness, and even temperature. Moreover, soundtracks that are congruent with specific taste/flavours are demonstrated to influence food/drink evaluation (the “sonic seasoning” phenomenon). Most importantly, this thesis highlights the multiple pathways by which sonic seasoning might occur. Mirroring the role of food-related auditory cues (such as the sizzle of the steak), soundtracks that can shape our sensory expectations before tasting as well as focus our attention on specific tastes/flavours during tasting. In addition, the emotions evoked by the soundtracks can also be transferred to the food that one happens to be eating. The results are discussed in terms of theoretical implications, practical applications, and directions for future research.

Extended Abstract

Thesis title: Assessing the mechanisms behind sound-taste correspondences and their impact on multisensory flavour perception and evaluation

By Qian (Janice) Wang, Jesus College, D.Phil. Student of Experimental Psychology, Hilary 2017.

Crossmodal correspondences are defined as the often-surprising associations between basic features of different sensory modalities that do not provide redundant identifying information about the same object/concept. The goal of the present thesis is to explore the different types of sound-taste/flavour correspondences and to try to add to the growing understanding the different mechanisms by which such correspondences might influence the tasting experience.

In Chapter 1, a general framework for crossmodal correspondences (Section 1.1) is presented, together with its reported influence on human information processing, some observed properties, and potential theories for how crossmodal correspondences are formed. Next, I review the literature related to sound-taste/flavour correspondences in particular (Section 1.2), both in terms of which correspondences have been demonstrated so far, as well as reported instances of background sounds influencing the evaluation of food/beverages (the phenomenon of “sonic seasoning”). Theories of multisensory flavour perception are presented in Section 1.3, and potential mechanisms underlying sonic seasoning are introduced in Section 1.4.

After reviewing the literature on crossmodal correspondences and multisensory flavour perception in Chapter 1, the studies reported in Chapter 2 explore the role of emotion mediation in the formation of sound-taste correspondences. The results of Experiment 1 reveal that participants’ choices of auditory pitch to match basic taste

solutions are partly mediated by their emotional valence and arousal ratings for the music and taste stimuli. In Experiment 2, the scope of the auditory stimuli is then extended from pitch to soundtracks in general; an analysis of 24 different soundtracks composed by various researchers and designers demonstrates that valence and arousal also influenced how participants matched basic taste words to soundtracks.

In Chapter 3, three experiments highlight the effects of taste/flavour-congruent soundtracks on a variety of foods/beverages. Experiment 3 reveals that varying the consonant/dissonant harmonies of a melody can alter the rated sweetness/sourness of juice samples. The results of Experiment 4 verify that people can consistently make matches between certain wines and pieces of music, and that listening to the latter can significantly influence the perceived acidity and fruitiness of a wine. Finally, Experiment 5 demonstrates that soundtracks designed to correspond with roughness and creaminess can influence the mouthfeel of chocolates.

The focus of the experiments reported in Chapters 4-8 is to assess various potential mechanisms underlying the sonic seasoning effect. Chapter 4 addresses the role of the timing of soundtrack presentation. More specifically, in Experiment 6, sweet and bitter soundtracks influenced the participants' sweet/bitter ratings of chocolates, but only when the soundtracks were played *while* tasting. When they were played *after* tasting, i.e., during evaluation, no similar crossmodal modulation of taste effects was observed. This result demonstrates that the soundtracks are genuinely influencing people's taste experience in real time, and not just acting by, say, altering their taste memories.

In Chapter 5, the ability of taste/flavour-congruent soundtracks to modify sensory expectations is assessed via two experiments. Experiment 7 provides evidence that a spicy soundtrack could enhance expectations about spiciness as well as the actual

perceived spiciness of the dish itself. Specifically, the results of Experiment 7D demonstrates that the influence of a spicy-corresponding soundtrack on participants' spiciness ratings can be explained via an assimilation-contrast model of expectation disconfirmation. Participants' spiciness ratings for salsa samples are enhanced when listening to a spicy soundtrack (compared to silence) only when the spiciness level of the salsa was fairly close to what was expected. No spiciness enhancement by music is observed for the mildly spicy salsa, possibly because the difference between the expected and actual spiciness levels was too great. The results of Experiment 8 further supports the theory of soundtracks biasing sensory expectations by showing that sweet and bitter soundtracks can influence the sweet/bitter ratings of chocolates, regardless of whether the soundtracks are heard only before tasting or only during tasting.

The role of attention is investigated in the two experiments reported in Chapter 6. In Experiment 9, cognitive load in the form of a memory task was introduced as a within-participant variable in a sonic seasoning study involving bitter and sweet soundtracks. The soundtracks influenced the participants' taste evaluation regardless of the level of cognitive load (high or low) and sound condition, which did not support the hypothesis that high cognitive load might reduce the influence of sonic seasoning. The results of Experiment 10, however, provide evidence that, as in the case of sounds orienting people's spatial attention (e.g., Chiou & Rich, 2012; Orchard-Mills et al., 2016), different soundtracks can draw people's attention to different taste/flavour attributes in a wine (Experiment 10A), as well as modify the perception of sweetness and sourness in real time (Experiment 10B).

Chapter 7 then extends Chapter 2's focus on emotion mediation between sound-taste correspondences in order to evaluate whether the emotional valence of what we hear

can influence what we taste. More specifically, Experiment 11 demonstrates that the valence of external stimuli, for both music and images, can influence participants' taste evaluation of beverages. In the experiment, a juice mixture is rated as tasting sweeter when participants are exposed to pleasant stimuli (a smiling image of a child or consonant music) as compared to unpleasant stimuli (an image of a crying child or dissonant music).

The goal of the experiment described in Chapter 8 is to investigate whether taste-congruent soundtracks can evoke the same kind of physiological reactions in the listener as the taste it is congruent with. Experiment 12 focuses specifically on the known fact that ingesting sour tastes can enhance salivary flow (Pedersen et al., 2002). Contrary to the hypothesis, however, the results of Experiment 12 do not provide evidence that a putatively sour soundtrack can increase people's rate of salivation.

Finally, the two experiments reported in Chapter 9 explore individual differences in sonic seasoning effects. Experiment 13 demonstrates that wine-tasting expertise does not moderate how music influenced wine perception, either in terms of basic taste qualities (e.g. sweet/sour) or in terms of more complex wine-tasting terminology (e.g., balance, length). In contrast, the results of Experiment 14 reveal that taster status – another source of individual variability – influenced the degree to which participants' chocolate taste ratings were influenced by bitter and sweet soundtracks. Specifically, those with higher sensitivity to phenylthiocarbamide (PTC) reported greater auditory modulation of taste with the bitter (85% cacao) chocolate, as compared to those with a lower sensitivity to PTC.

Taken together, then, the results of the various experiments reported in this thesis demonstrate that taste-congruent soundtracks can influence people's expectations

about the food that they are about to consume, shift their attention towards different taste/flavours in the food, and induce them to transfer their feelings about the soundtrack to the food they happen to be tasting. In this thesis, I hope to contribute to the understanding of multisensory tasting experiences and human information processing in general, by investigating how crossmodal correspondences can influence the way in which we perceive and integrate information from multiple sensory channels. In addition, this type of research can provide new insights for designers – both artistic and commercial – interested in the growing field of synaesthetic design, by helping them incorporate multisensory perception research into the creation of new experiences and products.

1 Background and conceptual framework

This thesis reports a series of 14 experiments designed to uncover the various ways in which people associate sound/music attributes with tastes/flavours, and the mechanisms by which such associations might influence our eating/drinking experiences. In recent years, “sonic seasoning” has started to emerge as a trend in restaurants, experience design, and marketing, whereby customized soundtracks are paired with food/beverages to putatively modify the eating/drinking experience. The experiments reported in this thesis aim to better understand the phenomenon, from the formation of sound-taste correspondences through to possible mechanisms by which such correspondences might influence taste/flavour perception in humans. The hope is that this will contribute to a better understanding of human information processing, in terms of how non-food-related auditory information can influence taste/flavour perception, as well as facilitate future innovations and the design of eating experiences.

In this introductory chapter (Chapter 1), crossmodal correspondences between taste/flavour and auditory attributes are explored and presented in a model describing how such correspondences might influence taste/flavour evaluation. Fourteen experiments (21 studies in all, as some of the experiments comprise multiple studies) are presented in the main body of the thesis (with one additional experiment presented in Appendix A). These experiments were designed to address potential emotional associations underlying sound-taste correspondences (Chapter 2), highlight the effect of sonic seasoning on a variety of foods/beverages (Chapter 3), and systematically assess different potential mechanisms and moderating factors underlying the auditory modulation of taste/flavour (Chapters

4-9). Chapter 10 then closes with a proposed model of sonic influences on taste/flavour perception, addresses a number of key caveats and limitations, and introduces directions for future research and real-world applications.

1.1 Crossmodal correspondences: What are they?

Crossmodal correspondences have been defined as the often surprising (at least, initially) crossmodal associations that people tend to share between features, attributes, or sensory dimensions in different sensory modalities (Spence 2011a). For example, most people associate brighter visual stimuli with louder sounds, higher pitched sounds with higher elevations in space and smaller objects, and sweet tastes with round shapes. Unlike synaesthesia, however, crossmodal correspondences tend to be shared across individuals, involve stimuli from different sensory modalities, and are experienced without any necessary sensory concurrent (see Deroy & Spence, 2013, for a review).

What are not considered to be crossmodal correspondences, in this case, are the cross-sensory associations that are formed based on semantic congruence, where the sensory stimuli can be seen as belonging to the same specific concept (or object). For instance, the image of a dog and the sound of a dog's bark are semantically congruent¹, since they reflect a common identity (Chen & Spence, 2010; Hein et al., 2007). However, most researchers would seem to want to distinguish such concepts from the crossmodal correspondences. An important difference between the two phenomena is that semantic congruence provides

¹ Assuming that the image of the dog roughly fits one's mental image from hearing the bark. For instance, seeing a Great Dane and hearing the bark of a Chihuahua would not necessarily be semantically congruent.

redundant identifying information about the same concept (or sometimes specific object) whereas crossmodal correspondences refer to complementary/non-redundant information, and may even lack such a common binding object/concept (Parise & Spence, 2013). For instance, the correspondence between low pitch and the aroma of dark chocolate (Crisinel & Spence, 2011a), lacks any common object/event where both of those attributes may co-occur. Furthermore, semantic congruency tends to involve more complex sensory stimuli that can help identify specific objects/concepts (e.g., the barking of a dog), whereas crossmodal correspondences typically involve (at least one) simpler sensory feature, such as visual brightness or pitch (Spence, 2011a).

1.1.1 The role of crossmodal correspondences in the crossmodal binding problem

We perceive the world with all of our senses. Most perceptual events provide multiple channels of sensory information simultaneously. The question, then, is how inputs from different sensory modalities interact and merge together to form coherent object representations (known as the crossmodal binding problem, see Ernst, 2007; Spence, 2011a).

The research demonstrates that multisensory integration is influenced by both bottom-up stimuli-based factors and top-down cognitive factors (Spence, 2007; Spence & Chen, 2010). In their seminal 1993 book, Stein and Meredith demonstrated that two sensory stimuli need to be spatially and temporally coincident in order to elicit the highest number of impulses from a multisensory neuron in the brain of a cat. From this, they proposed the rules of spatial and temporal congruency as conditions for multisensory enhancement. In humans, for

instance, it has been demonstrated that participants' sensitivity to a below-threshold visual stimulus is enhanced when a short auditory tone is presented at (approximately) the same time and location, compared to when no sound is presented (Bolognini et al., 2005; Frassinetti et al., 2002). Besides spatiotemporal congruency, top-down cognitive factors such as semantic congruency and crossmodal correspondences can also help the perceptual system to integrate information from the appropriate sensory inputs by offering prior knowledge of which stimuli go together (Spence et al., 2010; Spence, 2011a).

1.1.2 Examples of multisensory interactions

Over the years, it has been demonstrated that the factors underlying crossmodal binding can also influence multisensory interaction effects. Different sensory stimuli that are perceived at the same time/location tend to be integrated (although the importance of spatial correlation in multisensory integration is task dependent, usually for those tasks requiring an orienting response, see Spence, 2013). Another fact is that certain senses can dominate others depending on the task at hand (e.g., Botvinick & Cohen, 1998; Durgin et al., 2007; Rock & Victor, 1964). A variety of perceptual illusions have been documented showcasing the consequence of these principles. For instance, vision typically dominates over sound in the perception of spatial location; this tendency can be seen in the ventriloquist effect, where the source of an auditory stimulus is incorrectly attributed to the location of a simultaneously (or nearly simultaneously) presented visual cue (Alais & Burr, 2004; Howard & Templeton, 1966). Another instance is the McGurk effect, where people mistakenly perceive the syllable "ba" as "da" when simultaneously watching lip movements of "ga" (e.g., McGurk & MacDonald, 1976). When it comes to the temporal domain, audition is dominant over vision, such as in the

two-flash illusion, whereby observers perceive two flashes instead of one, if the visual flash is accompanied by two brief beeping sounds (e.g., Shams et al., 2000, 2002).

In addition, cognitive factors such as semantic congruency has been demonstrated to effect information processing (see Section 1.1.3 for the influence of crossmodal correspondences). Jackson (1953) first demonstrated that the spatial ventriloquism effect was extended over a larger area for semantically congruent pairs of stimuli (e.g., the sight and whistling sound of a steaming kettle) than for arbitrary pairs of stimuli (e.g., the sound of a bell and a spark)². In addition, performance in visual search tasks is facilitated when the visual target is accompanied by a semantically congruent sound (Iordanescu et al., 2008, 2010; Knöferle et al., 2016). There is also evidence that semantically congruent sounds (e.g., the sound of a barking dog) can help participants identify rapidly presented pictures (Chen & Spence, 2010, 2011a, b).

In the next section, and for the remainder of the thesis, I will focus specifically on the (relatively less studied) multisensory interaction effects pertaining to crossmodal correspondences.

1.1.3 Influences of crossmodal correspondences on human information processing and perception

A large body of empirical research has been conducted over the last half century or so regarding the impact of crossmodal correspondences on human information processing. The speeded classification task was typically used to this end, where

² An alternative interpretation, based on replication failures, suggests that temporal correlation, not semantic congruency, was the cause of the spatial ventriloquism effect since the semantically congruent pairings of audiovisual stimuli were also presented closer in time.

participants need to identify a particular characteristic of a target stimulus as quickly as possible while ignoring irrelevant stimuli (Marks, 2004). It is worth pointing out here that the majority of the speeded classification studies involved audiovisual correspondences. For instance, Bernstein and Edelman (1971) revealed a correspondence between auditory pitch and visual elevation when they found that participants responded more slowly to visual stimuli when their elevation was incongruent with the pitch of a simultaneously-presented irrelevant sound. Extending upon this work, Marks and his colleagues demonstrated that a variety of crossmodal correspondences influenced response latencies; these correspondences include pitch and size (Marks et al., 1987), pitch and lightness/darkness (Marks, 1987), pitch and angularity (Marks, 1987), pitch and brightness (Marks, 1989), and loudness and brightness (Marks, 1989). Parise and Spence (2012) used a version of the Implicit Association Test (Greenwald et al., 1998) in order to demonstrate a number of known and unknown audiovisual correspondences. Because only one stimulus was presented on each trial, the congruity effects on participants' response latencies must have been attributable to their automatic awareness of crossmodal correspondences, rather than to processes associated with selective attention or multisensory integration.

Beyond any decisional effects, there is also evidence that crossmodal correspondences influence multisensory integration as well. More specifically, the congruency of a pair of sensory stimuli can modulate spatiotemporal integration. For instance, Parise and Spence (2008) reported that participants find it more difficult to report whether a visual or auditory stimulus is presented first, when they are crossmodally congruent (in terms of the pitch-size correspondence) as compared to when the stimuli are incongruent. Furthermore, Parise and Spence

(2009, Experiment 3) found that audiovisual correspondences between pitch and size could also modulate spatial integration, where participants experienced spatial ventriloquism effects when reporting whether the auditory stimuli came from the left or right of the simultaneously presented visual stimuli.

Audiovisual correspondences can also influence human sensory perception. For instance, when asked to look at ambiguous visual motion displays, people were more likely to judge a display as moving upwards while listening to an ascending pitch, compared to a descending pitch (Maeda et al., 2004). The association between gender and pitch can even influence participants' judgements of the gender of androgynous faces (Smith et al., 2007). Participants in Smith et al.'s study were more likely to judge an androgynous face as female when the simultaneously-presented pure tone was in the female-speaking frequency range, and more likely to judge it as male when the tone was in the male-speaking frequency range. Furthermore, the correspondence between auditory pitch and visual brightness has been demonstrated to enhance performance on visual search tasks by means of top-down facilitation. Namely, auditory cues that were congruent with target brightness helped participants respond more quickly if they had been told about the crossmodal correspondence before the start of each block of experimental trials (Klapetek et al., 2012).

However, unlike the examples reported thus far, the association between pitch and spatial elevation does not seem to modulate the haptic perception of height, as reported by Geronazzo et al. (2015). In the experiment, participants had to estimate the height of a virtual stair step, rendered haptically through a Sensable PHANTOM Desktop device and explored via a stylus held in the hand. In addition, each trial was accompanied by one of four auditory feedback conditions

(two informative conditions where pitch heard is correlated with step height, random pitch, and silence). Surprisingly, the participants' height estimates (or rather, the size of the stair step), given both verbally to the nearest mm and haptically by opening the thumb and index finger, did not differ significantly under the different feedback conditions. The authors argued that this was possibly because haptic height estimation is a robust and precise procedure that does not require any redundant information supplied by audition, as demonstrated by the fact that participants estimated height accurately even in the silent control condition. In other words, crossmodal correspondences might only influence sensory perception for relatively ambiguous or difficult tasks such as judging bistable visual motion and androgynous faces.

1.1.4 Some properties of crossmodal correspondences

1.1.4.1 Bi-directionality

Bi-directionality refers to the tendency, in the area of crossmodal correspondences research, for one stimulus in a pair to be matched to the other, regardless of which stimulus is being matched. For instance, in terms of the correspondence between auditory pitch and visual elevation, Evans and Treisman (2010) conducted nine speeded classification experiments demonstrating that participants responded more quickly when the simultaneously-presented auditory and visual stimuli were congruent (i.e., high frequency tone with visual cue positioned above fixation; low frequency tone with visual cue below fixation), regardless of whether the participants were reporting on the position of the grating or the pitch of the tone. Similar results have also been demonstrated by Ben-Artzi and Marks (1995), Melara and O'Brien (1987), and Patching and Quinlan (2002). In addition, bi-

directional associations have been documented between auditory pitch and visually assessed size (Evans & Treisman, 2010), thickness (Evans & Treisman, 2010), and brightness (Marks, 1987; Melara, 1989).

Karwoski et al. (1942) first proposed a theory whereby crossmodal (or indeed, even unimodal) correspondences arise from bi-directional cross-activations between polar dimensions involving shared features³ such as brightness, thickness, and smoothness (see Section 1.1.5 for a more detailed discussion on possible mechanisms underlying crossmodal correspondences). According to this theory, all correspondences should be bi-directional. Some evidence on this point was provided by Peter Walker et al. (2016b), who demonstrated that, just as darker objects and lower pitch sounds are associated with heaviness (Walker, 2012b), heavier objects are judged to be darker and to make lower pitch sounds.

1.1.4.2 Relativity

Meanwhile, Gallace and Spence (2006) demonstrated that the crossmodal correspondence between auditory pitch and visual size is relative. The participants in their study were instructed to determine whether the second of two successively presented circles was smaller or larger than the first. At the same time, a task-irrelevant auditory tone of either 300 or 4,500 Hz was presented simultaneously with the presentation of the second circle. In their first experiment, the high and low frequency sounds were randomly presented in a mixed fashion in the same blocks. In their second experiment, the different frequency sounds were presented

³ It is worth noting here that these features are shared in the sense that people use the same linguistic term to describe sensory attributes in multiple modalities (e.g., a “bright” sound, a “bright” light). What these shared features are *not necessarily* referring to is perceiving the same environmental stimulus via different modalities (e.g., auditory brightness is unrelated to visual brightness). However, visual thickness and tactile thickness *do*, in fact, refer to the same physical attribute in the physical world.

in isolation in separate blocks of experimental trials. Gallace and Spence observed a crossmodal congruency effect induced by the correspondence between pitch and size only when the two tones appeared mixed within the same block of trials, and not when each frequency was presented in a separate block. Consequently, they reasoned that the mapping of those feature values across the two sensory modalities is relative in nature. That is, for the correspondence between pitch and size to induce a congruency effect, there needs to be an explicit comparison between multiple feature values. That said, it is worth pointing out that the two frequencies used in the study are both in the mid-range of human hearing (out of a possible 20-20,000 Hz, see Rosen, 2011, p. 163). It is certainly possible that if more extreme frequency values had been used, say, 100 and 10,000 Hz, it might have been easier for the participants to establish that one frequency was low and the other one was high (even if each frequency were to have been presented in isolation for the whole block).

Chiou and Rich (2012) found that the correspondence between pitch and spatial location is also relative, with the same sound eliciting attentional shifts in opposite directions depending on the relative pitch of a second sound. Similarly, Walker et al. (2015) demonstrated that the same grey-coloured circle can interact with the size of response keys as being either bright or dark, depending on the brightness of the other visual stimuli with which it appears. Meanwhile, Guzman-Martinez et al. (2012) have argued that the crossmodal mapping between auditory temporal amplitude-modulation (AM) rates and spatial frequency is absolute, not relative. They demonstrated via a series of experiments that participants made consistent monotonically increasing mappings between visual spatial frequency and auditory AM rates. However, this view has been contested by Orchard-Mills et al. (2013).

The latter researchers demonstrated that relative spatial frequency-auditory AM rate matches (where the auditory AM rate did not match the spatial frequency exactly, but were either relatively high or low) were sufficient to facilitate people's performance in visual search tasks. In other words, all of the crossmodal mappings discussed thus far – between pitch and spatial location, size and visual brightness, and auditory temporal AM rates and spatial frequency – would seem to be more relative than absolute.

1.1.4.3 Transitivity

Transitivity is a feature of associative learning, whereby if one stimulus (B) is associated with two different stimuli (A, C), then the latter will be associated with each other. Logically, if A implies B, and B implies C, then A implies C. In other words, two unrelated stimuli can be indirectly related to each other through a third stimulus with which they are both directly related.

The theory of transitivity was tested in a speeded classification task by Laura and Paul Walker (2012). Given that high pitch is associated with both smaller size and increased visual brightness, they predicted that smaller shapes would also be associated with brighter stimuli. Indeed, the results revealed that the participants classified bright circles more quickly when they needed to press the smaller of two buttons.

That said, there seem to be some examples of correspondences that simply do not follow this rule (note also that the above example is not an example of crossmodal correspondence *per se*, since both stimuli are presented visually). For instance, sweet fruity odours are mapped to both high pitch (Crisinel & Spence, 2011) and

round shapes (Seo et al., 2010), but high-pitched sounds are mapped to more angular – not rounded – shapes (Marks, 1987).

1.1.5 Theoretical accounts of crossmodal correspondences

A number of underlying causes have been put forward to explain the existence of crossmodal correspondences (Deroy et al., 2013; Knöferle & Spence, 2012; Spence, 2011; Walker et al., 2012). These include statistical co-occurrences, structural commonalities, semantic associations, connotative meaning, and common emotional associations. It is important to note here that, instead of there being only one mechanism which might explain all available correspondences, multiple accounts may have some explanatory power, and certain correspondences might have multiple explanatory causes.

1.1.5.1 Statistical

Certain correspondences, such as that between auditory pitch and elevation, may reflect the internalisation of the statistical regularities of the environment. For instance, Parise et al. (2014) investigated the mapping of auditory pitch to vertical elevation by measuring natural sounds from the environment (both urban and rural). Statistical associations between auditory frequency and spatial elevation were analysed from recordings, both as the sounds exist in the environment and after the sounds have been filtered by the outer ear. In both cases, there was a clear correlation between sound frequency (at least, up to 7000 Hz) and the average elevation of the sound source in the environment. Similar types of statistical learning likely explain how some aromas, such as that of vanilla or strawberry,

come to smell sweet after repeated co-exposure (Stevenson et al., 1998; Stevenson & Boakes, 2004).

Furthermore, statistical associations can involve surprisingly subtle interactions between multiple factors; for instance, Pitteri et al. (2017) recently reported that the association between spatial elevation and auditory pitch is actually based on a combination of pitch height (defined as the tone's fundamental frequency) and tone brightness (defined by the auditory spectral centroid). The participants in this study were given a speeded task involving classifying tones of different pitches using buttons arranged vertically or horizontally. When the tones varied coherently in pitch and tone brightness, a congruency effect between the pitch and vertical elevation (of the buttons) was observed. However, when pitch height and tone brightness were manipulated independently, no congruency effect was observed. Similarly, Grassi et al. (2013) found that the ability of participants to estimate the size of an object based on the sound it makes when dropped onto a plate is based not only on the loudness or frequency of the sound, but on a combination of both variables. This result further demonstrates that people's learned associations can involve multiple correlated attributes, beyond simple one-to-one feature mappings.

1.1.5.2 Structural

A second account of the origin of crossmodal correspondences is structural in nature, whereby the brain may either encode information from different sensory modalities in adjacent areas, or process them in the same way. It has been theorised that increasing stimulus intensity is coded by increased neural firing, regardless of the sensory input under consideration (Stevens, 1957; see Nehr Korn et al., 2015, for a more nuanced non-monotonic intensity coding model). As such,

this could explain the correspondence between visual brightness and auditory loudness, or taste intensity and auditory loudness, where there might be a common code for matching sensations in the way in which the brain codes information. More recently, a kind of generalised magnitude system has been proposed whereby time, space, and quantity (essentially, stimuli that vary along dimensions that can be ranked as more or less), are processed in a similar fashion in the parietal cortex (Crollen et al., 2013; Pinel et al., 2004; Walsh, 2003; though see Van Opstal & Verguts, 2013, for a critical review).

The question of whether some crossmodal correspondences might be innate is often explored in experiments involving infants. For instance, Walker et al. (2010) found that 3-4 month-old infants are sensitive to correspondences between auditory pitch and both spatial elevation and visual sharpness⁴. Based on this evidence, they speculated that those correspondences are more likely to be innate than learned because the correspondences involved most likely originated in language rather than the environment; and even if the correspondences did exist in the real world, there might not have been enough time for learning to take place. In contrast, Fernandez-Prieto et al. (2014) observed pitch-visual size correspondence effects in 6-month-old infants but not 4-month-old infants. They therefore suggested that the pitch-visual size correspondences may not be innate, but requires more experience with the environment to form the learned association. Certainly, one cannot not rule out the possibility that learning could have occurred in infants, even those as young as 3-4 months. Further studies are undoubtedly

⁴ However, note that Lewkowicz and Minar (2014) argued that infants looked longer at congruent conditions in Walker et al.'s (2010) study, not because of crossmodal congruency, but because the congruent condition was more salient and therefore attracted more attention from the infants. Walker et al. (2014) came back with a rebuttal pointing out that the auditory stimuli used by Lewkowicz and Minar differed significantly from the original in a way that undermined their attempt at replication.

needed before conclusive validation can be offered for Walker et al.'s (2010) claim that the pitch-elevation and pitch-sharpness correspondences are innate in origin.

1.1.5.3 Affective

There is growing evidence in support of the existence of affective correspondences, whereby people associate disparate stimuli from different sensory modalities based on a common feeling engendered by, or associated with, each of the component stimuli. For instance, shapes that are perceived to be more pleasant (i.e., rounded shapes) tend to be matched with tastes that are more pleasant (i.e., sweet tastes)⁵. Indeed, a number of studies have already highlighted the role of emotion in mediating crossmodal correspondences between colour and music (Barbiere et al., 2007; Bresin, 2005; Palmer et al., 2013), colour and aroma (Schifferstein & Tanudjaja, 2004), and shape and taste (Velasco et al., 2015).

1.1.5.4 Semantic

A fourth account of crossmodal correspondences is in terms of shared words or concepts used to characterise unisensory stimuli arriving from different modalities (Gallace & Spence, 2006; Walker, 2012b; Walker & Smith, 1984). Such correspondences would be expected to emerge only after language acquisition (Martino & Marks, 1999), and to influence the decisional stages of information processing rather than perception (see Spence, 2011, for a review). For instance, the same descriptor “sharp” can be used to describe a razor blade, a 10,000 Hz tone, or the flavour in an aged cheddar cheese, say. The terms “high” and “low”

⁵ The connection between sweetness and round shape can perhaps also be explained by the fact that the intensity of a sweet taste experienced in the mouth, as it changes over time, is smooth and gradual (Obrist et al., 2014).

notes are used in describing perfume (see Deroy et al., 2013, for a review), which can plausibly explain, for instance, the association between high auditory pitch and citrus, and between low pitch and musky aromas (Crisinel & Spence, 2010b). In terms of taste, the word “sweet” is not only used to describe a basic taste, but also certain aromas, such as vanilla and strawberry, and a certain musical style (“dolce”, see Fallows, s.a.). The pervasiveness of such linguistic similarities may account for many instances of crossmodal matches, but not those correspondences that are also experienced by non-human mammals (like between luminance and pitch, see Ludwig et al., 2011; and Ratcliffe et al., 2016, for a recent review).

1.1.5.5 Polarity-based associative network

Proctor and Cho (2006) suggested the principle of polarity correspondence as an explanation for the congruency effects that were first noted in speeded binary classification tasks (Section 1.1.3). The principle assumes that people code stimuli as positive or negative polarity relative to a reference point, possibly along multiple dimensions (in which case the net polarity is determined by a summation of polarity values). Correspondences are formed (or at least, congruency effects are observed in terms of more rapid classifications) when the polarity associations from the two stimuli match. For instance, the association between a relatively high-pitched tone and an up or right response, and between a relatively low-pitched tone and a down or left response, can be explained by the fact that high pitch, left, and up are all coded as positive (+) polarity, whereas low pitch, right, and down are coded as negative (-) polarity (Chang & Cho, 2015).

Similarly, Laura and Peter Walker have proposed that crossmodal correspondences can be explained via a cross-activation network⁶ of bi-directional polar dimensions encoded in different sensory modalities (L. Walker et al., 2012; P. Walker, 2012a; P. Walker & L. Walker, 2012). For example, L. Walker et al. (2012) revealed the commonality of these features by demonstrating that polar pairs of feature dimensions such as big/small, sharp/smooth, light/heavy, thin/thick can be conveyed via visual, auditory, and tactile stimuli. Moreover, these feature dimensions are aligned with each other in a consistent way, such that, for instance, smallness is also associated with sharpness, lightness, and thinness (P. Walker, 2012; see Table 1.1 for a list of aligned conceptual features). For instance, hearing a low-pitched sound can cross-activate features of thickness, heaviness, darkness, and slowness (see P. Walker, 2016a, for a review specifically relating to sound-based correspondences).

Thin	Thick
Fast	Slow
Light	Heavy
Sharp	Smooth
High	Low
Bright	Dark

⁶ Perhaps similar to the spreading activation model commonly used to represent semantic memory retrieval, where concepts are represented as nodes in a network, and associated nodes are connected by an edge whose length reflects the strength of their shared associations (Anderson, 1983; Collins & Loftus, 1975; Foster et al., 2013).

Small	Big
-------	-----

Table 1.1. The network of aligned feature dimensions (Walker, 2012). Each row denotes a pair of polar adjectives. Features in the same column are subject to cross-activation, so for instance smaller objects are also brighter, higher in elevation, and faster than large objects.

The results of various speeded reaction tasks have revealed that the same network of correspondences between these polar feature dimensions have emerged regardless of the sensory modalities used in the studies under consideration. For example, high visual angularity is associated with hardness, high pitch, and visual brightness (Parise & Spence, 2009, 2012; P. Walker 2012), large visual size is associated with brightness (P. Walker et al., 2015), and thicker letter strokes are associated with lower auditory pitch and lower visual brightness (P. Walker, 2016b).

It should, however, be noted that neither system explains how the polar associations are formed in the first place. Certainly, the task of projecting polar adjective onto a variety of sensory cues (e.g., Watt & Quinn, 2007) can be reliably performed by people, even children (Gardner, 1974). Walker et al. (2017) have proposed that the correspondences may originate either in language (Section 1.1.5.4) and/or statistical learning (Section 1.1.5.1).

1.1.6 Unimodal correspondences?

Thus far, only correspondences between stimuli of two *different* sensory modalities have been discussed. However, there is no reason to think that such correspondences cannot also take place within the same sensory modality. After all, the plausible mechanisms underlying crossmodal correspondences (statistical,

structural, affective, semantic, and polar) delineated in the previous section could be equally applicable to stimuli from the same modality. Walker (2012, Experiment 4) investigated the hypothesis that two concurrent features in the same modality can correspond if they share similar connotative meanings. In a speeded classification task, participants classified differently shaded geometric shapes as either brighter or darker than the medium-grey background. The angularity of the shapes themselves was the task-irrelevant feature that interacted with target brightness, where participants responded more quickly when the brightness of the shape was congruent with its angularity (brighter shades with more angular shape). This congruity effect demonstrated that cross-sensory correspondences do not necessarily involve two distinct sensory modalities, at least when it comes to visual brightness and angularity.

1.1.7 Interim Summary

Thus far, an overview of empirical research on crossmodal correspondences over the last century has been presented. This field of research, the vast majority of which focuses on audiovisual correspondences, has revealed the impact of such correspondences both in terms of influencing the speed/accuracy of human information processing, and (on purely a perceptual level) in terms of influencing spatiotemporal integration and sensory perception. Moreover, there is some evidence to suggest that crossmodal correspondences are typically bi-directional, transitive, and typically seem to operate in a relative manner. Thus far, researchers have offered a myriad of potential mechanisms underlying crossmodal correspondences. It is unlikely that there is a single mechanism that can explain all types of correspondences. Instead, it is more probable that there are different types of crossmodal correspondences with one or more explanatory mechanisms. Some

intriguing questions on this front are whether some correspondences might be easier to learn/acquire than others, and whether different types of correspondences will behave differently in terms of decisional or perceptual influences.

Unlike the vast majority of research devoted to audiovisual correspondences and interactions, relatively little has been studied about correspondences involving gustation. The next section focuses specifically on correspondences between sound and taste/flavour, and reports on both known associations and ways in which sound has been reported to influence taste/flavour perception in the laboratory as well as in the real world.

1.2 Associations between taste/flavour and sound attributes

1.2.1 Establishing the correspondences between taste/flavour and audition

Researchers have recently started to reveal the extensive range of crossmodal correspondences that exist between auditory attributes on the one hand, and tastes, aromas, and flavours on the other. Much of the empirical work began with single auditory tones, and gradually expanded to short musical excerpts and full-length soundtracks, which can be rated as sweet, salty, sour, or bitter depending on certain parameters of their composition, such as their pitch, articulation, tempo, loudness, etc.

The first studies to investigate the matching of sound and (putatively) flavour were conducted by Holt-Hansen (1968, 1976). He reported that participants ($N = 16$) reliably matched two types of Carlsberg beer (Carlsberg regular and Elephant Beer) to pure tones of different frequencies (the former was placed at 510-520 Hz, whereas the latter, higher alcohol beer, was matched with tones in the 640-670 Hz

range). What is more, drinking the beer while listening to the matching tone led to higher pleasantness ratings for certain of the participants (Holt-Hansen, 1976). Some years later, a follow-up study by Rudmin and Cappelli (1983) consisting of the same Carlsberg lagers, plus non-alcoholic beer, grapefruit juice, hard candy, and dill pickle revealed that participants (N=10) chose significantly higher frequencies for the grapefruit juice, candy, and pickle, as compared to the two beers. This confirmed the idea that people can consistently match different pitches to different food/drinks, and in fact, revealed that sweet and sour tastes are matched to higher pitches compared to the beers, which are presumably more bitter (see Experiment 1A and 1B).

A more rigorous series of tests involving basic tastes was conducted by Anne-Sylvie Crisinel at the Crossmodal Research Laboratory here in Oxford. Implicit Association Tests revealed an association between high pitch and sweet- and sour-food names, as well as an association between low pitch and bitter food names (Crisinel & Spence, 2009, 2010a). A potential confound here though is that participants might have matched pitches to the linguistic features of the food names themselves, rather than the (imagined) tastes of the foods. Simner et al. (2010) demonstrated that phonetic features were reliably matched to basic tastes at two different concentrations, especially with sweet tastes matching to lower values in terms of vowel height, vowel front/backness (where lower values correspond to more back in vowel space), and spectral balance compared to sour tastes.

In order to make sure that participants were matching sounds to imagined food tastes rather than of linguistic features of the food names, Crisinel and Spence (2010b) conducted another study using actual taste and aroma solutions. In the study, participants had to match each taste sample to a musical note (one of 13

notes from C2 to C6, in intervals of two tones) and a class of musical instruments (piano, strings, winds, and brass). Participants were seated in front of a virtual keyboard that allowed them to play each one of the 52 possible sounds (i.e., 4 instruments x 13 pitches) in order to find the best match. The results demonstrated that for a number of the tastes and aromas, the participants were consistent in terms of the notes and instruments that they felt went especially well together. Sweet and sour tastes were mapped to higher-pitched sounds, while bitter tastes were mapped to lower-pitched sounds. In addition, sweet tastes were mapped to piano sounds whereas bitter and sour tastes were mapped to brass instruments. In terms of aromas, fruity notes such as apricot, blackberry, and raspberry were all matched with higher (rather than lower) musical notes, and with the sounds of the piano and often also woodwind instruments, rather than with brass or string instruments. By contrast, lower pitched musical notes were associated with musky, woody, dark chocolate, and smoky aromas, bitter tastes, and brassy instruments instead.

A follow-up study (Crisinel & Spence, 2011b) was conducted using flavoured milk solutions (lemon, orange flower, vanilla, and plain) having different fat concentrations (0.1%, 1%, 2%, 4%). While the participants still made consistent associations between flavours and both pitch and instrument type, the fat content did not influence their auditory choices (possibly because a control study revealed that participants were not able to detect the difference between milks of different fat contents).

Approaching the sound-taste correspondence problem from a somewhat different angle, Mesz et al. (2011) asked 9 professional musicians to improvise freely on the theme of basic taste words (bitter, sweet, sour, and salty). The resulting

improvisations were analysed, and revealed consistent musical patterns for each taste. Bitter improvisations were low-pitched and legato, salty improvisations were staccato, sour improvisations were high pitched and dissonant, and sweet improvisations were consonant, slow, and soft. A follow-up experiment had 57 non-musicians choosing a basic taste word that best matched a subset of the improvisations. The participants performed significantly better than chance (around 68% correct, compared to chance level of 25%). Similarly, Knoeferle et al. (2015) conducted a study in which the participants matched auditory properties (pitch height, roughness, sharpness, discontinuity, tempo, sharpness, and attack) to basic taste words (sweet, sour, salty, and bitter) by using a series of sliders to control auditory properties of a short chord progression⁷.

Beyond correspondences between sounds and basic tastes, there is also evidence of reliable associations between certain pieces of music and wines. For instance, the participants in Spence et al.'s (2013) study found that Domaine Didier Dagueneau's Pouilly Fumé Silex (a crisp, mineral white wine) matched significantly better with Mozart's Flute Quartet in D major - Movement 1 than with Tchaikovsky's String Quartet No 1 - Movement 2; but the exact opposite result was observed for a glass of Chateau Margaux (a complex, balanced red wine). Furthermore, in a subsequent experiment, tasting the wines while listening to matching music resulted in a small but significant increase in the participants' wine enjoyment ratings, as compared to tasting the same wines in silence.

On the topic of wine and music pairing, Wang and Spence (2016) recently reported a study in which that a high acid and tannic Montefalco Rosso wine was

⁷ Loudness was controlled for all auditory stimuli to avoid confounding volume levels with auditory properties like pitch (Suzuki & Takeshima, 2004).

rated to match better with a rousing selection of Carmina Burana than Dussek's soothing Harp sonata. In contrast, a fruity and soft Australian Blind Spot Grenach-Syrah-Mouvedre blend was rated to match better with Rachmaninoff's Romance than Dessek's Harp Sonata. Moreover, they found evidence that wine-music matching is partly mediated by the emotional ratings of the wine and the music (see the affective account of crossmodal correspondence origins in Section 1.1.5.3).

1.2.2 Modifying the tasting experience with sound and music

It is important to note that there are a number of ways in which what we hear could potentially influence what we think about what we taste (see Table 1.2). There could, for example, be a generalized effect of music, such that classical music would make diners spend more and stay longer at restaurants (North et al., 2003; North & Hargreaves, 1998; Wilson, 2003). It has been demonstrated that ethnically appropriate music can help to bring out the ethnicity in a dish of food⁸ (Yeoh & North, 2010), and change people's food choice and shopping behaviour (e.g., North et al., 1998, 2003; Zellner et al., 2017). A growing body of scientific evidence also demonstrates that modifying the sound of mastication can influence the perceived texture of crunchy foods (Demattè et al., 2014; Zampini & Spence, 2004)⁹, and sounds that are too loud can lead to a crossmodal suppression of taste and alcohol perception (e.g., Stafford et al., 2012, 2013; see Spence, 2014b, for a review).

⁸ Although ethnic music does not necessarily influence people's liking for matching ethnic dishes (Zellner et al., 2017).

⁹ But amplifying high frequency components of carbonated water samples does not change the degree of perceived carbonation in the mouth (Zampini & Spence, 2005).

Auditory source	Sound	Effect	Study
Sound of mastication	Modified chewing sounds	Crunchier texture	Demattè et al. (2014); Endo et al. (2016); Zampini & Spence (2004)
Background noise	White noise	Lower sweetness and saltiness, higher crunchiness	Woods et al. (2011)
	Airplane cabin	Decreased sweet sensitivity, increased umami sensitivity	Yan & Dando (2014)
	News stories + Hardcore music	Impaired alcohol discrimination	Stafford et al. (2012), (2013)
Music	Ethnic music (Malay vs. Indian; Spanish vs. Italian)	Food choice	Yeoh & North (2010); Zellner et al. (2017)
	Classical music, pop music, silence	Increased spending	North et al. (2003)
	Music of different styles (e.g., powerful and heavy, subtle and refined)	Altered ratings (e.g., powerful and heavy, subtle and refined)	North (2012)
	Taste-congruent soundtracks	Altered taste evaluation	Crisinel et al. (2012); Reinoso Carvalho et al. (2015b); Spence et

al. (2013, 2014);
Velasco et al. (2013);
Wang & Spence
(2015b)

Table 1.2. Summary of some of the ways in which sound affects taste.

Although there is a large and growing body of literature documenting the various ways in which what we hear, specifically music, influences our general experience and evaluation of food and drink (see North & Hargreaves, 2008; Spence, 2015a; Spence & Shankar, 2010 for reviews), the focus in this thesis, specifically, is the question of whether music that in some way corresponds crossmodally with specific tastes/flavours can help to bring out specific sensory attributes in the tasting experience.

1.2.2.1 Experimental results

It was Crisinel et al. (2012) who first demonstrated that, over and above any crossmodal associations (or correspondences) between sounds and taste words, auditory stimuli might have a genuinely perceptual effect on taste/ flavour evaluation. In their study, participants were given samples of bittersweet cinder toffee to evaluate while listening to one of two soundscapes. The soundscapes were composed to match either sweetness or bitterness. Listening to the higher pitched sweet soundscape resulted in the toffee being rated as tasting significantly sweeter and less bitter than while listening to the lower-pitched bitter soundscape.

Since then, large scale experimental studies conducted with whisky (Velasco et al., 2013), vodka (Wang & Spence, 2015b), and wine (North, 2012; Spence et al., 2013, 2014) have all demonstrated that what people hear in the background,

namely specifically chosen/composed music (or soundscapes), can significantly alter their perception (or at the very least their ratings) of a variety of taste/flavour characteristics, such as the woodiness of whisky and the freshness/fruitiness of wine.

Reinoso Carvalho et al. (2015a) performed a chocolate tasting test with three soundtracks (sweet, medium, bitter) and three types of chocolate (bitter, medium, sweet), but first asked participants to match the soundtracks with the chocolates. The results of the test confirmed earlier findings that it is possible to significantly influence gustatory ratings using sonic cues corresponding to different tastes. Moreover, by letting the participants match their own stimuli, greater differences between the conditions presented was observed (e.g., the bitter chocolate tasted sweeter when participants listened to the soundtrack they individually matched with the sweet chocolate, rather than the soundtrack that the experimenter designed to match with the sweet chocolate). From this, it can be argued that using customised soundtracks (or at least pre-tested by the participant) may lead to greater auditory taste modulation effects.

Kontukoski et al. (2015) took a behavioural approach to measuring the influence of sound on eating/drinking. Instead of studying the effect of music on taste ratings, they instructed participants to mix a drink from a variety of juices (orange, lemon, grapefruit, mango, pineapple, and liquid honey) while listening to one of four possible soundtracks. Two of the soundtracks were selected in a pre-test to match with sweetness, and two of the soundtracks with sourness. A chemical analysis of the juice mixtures in terms of sugar and acid concentrations then revealed that drinks prepared while listening to the sweet soundtracks contained

more sugar, and drinks prepared while listening to the sour soundtracks contained more acid.

On a related note, there have been a few studies focused on the impact of white noise. Woods et al. (2011), for instance, found that participants who consumed different foods reported lower sweetness and saltiness, but higher crunchiness ratings, when exposed to loud white noise at 75-85 dB as compared to quieter noise at 44-45 dB. In a more detailed study, Yan and Dando (2014) exposed participants to airplane cabin noise at 80-85 dB for half an hour before tasting basic taste solutions (sweet, bitter, salty, sour, and umami) in a variety of concentrations. Compared with their performance in a silent condition, participants experienced decreased sweetness sensitivity but increased umami perception.

1.2.2.2 The use of sonic seasoning in real-world applications

For a more general examination of the role of music in the history of dining in Western tradition, see Wang (2013). The following review is restricted to modern applications, where music is explicitly used to enhance or alter the evaluation of food/beverages.

The most famous example of sonic seasoning is plausibly “The Sound of the Sea”, which has been the signature dish on the tasting menu at Heston Blumenthal’s The Fat Duck restaurant for a number of years now. With this dish, the waiter arrives at the table holding in one hand a plate of seafood that looks like the seashore – complete with “sand” and “sea foam” – and, in the other, a seashell out of which dangles a pair of iPod earphones. The diners are instructed to insert the earphones before starting to eat, whereupon they hear the sounds of the sea: the waves crashing gently on the beach together, the cry of seagulls, and the faint voices of

holiday makers in the background (<http://condimentjunkie.co.uk/case-study/sounds-sea>). The introduction of the soundtrack transforms the dining experience, putatively both by enhancing the flavour of the food itself and by getting the diner to pay more attention to the gustatory (and auditory) experience. Some diners have even been known to find the multisensory experience so powerful that they have been brought to tears while tasting this dish (e.g., de Lange, 2012; see Spence & Wang, 2015d, for a review of emotional responses from music-food/beverage combinations).

The House of Wolf, an experiential dining restaurant in London, once featured a “Sonic Cake Pop” dish where, along with the chocolate-covered bittersweet toffee, diners were provided with a number to call. Depending on which option they chose after making the call, the diner could either listen to a sweet-congruent or bitter-congruent soundtrack, designed to either bring out the sweet or bitter notes in the dessert (Spence & Piqueras-Fiszman, 2014).

Besides restaurants (albeit the high-end and experimental kind), in-flight dining offers a unique opportunity for sonically augmented meals. In 2014, British Airways launched a “Sonic Bites” soundtrack for its long-haul passengers, where soundtracks were designed to complement the starters, mains, and desserts (<http://www.ba-touchdown.com/2014/10/scrumptious-sound-bites-on-board/>). The dessert soundtracks featured high-pitch tones, putatively to bring out sweet flavours in the dessert. Similarly, rock music was served with red wine to make it appear robust and heavy (e.g., North, 2012), while classical music was paired with the white wine to enhance the perception of quality.

A second area where music and taste are increasingly being brought together is in customised soundtracks – often embedded in sensory apps – to support and

complement special brand experiences. The champagne house Krug, for example, recently launched its Music Pairing series (<https://www.krug.com/>). With this app, the consumer can scan their wine label, or else type the individual number appearing on the back of their bottle, in order to access a selection of music that has been chosen to match the specific wine. Krug's "Music Pairing" uses musical selections from especially selected musicians to accompany six particular varieties of Champagne. At present, the musical matches that are offered are idiosyncratic, since they are based solely on the preferences of individual musicians selected by Krug.

In another example of sensory applications, Courvoisier developed "Le Nez de Courvoisier", an app to musically enhance the cognac drinking experience. Courvoisier commissioned a composer to compose a series of short musical clips to match the key notes (crème brûlée, ginger biscuits, candied orange, coffee and iris) found in their XO Imperial Cognac. The idea here was that consumers could listen individually to each of these musical pieces. Once they had established (or, better still, correctly guessed) the association between the pieces of music and the various aromas, presumably they could then highlight specific notes in their cognac by listening to the right piece of music. In addition, consumers could also listen to another musical composition in which each of these individual musical elements had been skilfully integrated into a single, more complex, arrangement, similar to the complex interplay of aromas found in a fine cognac.

More recently, Cadbury launched the Cadbury Favourites project (<http://cadburyflavourites.co.uk>), a music album available on Spotify with nine unique soundtracks designed to enhance each of its Dairy Milk chocolate range. The music was composed by the London Contemporary Orchestra, based on data

collected from consumer research regarding which sounds best complement each type of chocolate. For instance, their “Crunchie Bits” chocolate is paired with high-pitched bright tone and rhythmical variation, where as the Caramel chocolate is paired with a smooth sound at a moderate pitch. According to Cadbury, listening to the corresponding soundtrack while eating a chocolate bar can enhance the taste experience and provide more pleasure (although of course, no empirical evidence is given on that account).

However, the success of such sonic branding efforts is by no means guaranteed. For instance, when Crisinel et al. (2013) tested a subset of the Courvoisier soundtracks and odours (candied orange, crème brûlée, and ginger cookies) in an experimental setting, participants only reliably matched one of the odours (candied orange) to its intended soundtrack. This demonstrates the need for further research and illuminates the tension between marketing and research, where attractive pronouncements are prioritised over demonstrable and reproducible results.

1.3 Multisensory influences on flavour perception

Eating is one of the most multisensory experiences we experience on a daily basis. When people think of eating, the intrinsic senses of taste and smell usually come to mind. However, recent research has demonstrated that all our senses play a role in influencing flavour perception (see Auvray & Spence, 2008; Delwiche, 2004; Stevenson, 2014; for reviews). For instance, recalling the experience of eating an apple would bring up not just the taste and smell, but also the colour of the apple, its weight and shape, its crunchiness and juiciness, and even the sound of chewing.

1.3.1 A unified flavour perception

The psychologist E. B. Titchener wrote over a century ago about the “curiously unitary character” of the senses when eating a peach, where the senses can be separated individually, but still blend together to give a unitary sense of peach flavour:

“Think, for instance, of the flavour of a ripe peach. The ethereal odor may be ruled out by holding the nose. The taste components — sweet, bitter, sour — may be identified by special direction of the attention upon them. The touch components — the softness and stringiness of the pulp, the pucker feel of the sour — may be singled out in the same way. Nevertheless, all these factors blend together so intimately that it is hard to give up one’s belief in a peculiar and unanalyzable peach flavour...” (Titchner, 1909, p. 135).

One intriguing question is just where the sensation of flavour comes from. A large body of research from the past century reveals that flavours appear to originate from the oral cavity, even if olfactory stimuli – especially retronasal olfaction – are detected in the nose (e.g., Hollingworth & Poffenberger, 1917; Rozin, 1982; Stevenson et al., 2011). In addition, the phenomenon of oral referral appears to go beyond merely changing the perceived location of olfactory stimuli; in fact, they are combined with taste information from the tongue to form integrated flavour percepts (Rozin, 1982; Spence et al., 2015). Notably, people find it difficult to attend selectively to olfactory stimuli after the stimuli have been localised in the mouth (Ashkenazi & Marks, 2004). The loss of the source of olfactory information is most likely a result of gustatory capture (argued by Spence, 2016a), where the most intense stimulus (normally taste) directs one’s attention to the spatial location where that stimulus comes from. This is supported by studies indicating that the

degree of oral referral is proportional to the intensity of the tastants, and inversely proportional to the intensity of olfactory stimuli (Stevenson et al., 2011).

Intriguingly, the occurrence of oral referral also seems to be related to the degree of congruency between the oral and taste stimuli. Lim and Johnson (2011) demonstrated that, when participants were introduced to a simultaneous retronasal odour (soy sauce, vanilla) and a taste solution (sweet, salty, water), they rated the odours as coming from the mouth significantly more often when the odour-taste combination was congruent (vanilla – sweet, soy sauce – salty) than when the solution was neutral or when the combination was incongruent. Further studies conducted with solid gelatine disks instead of liquid solutions (Lim & Johnson, 2012), and with more ecologically valid stimulus combinations (citral aroma with sweet or sour tastants, coffee aroma with sweet or bitter tastants; Lim et al., 2014) revealed similar results where self-reported smell-taste congruency enhanced oral referral. In addition, the latest research supports the hypothesis that retronasal enhancement of odour by taste is dictated by the nutritive value of the tastants in addition to odour-taste congruency; sweet, salt, and umami tastes – which signal the presence of elements essential for survival – presented evidence of enhancing retronasal odour, but no such effect was seen for sour or bitter tastes (Linscott & Lim, 2016).

1.3.2 Neural substrates underlying flavour

In humans, taste is first processed in the primary taste cortex in an area of the anterior insula (see Small et al., 1999, and Small, 2010, for reviews), along with oral texture and temperature (de Araujo & Rolls, 2004; Guest et al., 2007). Smell, on the other hand, is first processed in the pyriform cortex (Poellinger et al., 2001;

Sobel et al., 2000; see Gottfried, 2010, for a review). In support of the idea that the different information channels which contributes to flavour are bound together in an unified entity, neuroimaging evidence has demonstrated that convergence for taste and odour occur in the orbitofrontal and anterior cingulate cortex (de Araujo et al., 2003; Small et al., 2004; Small & Prescott, 2005; see Rolls, 2015, for a review). Indeed, increased blood oxygen level was observed in the orbitofrontal cortex and amygdala when taste and odour stimuli are presented in combination, compared to the summed activity of taste and smell pretended alone (Small & Jones-Gotman, 2001). Furthermore, there is evidence that visual inputs associated with food also converges with olfactory and taste information in the orbitofrontal cortex (O'Doherty et al., 2002; Simmons et al., 2005; Wang et al., 2004; see Spence, 2016c, 2017, for reviews).

1.3.3 Flavour as a multisensory perceptual system

Of course, flavour involves more than just the senses of smell and taste. The trigeminal system plays an important role in flavour formation, whereby chemical irritants are detected as feelings of stinging or burning that are usually associated with warming or cooling sensations (Green & Lawless, 1991). Trigeminal irritants such as capsaicin have been demonstrated to inhibit taste and odour perception (Cain & Murphy, 1980; Prescott et al., 1993; Prescott & Stevenson, 1995). Touch, in terms of food texture and mouthfeel, has also been reported to interact with flavour. For instance, increasing the viscosity of a liquid can lead to decreased taste and flavour intensity (Arabie & Moskowitz, 1971; Christensen, 1980; Cook et al., 2003), and rough-textured foods are rated as more sour than smooth-textured foods (Slocombe et al., 2016). In addition, visual attributes such as colour and

shape can influence odour identification, taste intensity, and flavour description (e.g., Clydesdale et al., 1992; Fairhurst et al., 2015; Johnson & Clydesdale, 1982; Morrot et al., 2001; Zampini et al., 2008; see Spence et al., 2010, and Spence, 2015, for reviews). Finally, the previous section (1.2.2) reviewed the evidence found in literature for the role of audition in flavour perception.

Gibson (1966) proposed an ecological model whereby information about an object is processed and interpreted via different sensory channels, as a part of an active process to acquire information about the environment (Auvray & Spence, 2008). Flavour perception, then, may be considered as a system that controls ingestion, with the goal of picking up all available information about the food that is about to enter the body in order to secure an adequate supply of nutrients and avoid poisons (Stevenson, 2009). Moreover, this information intake can be considered in multiple stages: first, there is the pre-ingestion period when food is identified and expectations are formed – this is probably most naturally gathered via visual information, together with some degree of tactile (e.g. weight, surface texture, hardness), olfactory (orthonasal), and auditory information (e.g., sizzling, fizzing, bubbling). Then, there is the actual eating/mastication period where additional properties of the food – such as taste, (retronasal) aroma, texture, temperature, and piquancy – are detected by various taste and somatosensory receptors. These receptors serve to detect nutrients and poisons in the food (Chalé-Rush et al., 2007; Green & Lawless, 1991; Schiffman, 2000). At the same time, hedonic judgments are made continuously during ingestion as a way of motivating and curtailing ingestion (e.g., Hetherington, 1996). Finally, learned associations can form between different sensory stimuli as a result of the eating process (e.g., red coloured fruits are ripe and sweet; Stevenson, 2014).

Just as the tactile system combines disparate information from various parts of the body to register invariant stimuli, this proposed flavour system combines information from all the senses in order to form flavour percepts that ultimately maximises nutrient intake. Viewed from this perspective, extrinsic information such as sound can act to provide extra information about the food that one is about to taste or is currently tasting. In fact, one could argue that, when background music is heard while eating, the brain might – consciously or otherwise – incorporate the auditory information as part of the flavour evaluation process. Certainly, external information such as the weight and colour of the tableware can influence people’s perception of food and drink (Harrar et al., 2011; Piqueras-Fiszman et al., 2012; Tu et al., 2016; see Spence et al., 2012, for a review). According to the Bayesian model of sensory integration, the brain uses prior knowledge about what sensory signals go together – whether inborn or explicitly learned – to integrate appropriate sensory stimuli with the goal of maximising the reliability of perceived information (Ernst, 2008; Parise & Spence, 2009; Spence, 2011). Sound-taste/flavour correspondences, then, could act as a conduit (i.e., in the form of a Bayesian prior) to help the brain interpret auditory cues in order to optimise taste/flavour evaluations.

1.4 Possible mechanisms underlying auditory modulation of taste/flavour

With the object-based model of multisensory flavour perception in mind, the question remains as to the possible mechanisms underlying the sonic seasoning effects observed in Section 1.2. As mentioned in Section 1.3, our sense of hearing is involved in shaping our expectations about the food we are about to consume

(see Spence & Wang, 2015a, for a review on the impact of container opening sounds on people's evaluations of the identity and sensory properties of the beverage inside), as well as gathering information about the food itself (e.g., Velasco et al., 2013; Zampini & Spence, 2004).

Here, I will give a brief overview of some of the possible ways in which crossmodal correspondences between taste and sound might be used, implicitly or explicitly, to influence people's perception and evaluation of their food/drink. Each of the possible mechanisms is discussed in greater detail in Chapters 4-8.

1.4.1 Sensory expectations

Referring back to the canonical Crisinel et al. (2012) study, it is possible that the taste-congruent soundtracks, which could be heard a few seconds before the participants actually started to taste the toffee sample, may have induced sensory expectations about the flavor of the toffee (conscious or otherwise). These sensory expectations, could, in turn, influence the perceived taste of the toffee. This explanation is supported by previous research underlining the powerful effects that expectations can have on sensory evaluations (e.g., see Deliza & MacFie, 1996; and Piqueras-Fiszman & Spence, 2015, for reviews). The idea here is that when a stimulus is presented in one sensory modality, it may prime those sensory attributes that correspond crossmodally with it. Therefore, listening to the high-pitched sweet soundtrack may have primed sweetness in the minds of the participants in Crisinel et al.'s study, just as listening to the low-pitched bitter soundtrack may have led to expectations of a more bitter food product.

1.4.2 Attention

According to the attentional account of sonic seasoning, taste-congruent soundtracks work by drawing the listener's attention towards the taste/flavour that happens to correspond to the soundtrack, rather than to another taste/favour. It has been long established that attention can enhance the salience of the attended stimulus/ feature, relative to when the same stimulus/ feature is unattended (see Driver, 2001; Spence, 2014b; though see also Laing & Glenmarec, 1992; Stevenson, 2012). What is also known is that the effects of attention on awareness tend to become more pronounced as the perceptual input becomes more challenging/complex (Lavie, 1995, 2005). This may help to explain the hypothesis that the sonic seasoning effect tends to be more pronounced for foods with complex flavours (Spence et al., 2011).

In addition, such crossmodal effects could occur in either a stimulus-driven (exogenous) or voluntary (endogenous) manner. In all the experiments reviewed thus far (see Section 1.2.2.1), the fact that music/sound was relevant to the tasting experience was made apparent to the participants. It is an open question, therefore, as to whether similar crossmodal effects would be observed if the participants are not paying attention to the music, or whether these kinds of effects require the participant to actively link their tasting experience with what they are hearing (Spence, 2015c).

1.4.3 Neurophysiological response

Of course, there is the possibility that any crossmodal effects of music on taste perception may have, at least in part, a direct, low-level, neurophysiological basis (i.e., one that does not require the involvement of attention). Such a suggestion is

inspired by Wesson and Wilson's (2010) surprising discovery of direct connections between the ear and the olfactory tubercle in mice (see Wesson & Wilson, 2011, for a review of the multiple functions of the olfactory tubercle). It is certainly feasible that such a crossmodal connection might exist in humans.

A different possible neural connection was suggested by Brown et al. (2011), who demonstrated that the processing of aesthetic stimuli – be they paintings, music, or food – overlaps within the primary gustatory cortex. The authors propose that the aesthetics system evolved first for the appraisal of objects necessary for survival, such as evaluating the suitability of food/energy sources; over time, the same neural circuitry was co-opted for the appreciation of artworks. Therefore, the fact that our evaluation of music and food are processed in overlapping brain areas might potentially account for the associations that we make between them, not to mention how the evaluation (especially hedonic) of stimuli in one sensory modality might influence the evaluation of another.

Furthermore, it is possible that certain sounds might trigger an automatic physiological reaction in the listener, similar to that induced by ingesting foods with a particular taste (Alaoui-Ismaili et al., 1997; Robin et al., 2003; Rousmans et al., 2000). This “embodiment effect” might then influence taste perception. For instance, listening to high-pitched, dissonant music might enhance the rate of salivation, which in turn makes the food seem sourer (because the participant might misattribute the increased salivation to the food instead of the music).

However, it is important to note here that synaesthesia is not suggested as a possible mechanism underlying the sonic seasoning effect (although some, like composer Nick Ryan, use their synaesthesia to create multisensory taste experiences). While there are cases of individuals who might experience a specific

taste/flavour upon hearing specific sounds, such individuals are extremely rare (Simner et al., 2010). Furthermore, the synaesthetes themselves differ in the nature and extent of their experiences (e.g., Beeli et al., 2005; Luria, 1968). Therefore, it would be of limited relevance to draw connections between the associations experienced between the small number of sound-taste synaesthetes and the general sound-taste correspondences experienced by most people.

1.4.4 Emotion Mediation / Sensation Transfer

By comparison, according to the sensation transfer account, a soundtrack can have a halo/horns effect on participants' food evaluations (Kappes et al., 2006). Music can have emotional priming effects, for instance, on the processing of facial expressions (Logeswaran & Bhattacharya, 2009). In terms of participants from Crisinel et al.'s (2012) study, they might have rated the toffee samples as tasting sweeter while listening to sweet music compared to the bitter music, because they liked the sweet soundtrack more than the bitter soundtrack. The participants might have experienced positive feelings from listening to the sweet soundtrack. These positive feelings could have resulted in a more positive affective response to the toffee samples, which would have boosted the rating of pleasant (sweet) tastes.

1.4.5 Response bias

Finally, it is possible that taste-congruent soundtracks may act only to alter participants' self-reported ratings without having a genuine perceptual effect. For instance, participants in Crisinel et al.'s (2012) study could have rated the toffee samples differently because listening to the sweet soundtrack made them feel like they ought to rate the taste of the toffee higher on the bitter-sweet scale, compared to when listening to the bitter soundtrack. Nevertheless, this explanation does not

necessarily nullify the effect of crossmodal correspondences, since the taste-corresponding soundtracks are still modifying taste/flavour evaluation.

1.5 Aims and outline of the following chapters

The goal of the experiments reported in this thesis is to explore the different types of sound-taste correspondences, and to understand how such correspondences might go beyond a theoretical interest to influence real-life tasting experiences. In particular, the experiments reported in Chapters 4-8 address each of the potential mechanisms in turn. Chapter 9 then takes a lateral approach by assessing the role of individual differences, such as tasting expertise and taster status, in moderating sonic seasoning effects. Finally, Chapter 10 presents a comprehensive framework in which all the potential mechanisms can be placed. Through this thesis, I hope to contribute to the understanding of multisensory tasting perception and human information processing in general, by investigating the way in which we perceive and integrate taste/flavour information from multiple sensory streams. In addition, this type of research can provide new insights for designers – both artistic and commercial – interested in the growing field of synaesthetic design, by helping them incorporate multisensory perception research in the creation of new experiences and products.

2 How are crossmodal correspondences between sound and taste formed? The emotion mediation hypothesis

2.1 Introduction

Researchers have recently started to uncover the extensive range of crossmodal correspondences that exist between sound (and music) on the one hand, and tastes, aromas, and flavours on the other (Belkin et al., 1997; Crisinel & Spence, 2011a, b; see Knöferle & Spence, 2012, for a review). The research shows that soundtracks can be rated as sweet, salty, sour, or bitter depending on certain parameters of their composition, such as their pitch, articulation, loudness, etc. (e.g., Crisinel & Spence, 2010a; Knoeferle et al., 2015; Mesz et al., 2011). For example, sweetness and sourness tend to correspond with sounds that are higher in pitch while bitterness corresponds with lower-pitched sounds instead (Crisinel & Spence, 2009, 2010b). However, while the reliability of such crossmodal matches has become increasingly clear, the most appropriate explanation for such surprising crossmodal correspondences between seemingly unrelated stimuli in different sensory modalities has yet to be reliably determined.

2.1.1 Emotional mediation of crossmodal correspondences

One potential explanation for the crossmodal matching of sound with taste is in terms of emotional mediation. The suggestion here is that certain crossmodal correspondences may reflect the common emotional dimensions (such as pleasantness or arousal, see Collier, 1996, for a reduction of emotion space to two dimensions)

shared by the various stimuli involved. So, for instance, the correspondence between consonant musical harmony and sweetness (see Wang & Spence, 2016) may be attributable to people finding both stimuli pleasant. Such a hedonic matching account between seemingly unrelated stimuli presented in different sensory modalities explains, at least in part, colour-music matching (Palmer et al., 2013), colour-odour matching (Schifferstein & Tanudjaja, 2004), and even shape-taste matching (Velasco et al., 2016).

Relevant evidence pertaining to the case of crossmodal correspondences between audition and taste has, however, been limited to the pleasantness account. In an experiment designed to evaluate whether pleasantness mediates the crossmodal mapping between chocolate and sounds varying in their pitch and timbre (i.e., instrument type), Crisinel and Spence (2012) reported that while the type of instrument sound chosen by their participants could be predicted on the basis of the pleasantness ratings they gave to the dark chocolate that they sampled, their choice of pitch could not. It should, however, be borne in mind that pleasantness constitutes but a single dimension of emotional space (Collier 1996). Importantly, to date, no one has yet examined the potential mediating role of emotional arousal in crossmodal correspondences between audition and taste.

2.1.2 Measuring emotions

So far, a model of valence and arousal has been used to define emotion, as opposed to a more categorical definition of basic emotions (Ekman, 1992). Experiments 1A and 1B both focused on two dimensions of valence and arousal as opposed to, say, a set of six basic emotions (happiness, surprise, sadness, fear, anger, and disgust). This was because a valence/arousal model is easier to measure, and because Wang et al. (2015)

recently observed that participants uniquely associate each taste (albeit imagined) with different valence/arousal values (see also Cavanaugh et al., 2015, for evidence that different perceptual attributes, including taste, can be used to differentiate emotions with different valence and arousal measurements). Using the valence/arousal definition of emotion certainly makes it easier to test the pleasantness mediation theory that has been documented for other crossmodal correspondences, since pleasantness/valence is already one dimension that has already been measured.

2.1.3 Taste and emotion

In order to evaluate whether auditory-gustatory crossmodal associations are mediated by emotion, the emotional associations that people have with basic tastes must be established first. Although there has been much research on the influence of emotions on eating (see Machet, 2008; Singh, 2014, for reviews) there is limited evidence on the emotions that may be evoked by (or associated with) basic taste stimuli. Robin et al. (2003) established emotional responses to the basic tastes based on autonomic nervous system parameters, which they then transcribed into one of six basic emotions (happiness, surprise, sadness, fear, anger, and disgust). According to their research, sweetness corresponded with ‘happiness’ and ‘surprise’, bitterness was matched with ‘anger’ and ‘disgust’, and sourness and saltiness elicited a variety of different emotions (Robin et al., 2003). Meanwhile, Bredie et al. (2014) measured emotional responses to the basic tastes at multiple stimulus concentrations. However, their study was limited to facial expressions and self-reported pleasantness/valence ratings. Yamaguchi et al. (1984) conducted a more thorough investigation involving all five basic tastes (i.e., including umami) at various concentrations, but again limited to pleasantness ratings. In Experiment 1, therefore, emotional responses were

collected, both in terms of valence and arousal, for basic tastes solutions at multiple concentrations.

2.1.4 Pitch and intensity

For whatever reason, pitch has been one of the most frequently studied attributes in crossmodal correspondences research involving the auditory modality. Researchers have, for instance, highlighted the existence of crossmodal associations with elevation (Ben-Artzi & Marks, 1995; Evans & Treisman, 2010; Mudd, 1963; Parise et al., 2014), but also with brightness (Marks, 1987), lightness (Marks, 1987), size (Evans & Treisman, 2010; Gallace & Spence, 2006), aroma (Crisinel & Spence, 2012), and taste (Crisinel & Spence, 2009, 2010a). In terms of taste-pitch matching, one drawback is that the sound-taste studies that have been conducted to date have typically utilized only a single taste intensity. In real life, of course, foods have tastes/flavours that vary widely in terms of their intensities and associated pleasantness. Experiment 1 therefore assessed whether variation in perceived taste intensity would influence such taste-pitch correspondences.

Previously, it has been shown that the method of magnitude matching between loudness and taste intensity, where participants choose a sound volume to match the perceived taste intensity, can highlight individual differences in taste perception and be used to establish a common basis for comparison across different participants (Gescheider, 1997, pp. 285-287). In a study by Marks et al. (1988), taste intensity-volume matching was used to demonstrate that PROP (6-*n*-propylthiouracil) compounds were experienced differently by supertasters (who perceive a bitter taste) than nontasters (who don't perceive any taste). Elsewhere, Bartoshuk and colleagues used magnitude estimation to study the loss of taste perception in the elderly. They

found that the elderly matched dilute tastes to louder sounds than young adults, possibly due to the elderly having a chronic background taste in the mouth or having weakened hearing (Bartoshuk et al., 1986).

In Experiments 1A and 1B, participants are presented with basic taste solutions at three different concentrations, and had to choose a pitch that best matched each tastant. In addition to pitch, the participants are also asked to select loudness as a way of measuring subjective taste intensity in Experiment 1A, and directly asked for their perceived taste intensity in Experiment 1B.

2.1.5 Beyond pitch: Comparing taste soundtracks

As part of this growing movement to study sound-taste correspondences, a number of researchers and sonic designers have recently started to compose their own soundtracks to match (or, in some sense, to be associated with) specific tastes. Increasingly, such soundscapes/soundtracks are being used in both scientific research and a growing number of artistic performances (e.g., Crisinel et al., 2012; Knoeferle et al., 2015; Kontukoski et al., 2015; Mesz et al., 2012; Carvalho et al., 2015a, b; Wang, 2013). Such soundtracks could have either or both of the following aims: 1) To demonstrate a connection between taste and sound in terms of participants being able to associate certain soundtracks with specific tastes; and 2) To modify the perceived taste of food and drink while participants listen to such soundtracks.

Up until now, however, the soundtracks produced by various researchers and designers have never been tested in the same setting. Furthermore, some of the soundtracks, such as those used in Knoeferle et al. (2015) and Mesz et al. (2012), have only previously been tested in a constrained context in which the participants

were essentially given a four-alternative-forced-choice task (i.e., with four soundtracks to listen to and four taste words to choose from). Under such conditions, however, one could potentially sometimes answer correctly simply by virtue of the test format (e.g., imagine a case in which a participant experiences a strong association between three soundtracks and three of the tastes, then, by default, the fourth soundtrack is guaranteed to be correctly matched to the only remaining choice. This would be true even if the soundtrack did not clearly match a particular taste. Therefore, in Experiment 2A, a group of these recently created taste-inspired soundtracks were tested together in the same sound and taste word matching task. The participants were free to choose any basic taste word they wished (out of the 4 possible choices of sweet, salty, sour, and bitter) for each of the soundtracks, in order to determine which of them exhibited the strongest association with a given taste. In Experiment 2B, the same soundtracks were used in order to measure the role of emotion as a possible mediator in the soundtrack-taste matching task. The participants were not asked to match the soundtracks to taste words, but instead were asked for their emotion ratings (for pleasantness and arousal) of the soundtracks and of bitter/salty/sour/sweet-tasting foods. These emotional ratings were then combined with the soundtrack-taste matching results from Experiment 2A in order to calculate the potential role of pleasantness and arousal in mediating the crossmodal correspondence between audition and taste.

2.2 Experiment 1A: Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and auditory pitch

2.2.1 Methods

2.2.1.1 *Participants*

33 participants (19 women, 14 men) aged between 19-35 years ($M=24.03$, $SD=4.50$) took part in the study. The participants gave their written informed consent, and reported no cold or other impairment of their senses of smell, taste, or hearing. The participants were recruited according to the Experimental Psychology Research Participation Scheme and Oxford Psychology Research Participant Database, and each participant was awarded either £5 or 2 course credits upon completion of the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

2.2.1.2 *Auditory stimuli*

For the sound-matching task, there were 19 keys on the MIDI keyboard mapped from C2 to C8, with each consecutive key being 2 whole steps apart (so the keys were C2, E2, G#2, C3, E3, etc... up to C8). All of the notes used the Steinway piano synthesizer from Apple's GarageBand software. The volume level was controlled by a dial with 7 radial markings around it. The volume at level 4, the middle marking, was approximately 75dB. Each increase in level corresponded with approximately a 5 dB increase, with a maximum of approximately 90 dB at level 7. The participants wore HD-3030 stereo headphones during the sound-matching task.

2.2.1.3 Taste stimuli

10 mL samples of bitter, sweet, sour, salty, and umami solutions were prepared, in 3 different concentrations (weak, medium, and strong). The intensity of all five tastes for each concentration level (4=weak, 7=medium, and 10=strong) were matched according to taste intensity scales developed at the University of Minnesota in Saint Paul (Karalus et al., 2010; see Table 2.1 for ingredients and concentrations of each of the taste solutions). The taste intensity scales were created by means of a two-step process. First, a sourness scale was constructed by having the participants (N=32) rate the intensity of 13 samples with different citric acid concentrations, dissolved in drinking water (Premium Waters, Minneapolis). The best-fit regression line between concentration and intensity was used to determine citric acid concentrations for intensity values between 0-20. The sour scale was then used as a reference scale to create the bitter, sour, sweet, and umami scales (N=20).

The solutions for this experiment were presented in clear 50mL clear plastic cups. The solutions themselves were both colourless and odourless.

Taste	Ingredient	Weak concentration (g/L)	Medium concentration (g/L)	Strong concentration (g/L)
Bitter	Caffeine	0.56	1.22	2.21
Salty	Salt (NaCl)	2.10	5.81	9.61
Sour	Citric acid	0.63	1.37	2.40
Sweet	Sugar	33.47	86.14	138.80
Umami	Monosodium	0.94	18.73	44.95

	glutamate (MSG)			
--	--------------------	--	--	--

Table 2.1. Ingredients used to make each basic taste solution and the proportions used for each concentration level. The same solutions were used for both Experiment 1A and 1B. All of the solutions were mixed with distilled water.

2.2.1.4 Design and procedure

The experiment was conducted with participants sitting at a table in front of a computer monitor and a MIDI keyboard in an experimental booth. The experiment was programmed on the LimeSurvey online survey platform.

For the evaluation of the taste solutions, the participants were instructed to taste each sample by swirling the solution around their mouths for 3 seconds, then expectorating. As a practice trial, participants were given a medium intensity solution of a random taste.

During the actual test, the participants were presented with 15 samples, one for each trial. For each trial, the participants were instructed to taste a particular sample, then rate their emotional response and pick a note (out of 19 choices) and volume setting on the MIDI keyboard that best matched the taste. The emotional responses were in terms of arousal and valence, both on a scale from -5 to +5. In order to help the participants make emotion ratings (especially arousal), they were presented with a 2-D grid with valence on the x-axis and arousal on the y-axis. The four corners of the grid were anchored by excitement, relaxation, depression, and stress (see Figure 2.1). The order of all four questions for each trial, as well as the order of presentation of the taste stimuli, was randomised. The participants rinsed their mouths out with water

between every trial. After all 15 samples had been presented, the participants were asked a series of post-trial questions about how much they enjoyed eating {bitter/salty/sour/sweet/umami}-tasting foods to be answered on a 5-point scale.

The entire study lasted for approximately 30 minutes.

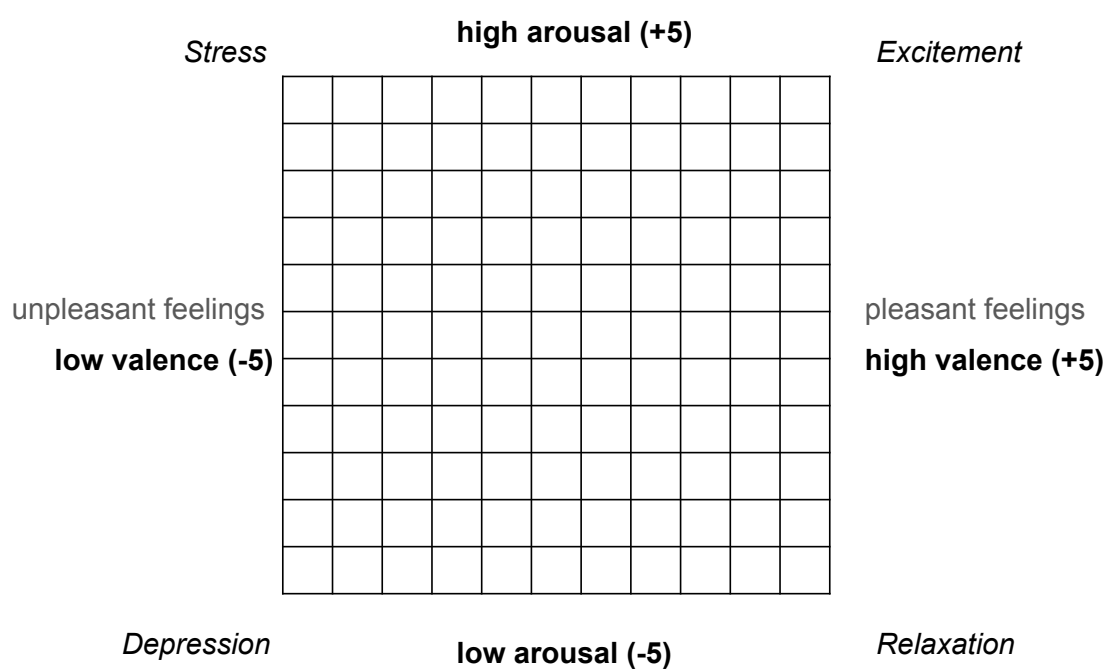


Figure 2.1. Grid used to help participants making valence/arousal ratings during Experiment 1A.

2.2.2 Results and discussion

2.2.2.1 Auditory parameter choices

The mean values of participants' choices for pitch and volume to match with various taste solutions at different concentrations are shown in Figure 2.2.

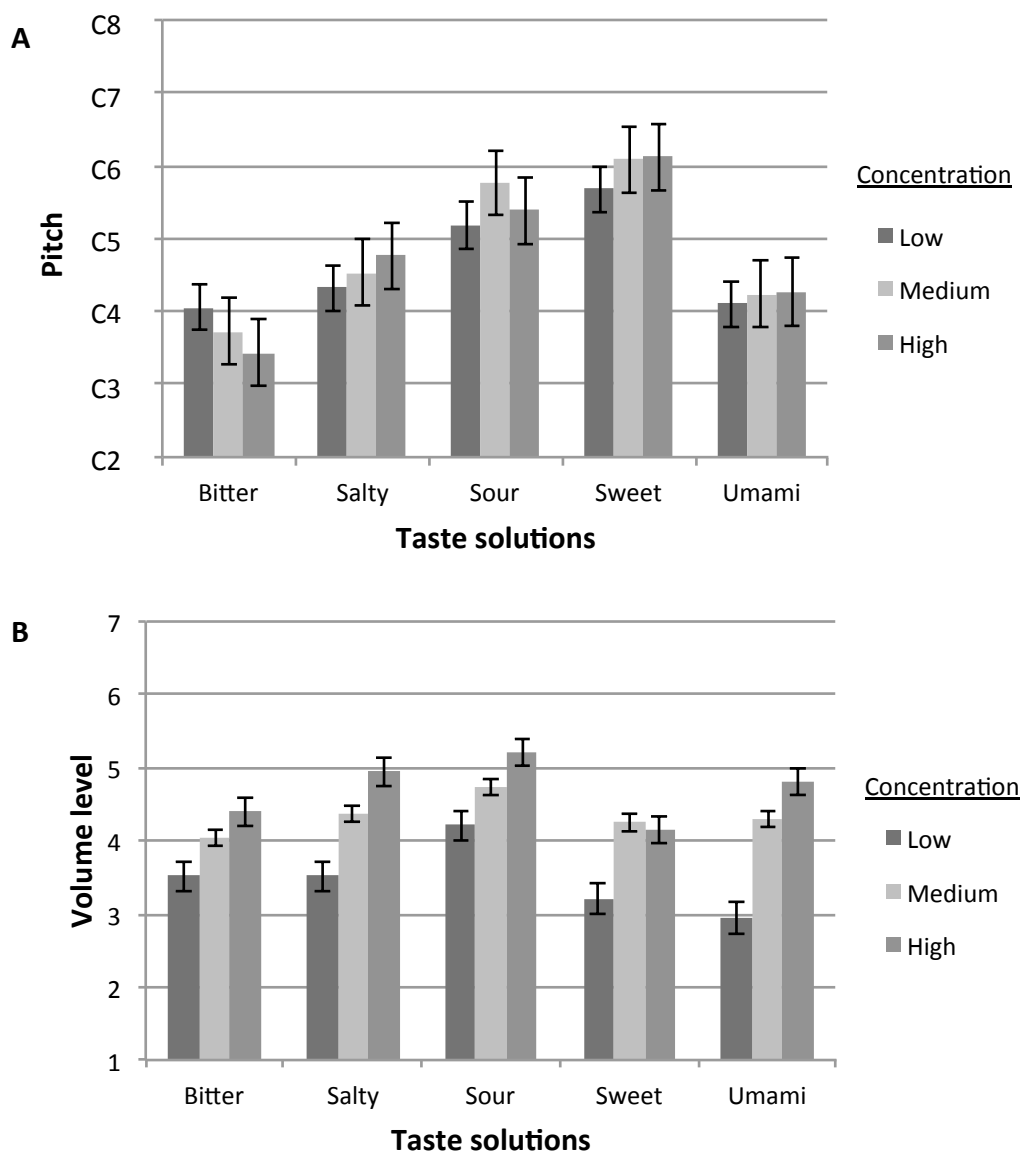


Figure 2.2. Mean values of participants' choice of pitch (A) and volume/loudness (B) that best matches with various taste solutions in different concentrations, in Experiment 1A. Pitch is shown in musical notation, with differences between each line (e.g., from C2 to C3) being one octave (A). Volume levels ranged from 1 (60dB) to 7 (90dB) (B). The error bars denote the standard error of the mean.

RM-ANOVAs with taste (bitter, salty, sour, sweet, umami) and concentration (low, medium, high) as factors were conducted on participants' pitch and volume choices. In terms of participants' choice of pitch (with scores ranging from 1, the lowest note,

to 19, the highest note), a main effect of taste solution type was observed, $F(4,128)=22.93$, $p<.0005$, $\eta^2=.42$. Pairwise comparisons revealed that sweet ($M=12.88$, $SD=3.27$) and sour ($M=11.33$, $SD=5.44$) solutions were matched to a significantly higher pitch than the bitter ($M=6.22$, $SD=4.36$), salty ($M=8.63$, $SD=4.23$), and umami ($M=7.60$, $SD=4.57$) solutions, regardless of their concentration ($p<.005$ for all comparisons). In addition, the salty solution was matched to a sound having a significantly higher pitch than the bitter solution ($p=.04$).

In terms of participants' choice of volume, a main effect of concentration levels was observed, $F(2,64)=67.57$, $p<.005$, $\eta^2=.68$, as predicted. Pairwise comparisons revealed that the low concentration solutions were matched to a significantly lower volume ($M=3.48$, $SD=1.64$) than the medium concentration ($M=4.33$, $SD=1.47$) solutions, which, in turn, were matched to a significantly lower volume than the high concentration solutions ($M=4.70$, $SD=1.47$; $p<.05$, for all comparisons). A significant main effect of taste solution type was also observed on volume ratings, $F(3.43,109.89)=5.25$, $p=.001$, $\eta^2=.14$, using Huynh-Feldt correction. Pairwise comparisons revealed that the sour solution ($M=4.72$, $SD=1.38$) was matched to a significantly higher volume than the salty ($M=4.27$, $SD=1.41$, $p=.010$), sweet ($M=3.87$, $SD=1.64$, $p=.001$), and umami ($M=4.02$, $SD=1.97$, $p=.017$) solutions.

Since volume can be interpreted as the participants' rating of the perceived intensity of a given taste solution (see Marks et al., 1988), there was reason to believe that it would more accurately reflect participants' individual assessments of the intensity of the solutions than the concentration levels been prepared. Pearson correlations were therefore calculated between pitch and volume ratings to assess any relationships between participants' perceived intensity and pitch choice. Overall, there was a significant positive correlation ($r_{495}=.12$, $p=.006$). More specifically, a significant

positive relationship between perceived intensity (volume) and pitch choice was documented for the salty ($r_{99}=0.30$, $p=.002$) and sweet ($r_{99}=0.31$, $p=.002$) solutions. On the other hand, a significant negative relationship was observed for the bitter solutions ($r_{99}=-0.20$, $p=.04$). No relationship was observed for the sour ($r_{99}=0.15$, $p=.15$) and umami ($r_{99}=0.17$, $p=.09$) solutions. In other words, for salty and sweet tastes, increased taste intensity was associated with higher pitch, while for bitter tastes, increased taste intensity was associated with a lower pitch instead.

2.2.2.2 Emotion ratings

In terms of participants' emotion ratings of various taste solutions, the mean ratings of valence and arousal are shown in Figure 2.3.

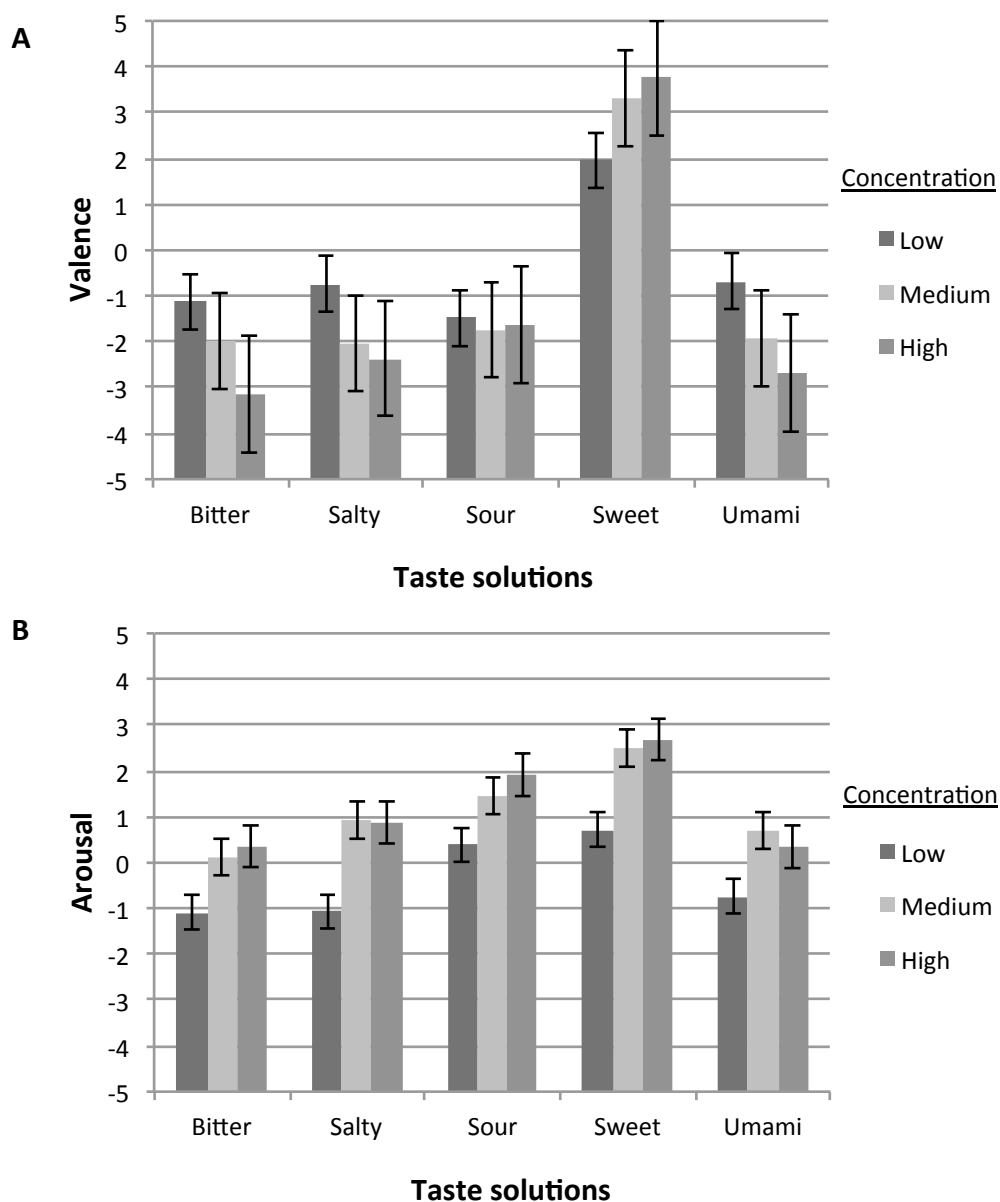


Figure 2.3. Mean values of participants' rating of emotional valence/pleasantness (A) and arousal (B) associated with various taste solutions in different concentrations, in Experiment 1A. Valence was measured on a scale from -5 (extremely unpleasant) to +5 (extremely pleasant). Arousal was measured on a scale from -5 (not arousing at all) to +5 (extremely arousing). The error bars denote the standard error.

To get a better understanding of participants' emotion ratings for the taste solutions having different concentrations, the average ratings of each taste were plotted for all

three concentrations on the same graph, with trendlines shown as vectors pointing in the direction of increased concentration (see Figure 2.4). Perhaps unsurprisingly, increasing concentration was associated with higher arousal levels. In general, bitter, salty, and umami solutions were rated as less pleasant and more arousing as the concentration increased. For the sour solutions, increases in arousal ratings were observed but little change in terms of pleasantness was seen. Only sweet solutions *increased* in terms of their pleasantness as the concentration increased. Presumably, these solutions did not reach the level of sweetness at which participants began to dislike them (see Giovanni & Pangborn, 1983).

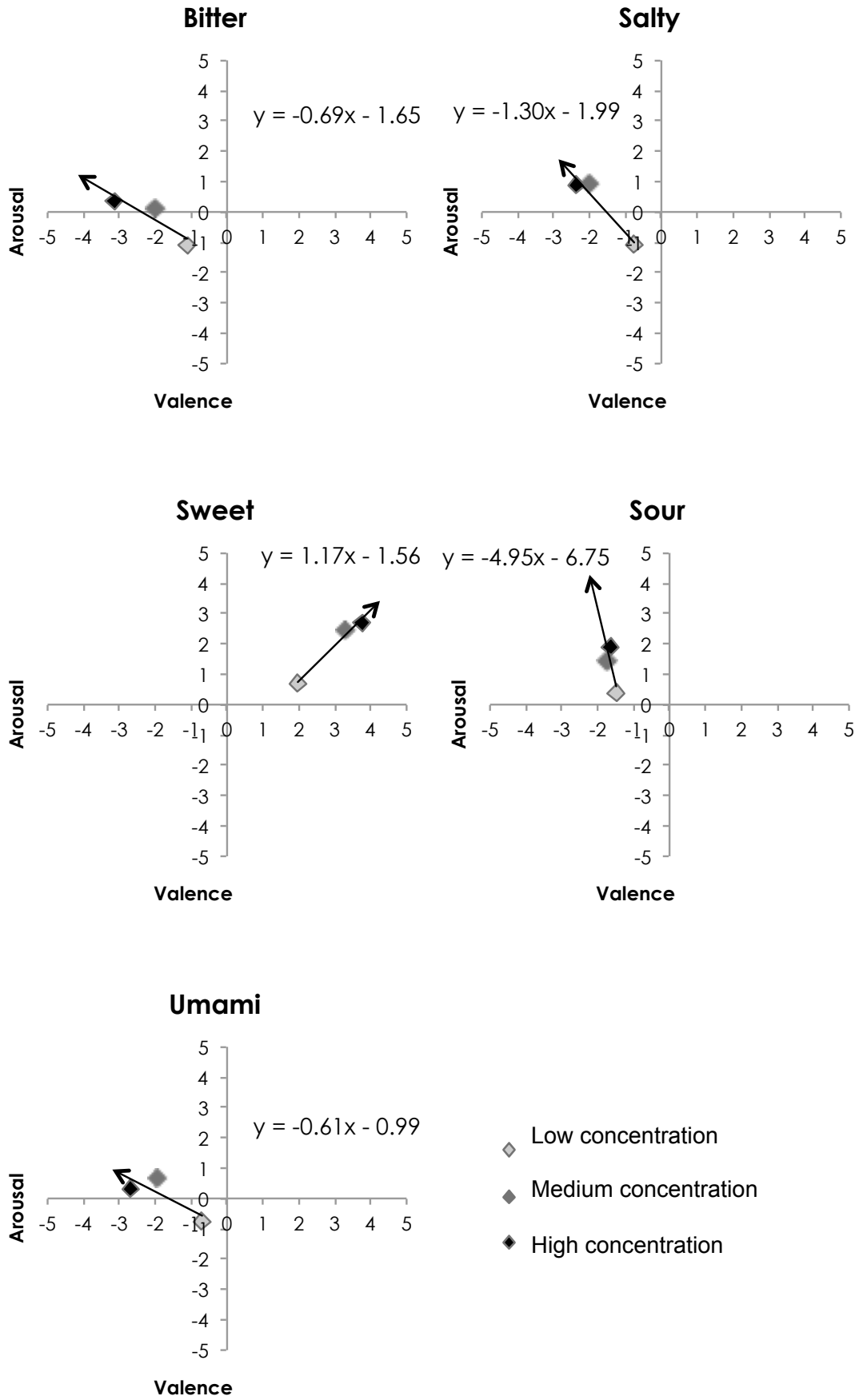


Figure 2.4. Mean values for participants' emotion ratings of each taste solution in Experiment

1A, plotted on a 2D valence-arousal graph. The horizontal valence axis ranges from +5 (extremely positive/pleasant) to -5 (extremely negative/unpleasant). The vertical arousal axis ranges from +5 (extremely arousing) to -5 (not arousing at all). For example, the top right corner (high valence and high arousal) would map to excitement, while the bottom left corner (low valence and low arousal) would map to depression. Different shading denotes different concentrations. A trendline is plotted for each taste (with its equation shown), with the direction of the arrow indicating the direction of increasing concentration.

Rm-ANOVA tests with taste and concentration as factors were conducted on participants' ratings of valence and arousal. In terms of participants' choice of valence, a main effect of taste solution type was observed, $F(2.78, 88.80)=65.16$, $p<.0005$, $\eta^2=.67$, using Greenhouse-Geisser correction. Pairwise comparisons revealed, unsurprisingly, that the sweet solution was matched to significantly higher valence ($M=3.01$, $SD=1.79$) than the bitter ($M=-2.09$, $SD=2.01$), salty ($M=-1.72$, $SD=1.97$), sour ($M=-1.62$, $SD=2.67$), or umami ($M=-1.78$, $SD=2.59$) solutions, regardless of the concentration at which the tastant was presented ($p<.0005$ for all comparisons). A significant main effect of concentration was also observed, $F(1.66, 53.01)=8.18$, $p=.002$, $\eta^2=.20$, using Huynh-Feldt correction. Pairwise comparisons revealed that the low concentration solutions ($M=-0.42$, $SD=2.26$) were rated significantly higher on the valence scale (i.e., as more pleasant) than the medium ($M=-0.88$, $SD=3.04$, $p=.013$) and the high concentration solutions ($M=-1.22$, $SD=3.39$, $p=.001$). An interaction effect was also observed between the taste of the solution and its concentration, $F(5.72, 183.08)=7.44$, $p<.0005$, $\eta^2=.19$. In particular, as concentration increased, valence ratings were higher for sweet solutions, but lower for bitter, salty, and umami solutions. Specifically, low concentration sweet solutions ($M=1.97$, $SD=1.96$) were rated as significantly less pleasant than medium ($M=3.30$,

SD=1.78, $p=.14$) and high concentration sweet solutions ($M=3.76$, $SD=1.00$, $p<.0005$). High concentration bitter solutions ($M=-3.15$, $SD=1.75$) were rated as significantly less pleasant than low ($M=-1.12$, $SD=2.00$, $p<.0005$) and medium concentration bitter solutions ($M=-2.00$, $SD=1.77$, $p=.027$). The high concentration umami solution ($M=-2.70$, $SD=2.74$) was rated as significantly less pleasant than the low concentration umami solution ($M=-0.70$, $SD=1.59$, $p=.003$). Finally, the low concentration salty solution ($M=-0.76$, $SD=1.50$) was rated as more pleasant than the medium ($M=-2.03$, $SD=1.70$, $p=.003$) and high concentration salty solutions ($M=-2.36$, $SD=2.29$, $p<.0005$).

In terms of participants' choice of arousal, a main effect of concentration levels was observed, $F(1.30, 41.51)=18.97$, $p<.005$, $\eta^2=.37$, using Greenhouse-Geisser correction. Pairwise comparisons revealed that the low concentration ($M=-0.36$, $SD=2.42$) solutions were rated as significantly less arousing than the medium ($M=1.13$, $SD=2.69$) and high concentration solutions ($M=1.24$, $SD=3.40$; $p<.0005$ for all comparisons). A main effect of taste solution type on arousal ratings was also observed, $F(2.84, 90.86)=7.49$, $p<.0005$, $\eta^2=.19$, using Greenhouse-Geisser correction. Pairwise comparisons revealed that the sweet solution ($M=1.96$, $SD=2.13$) was rated as significantly more arousing than the bitter ($M=-0.20$, $SD=3.08$, $p=.003$), salty ($M=0.24$, $SD=2.81$, $p=.049$) and umami ($M=0.09$, $SD=3.06$, $p=.005$) solutions; in addition, the sour solution ($M=1.25$, $SD=3.06$) was rated as significantly more arousing than the salty solution ($p=.036$) and almost as more arousing than the bitter solution ($p=.063$).

Pearson correlations between participants' choices of auditory attributes (pitch and volume) and their emotion ratings (valence and arousal) were then computed over all taste solutions. For pitch, positive moderate correlations between pitch and valence

($r_{495}=0.32$) and between pitch and arousal ($r_{495}=0.28$) were found. For volume, a moderate negative correlation with valence ($r_{495}=-0.25$) and a moderate positive correlation with arousal ($r_{495}=0.21$) was documented. All correlations were significant ($p<.0005$).

As many of the variables are correlated, partial correlation coefficients were calculated in order to control for the effect of possible third variables ($r_{495}=0.31$ between pitch and valence, $r_{495}=0.17$ between pitch and arousal, $r_{495}=-0.35$ between volume and valence, and $r_{495}=0.25$ between volume and arousal). All coefficients remained significant ($p<.0005$).

A multiple linear regression was used to test whether participants' rating of the valence and arousal of a given taste solution significantly predicted their choice of matching pitch or volume. The results of the regression indicated that for pitch, valence ($\beta=0.27$, $p<.0005$) and arousal ($\beta=0.22$, $p<.0005$) accounted for 14.8% of the variance, $R^2=0.148$, $F(2,492)=42.87$, $p<.0005$. For volume, valence ($\beta=-0.32$, $p<.0005$) and arousal ($\beta=0.29$, $p<.0005$) accounted for 14.0% of the variance, $R^2=0.140$, $F(2,492)=40.21$, $p<.0005$. As valence and arousal are correlated ($r_{495}=0.26$, $p<.0005$), they were also analysed independently. Valence alone accounted for 10.5% of the variance in pitch choice ($\beta=0.32$, $p<.0005$, $F(1,494)=32.67$, $p<.0005$) and 6.2% of the variance in volume choice ($\beta=-0.25$, $p<.0005$, $F(1,494)=57.76$, $p<.0005$). Arousal alone accounted for 8.1% of the variance in pitch choice ($\beta=0.28$, $p<.0005$, $F(1,494)=43.35$, $p<.0005$) and 4.3% of the variance in volume choice ($\beta=0.21$, $p<.0005$, $F(1,494)=22.11$, $p<.0005$).

Analysis of the post-experiment questions revealed that the average self-reported liking was 1.94 (SD=1.20) for bitter foods, 3.27 (SD=1.01) for salty foods, 2.82 (SD=1.13) for sour foods, 4.15 (SD=0.87) for sweet foods, and 4.12 (SD=0.86) for

foods associated with the taste of umami. A comparison of self-reported taste liking with the average valence ratings of each taste solution (averaged over all concentration levels) revealed that ratings were significantly different for salty ($t(98)=7.69$, $p<.0005$), sour ($t(98)=3.90$, $p<.0005$), and umami ($t(98)=14.76$, $p<.0005$) tastes.

Finally, some individual differences were observed in terms of taste preferences. 6 out of the 33 participants were identified as potential supertasters based on the criteria that they matched the weak concentration bitter solution with a volume level of at least 6 out of 7 (see Bartoshuk et al., 1994; Hall et al., 1975, for evidence that sensitivity to PTC predicts sensitivity to caffeine, which was used here for the bitter solutions). Interestingly, these “bitter-sensitive” individuals ($M=3.89$, $SD=4.52$) matched bitter solutions to significantly lower pitches than the remainder of the participants ($M=6.74$, $SD=4.18$), $t(97)=2.58$, $p=.011$. They also made higher valence ratings ($M=0.11$, $SD=2.74$) for sour solutions than the rest of the participants ($M=-2.00$, $SD=2.51$), $t(97)=-3.18$, $p=.002$.

The results of Experiment 1A therefore demonstrate the relationship between taste concentration, auditory pitch, volume, and both emotional valence and arousal. Differences in taste quality (bitter/salty/sour/sweet/umami) exerted a significant influence over the participants’ choice of both pitch and loudness as well as their ratings of arousal and valence. With increasing taste concentration, the participants chose louder sounds (as predicted), as well as higher arousal and lower valence ratings (except for sweet tastes, where increasing taste concentration increased valence ratings). For sweet and salty tastes, the choice of pitch was positively correlated with perceived taste intensity (as represented by loudness) whereas for bitter tastes, pitch choice was negatively correlated.

For those ratings where the concentration had a significant main effect (namely valence, arousal, and loudness), it is interesting to note that no significant differences were found between medium and high concentration for either valence or arousal ratings, although there was a significant difference between medium and high concentrations for loudness choices. Perhaps participants categorise taste concentrations as either high or low when it comes to expressing different degrees of valence and arousal, but when it comes to matching volumes, participants are better able to express differences. Alternatively, however, perhaps making an intensity match between taste concentration and loudness both physical properties, is easier than matching taste concentration with more abstract ideas such as valence and/or arousal (see Stevens & Marks, 1965, for an intensity matching account of the correspondence between brightness and loudness).

It is of interest for future research to note that people's self-reported liking for foods of a given taste may be different from their actual experience of the taste. Sour, salty, and umami were all liked significantly more as taste words (in the form of "sour/salty/umami-tasting foods") when actually consumed (albeit in liquid form). This means that for experimental studies involving hedonic measurements, it may be more appropriate to have one's participants consume real foods rather than using taste words (for instance, Velasco et al., 2014, showed a relationship between roundness/angularity ratings and liking for actual tastes, but not for taste words).

Overall, Experiment 1A showed a possible impact of both emotional valence and arousal on taste-note matching. However, because the participants chose pitch and volume and made emotion ratings in the same experimental block, it is possible that their choice of sound attributes were based on their emotion ratings. Therefore, in Experiment 1B, the experimental design was changed such that the participants now

made sound matches and emotion ratings in different blocks of trials. In addition, only indirect assessments of taste intensity have been made so far based on participants' volume selection. To double-check the assertion about loudness acting as a proxy for subjective intensity, participants rated the taste intensity of the solutions directly in Experiment 1B instead of selecting a loudness setting

2.3 Experiment 1B: Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and auditory pitch

2.3.1 Methods

In Experiment 1B, the possible concern that the participants used their emotion ratings to select notes by separating them into separate blocks was addressed. The participants were also instructed to make emotion ratings for the musical notes in order to make a more comprehensive emotional mediation analysis.

2.3.1.1 Participants

33 participants (23 women, 10 men) aged between 18-30 years ($M=21.67$, $SD=3.28$) took part in the study. They gave their written informed consent, and reported no cold or other impairment of their senses of smell, taste, or hearing. The participants were recruited in a similar manner as Experiment 1A, and each participant was awarded either £7 or 4 course credits upon completion of the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

2.3.1.2 *Auditory stimuli*

The musical note selection setup was identical to that used in Experiment 1A. In the music note emotion evaluation block, the participants listened to 2-second sound recordings of the same 19 notes available on the MIDI keyboard, recorded using the Steinway piano synthesizer from Apple's Garageband software.

2.3.1.3 *Taste stimuli*

The taste stimuli were identical to those used in Experiment 1A.

2.3.1.4 *Design and procedure*

Experiment 1B differs from Experiment 1A by separating sound-matching and emotion-rating tasks into different blocks of experimental trials. First, the participants tasted each sample in a random order, and had to choose a musical note (same selection process as Experiment 1A) that best matched the taste. On the next screen, the participants had to rate the perceived intensity of the sample that they had just tasted. The participants rinsed their mouths out with water between every trial. After the first block, participants were given a piece of cracker to eat to cleanse their palate while taking a 10 minute break. In the second block, the participants first practiced making valence/arousal ratings in response to images. Next, they tasted the samples again in a random order and made valence/arousal ratings for each sample, rinsing their mouths out with water between trials. Finally, in the third block, the participants listened to 2-second sound clips of all 19 music notes, in a random order, and made valence/arousal ratings for each note.

The entire study lasted for approximately 45 minutes.

2.3.2 Results and discussion

2.3.2.1 Auditory parameter choice

The mean values of participants' choices for pitch to match with various taste solutions at different concentrations are shown in Figure 2.5A.

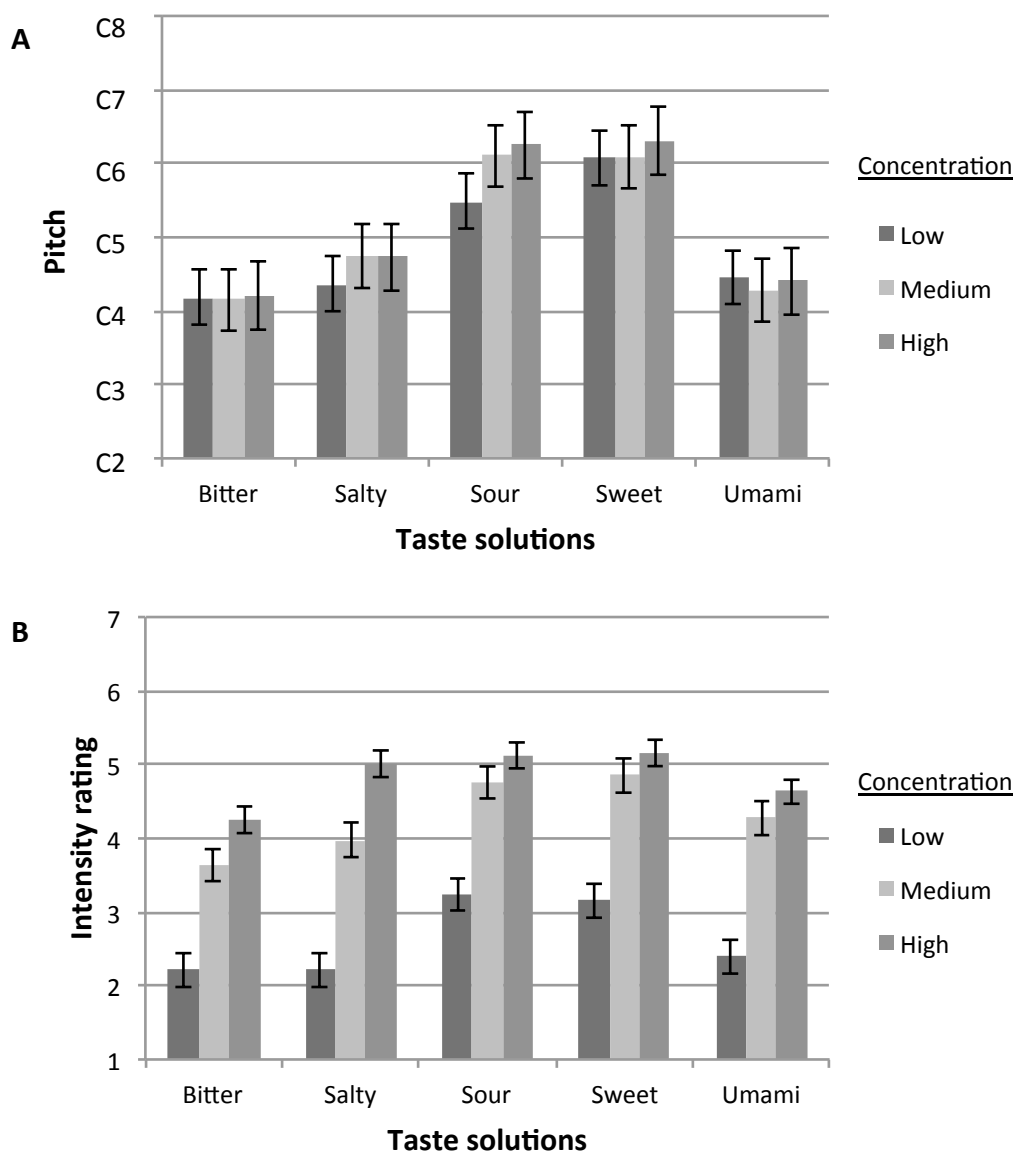


Figure 2.5. Mean values of participants' choice of best-matching pitch (A) and taste intensity (B) for various taste solutions in different concentrations, in Experiment 1B. Pitch is shown in musical notation, with differences between each line (e.g., from C2 to C3) being one octave (A). Intensity rating was on a scale from 1-7 (B). The error bars denote standard error.

An RM-ANOVA with taste (bitter, salty, sour, sweet, umami) and concentration (low, medium, high) as the factors was conducted on participants' choices of pitch. A main effect of taste solution type was observed, $F(4,128)=25.43$, $p<.0005$, $\eta^2=.44$. Pairwise comparisons revealed that sweet ($M=13.47$, $SE=0.55$) and sour ($M=12.84$, $SE=.55$) solutions were matched to a significantly higher pitch than the bitter ($M=7.54$, $SE=0.58$), salty ($M=8.83$, $SD=0.53$), and umami ($M=8.14$, $SD=0.45$) solutions, regardless of their concentration ($p<.005$ for all comparisons).

To validate that the participants had indeed perceived the solutions at different concentrations, their rating of intensity was assessed (see Figure 2.4B for mean rating values). As predicted, a main effect of concentration was observed, $F(2,64)=144.39$, $p<.0005$, $\eta^2=.82$. Pairwise comparisons revealed that the low concentration solutions were rated as significantly less intense ($M=3.48$, $SD=1.64$) than the medium concentration ($M=4.33$, $SD=1.47$) solutions, which, in turn, were rated as being less intense than the high concentration solutions ($M=4.70$, $SD=1.47$; $p<.05$ for all comparisons). A significant main effect of taste solution type on intensity ratings was also observed, $F(3.58,114.70)=6.25$, $p<.0005$, $\eta^2=.16$, using Huynh-Feldt correction. Pairwise comparisons revealed that the sweet solution ($M=4.38$, $SE=.20$) was overall significantly more intense than the bitter ($M=3.36$, $SE=.24$) and salty ($M=3.73$, $SE=.14$) solutions. In addition, the sour solution ($M=4.37$, $SE=.18$) was also significantly more intense than the bitter solution. This suggests that participants in fact did not perceive the solutions to be at equal intensity, as claimed by Karalus et al. (2010).

As in Experiment 1A, any relationship between participants' perceived taste intensity and pitch choice was assessed by calculating Pearson correlations between pitch and

intensity ratings of each taste. Unlike in Experiment 1A, participants' actual taste intensity ratings were used instead of a secondary measure (e.g., volume) that might reflect perceived taste intensity. Across all tastes, a significant weak positive correlation was shown between taste intensity and choice of pitch, $r_{495}=.12$, $p=.006$. More specifically, a significant positive relationship between perceived intensity and pitch choice was observed for the sour solution ($r_{99}=0.23$, $p=.023$). On the other hand, a significant negative relationship was observed for the umami solution ($r_{99}=-0.25$, $p=.011$). In other words, for sour solutions, increased perceived taste intensity was associated with a higher pitch, while for the umami solutions, increased perceived taste intensity was associated with a lower pitch instead. No significant correlations were observed for the other tastes (for bitter, $r_{99}=-0.15$, $p=.14$; for salty, $r_{99}=0.15$, $p=.15$; for sweet, $r_{99}=0.15$, $p=.15$).

2.3.2.2 Emotion ratings

In terms of participants' emotion ratings of various taste solutions, the mean ratings of valence and arousal are shown in Figure 2.6.

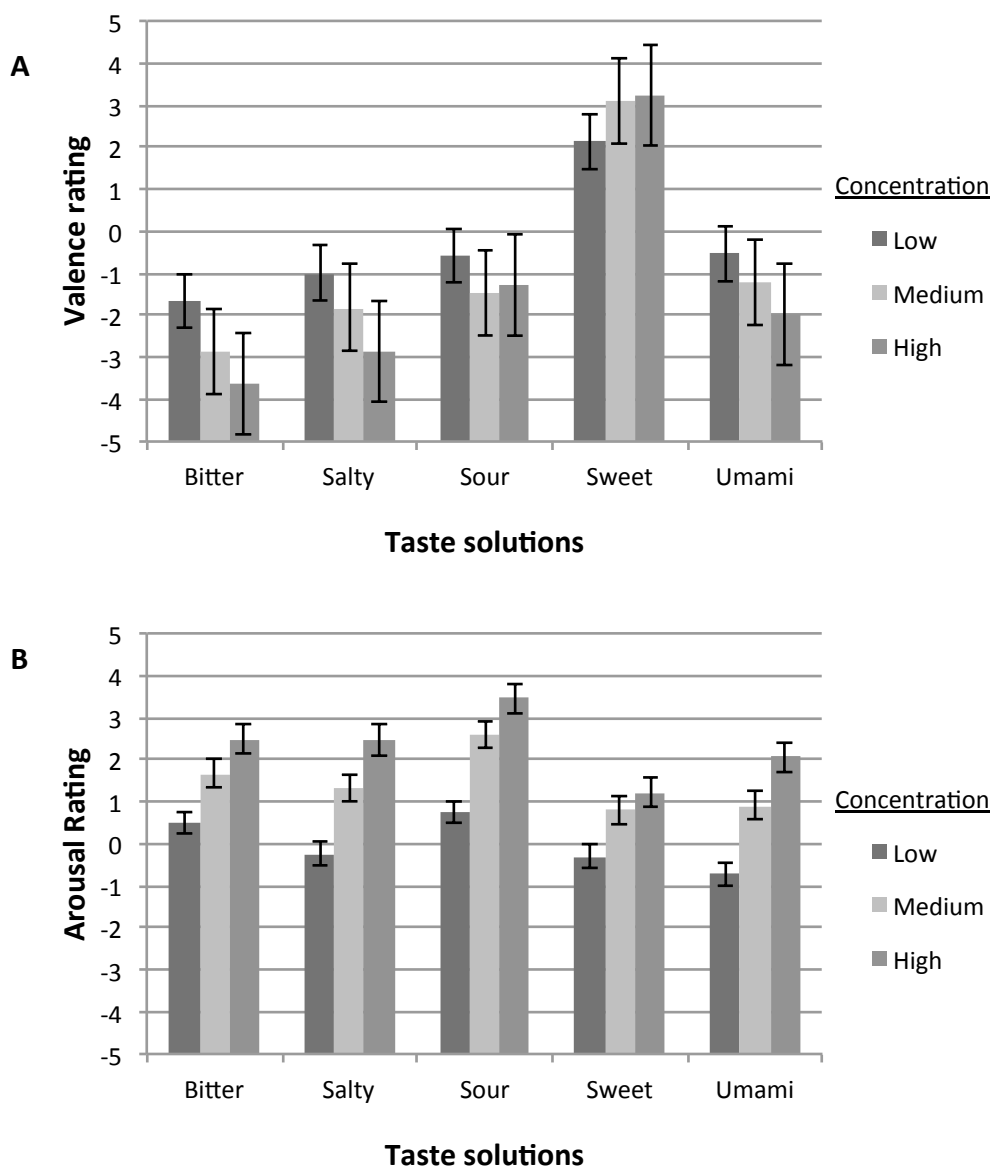


Figure 2.6. Mean values of participants' rating of emotional valence/pleasantness (A) and arousal (B) associated with various taste solutions in different concentrations, in Experiment 1B. Valence is measured on a scale from -5 (extremely unpleasant) to 5 (extremely pleasant). Arousal is measured on a scale from -5 (not arousing at all) to 5 (extremely arousing). The error bars denote the standard error.

RMANOVA tests with taste and concentration as factors were conducted on participants' ratings of valence and arousal. In terms of participants' choice of

valence, a main effect of taste solution type was observed, $F(3.43, 109.67)=62.55$, $p<.0005$, $\eta^2=.66$, using Huynh-Feldt correction. Unsurprisingly, pairwise comparisons revealed that the sweet solution was matched to a significantly higher valence ($M=2.82$, $SE=.24$) than the bitter ($M=-2.72$, $SD=.19$), salty ($M=-1.89$, $SD=.28$), sour ($M=-1.11$, $SD=.34$), and umami ($M=-1.25$, $SD=.32$) solutions, regardless of the concentration at which the tastant was presented ($p<.0005$ for all comparisons). In addition, the bitter solutions were significantly less pleasant than the sour, sweet, and umami solutions. A significant main effect of concentration was also observed, $F(2,64)=19.17$, $p<.0005$, $\eta^2=.38$. Pairwise comparisons revealed that the low concentration solutions ($M=-0.33$, $SE=.14$) were rated significantly higher on the valence scale (i.e., as more pleasant) than the medium ($M=-0.85$, $SE=.18$, $p=.003$) and the high concentration solutions ($M=-1.31$, $SE=.19$, $p<.0005$); in addition, the medium concentration solutions were more pleasant than the high concentration solutions ($p=.015$). An interaction effect was also observed between the taste of the solution and its concentration, $F(8,256)=8.68$, $p<.0005$, $\eta^2=.21$. In particular, as concentration increased, valence ratings were higher for sweet solutions, but lower for bitter, salty, and umami solutions. Specifically, low concentration sweet solutions ($M=2.12$, $SE=.25$) were rated as significantly less pleasant than the medium ($M=3.10$, $SE=.29$, $p=.006$) and high concentration ($M=3.24$, $SE=.33$, $p=.002$) sweet solutions. High concentration bitter solutions ($M=-3.64$, $SE=.23$) were rated as significantly less pleasant than the low ($M=-1.67$, $SE=2.7$, $p<.0005$) and medium concentration ($M=-2.85$, $SE=.25$, $p=.043$) bitter solutions; the medium concentration bitter solution was also rated as less pleasant than the low concentration solution ($p<.0005$). The high concentration salty solution ($M=-2.85$, $SE=.32$) was rated as less pleasant than both the low ($M=-1.00$, $SE=.25$, $p<.0005$) and medium concentration ($M=-1.82$, $SE=.44$,

$p=.01$) salty solutions. The medium concentration sour solution ($M=-1.46$, $SE=.35$) was less pleasant than the low concentration sour solution ($M=-.58$, $SE=.37$, $p=.009$). Finally, the high concentration umami solution ($M=-2.00$, $SE=.46$) was rated as significantly less pleasant than the low ($M=-0.55$, $SE=.24$, $p=.006$) and medium ($M=-1.21$, $SE=.42$, $p=.045$) concentration umami solutions.

In terms of participants' choice of arousal, a main effect of concentration was observed, $F(2, 64)=51.41$, $p<.0005$, $\eta^2=.62$. Pairwise comparisons revealed that the low concentration ($M=-0.006$, $SE=.17$) solutions were rated as significantly less arousing than the medium ($M=1.46$, $SE=.22$) and high concentration solutions ($M=2.33$, $SE=.20$), and that medium concentration solutions were less arousing than high concentration solutions ($p<.0005$ for all comparisons). A main effect of taste solution type on arousal ratings was also observed, $F(3.64, 116.55)=8.90$, $p<.0005$, $\eta^2=.22$, using Huynh-Feldt correction. Pairwise comparisons revealed that the sour solution ($M=2.27$, $SE=.21$) was rated as significantly more arousing than the salty ($M=1.18$, $SE=.25$, $p=.002$), sweet ($M=0.57$, $SE=.29$, $p<.0005$) and umami ($M=0.75$, $SE=.25$, $p<.0005$) solutions.

Pearson correlations between participants' choices of pitch, intensity ratings, and their emotion ratings (valence and arousal) were then computed over all taste solutions. As many of the variables are correlated, partial correlation coefficients were calculated in order to control for the effect of possible third variables. For pitch, positive moderate correlations between pitch and valence ($r_{495}=0.39$, $p<.0005$) and between pitch and arousal ($r_{495}=0.13$, $p=.004$) were found. For intensity, a moderate positive correlation was observed between intensity rating and arousal ($r_{495}=0.32$, $p<.0005$) but there was no significant correlation between intensity rating and valence ($r_{495}=0.034$, $p=.45$). In

addition, a significant correlation was observed between arousal and valence, $r_{495} = -.37$, $p < .0005$.

As in Experiment 1A, a multiple linear regression was used to test whether participants' rating of the valence and arousal of a given taste solution significantly predicted their choice of matching pitch. The results of the regression indicated that for pitch, valence ($\beta = 0.43$, $p < .0005$) and arousal ($\beta = 0.17$, $p < .0005$) accounted for 16% of the variance, $R^2 = 0.156$, $F(2, 492) = 45.53$, $p < .0005$.

Analysing the emotional factors separately, valence alone accounted for 13% of the variance in pitch choice, $\beta = 0.36$, $p < .0005$, $F(1, 494) = 73.96$, $p < .0005$. By contrast, arousal alone did not predict the participants' choice of pitch.

Extending the results of Experiment 1A, data about participants' emotion ratings of all 19 musical notes from C2 to C8 was also collected. Mean values of valence and arousal ratings for the notes can be seen in Figure 2.7. It is worth noting that valence ratings make an inverted U-shape, where pitches that are at the low (approximately pitches 1-5, or C2 to E3) or high (approximately pitches 16-19, or C7-C8) end are more unpleasant than the notes in the middle. Inversely, arousal ratings make a U-shape, where pitches at the low and high end of the keyboard are more arousing than those notes in the middle.

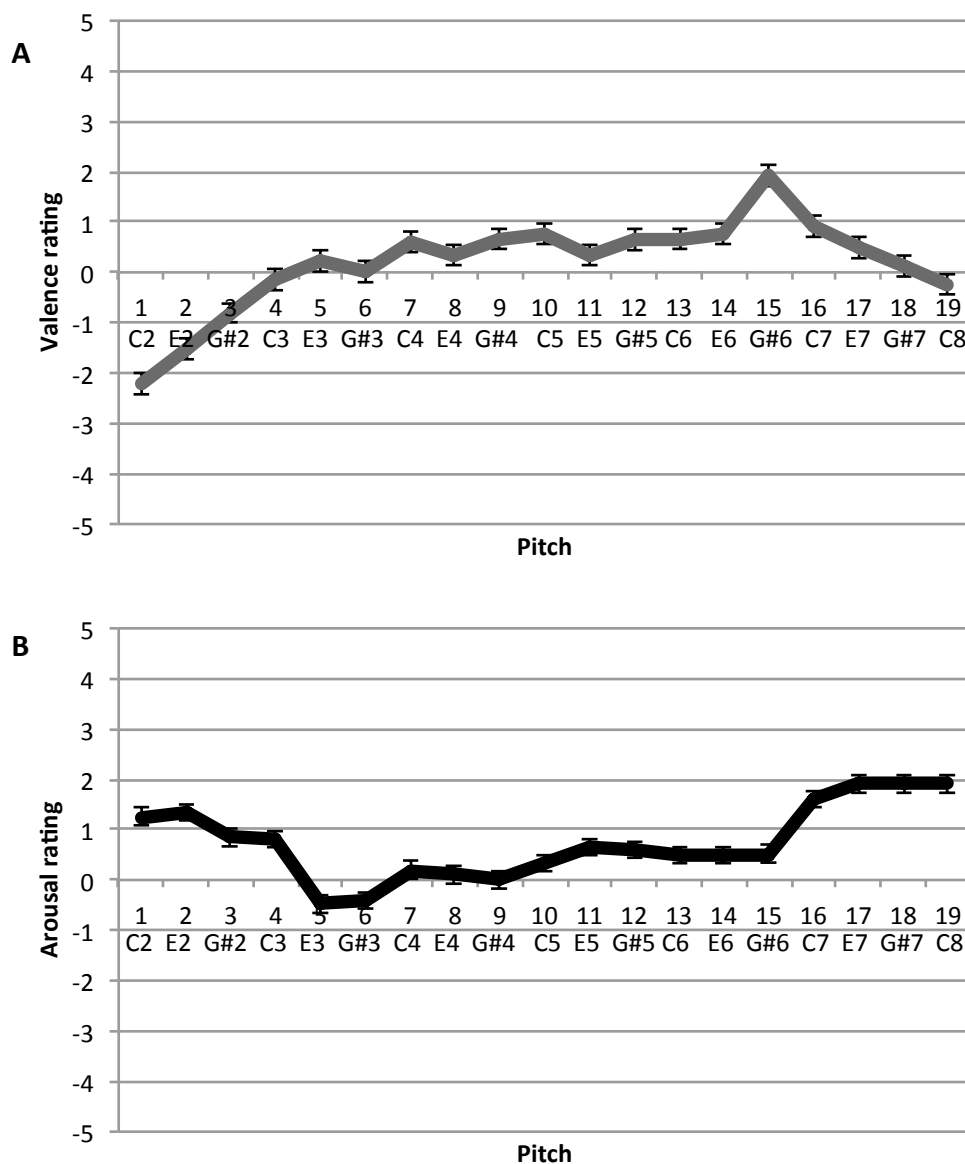


Figure 2.7. Mean values of participants' rating of emotional valence/pleasantness (A) and arousal (B) associated with music notes of different pitch (1=C2, 19=C8), in Experiment 1B. Valence was measured on a scale from -5 (extremely unpleasant) to 5 (extremely pleasant). Arousal was measured on a scale from -5 (not arousing at all) to 5 (extremely arousing). The error bars denote the standard error.

Using Palmer et al.'s (2013) method, the correlation between the emotion ratings of each taste and the emotion ratings of the note that it matched with were examined (see Figure 2.8). There were two separate analyses for valence and arousal. In terms of valence, there was a significant positive correlation between valence ratings for each

taste sample and the note that best matched the sample, $r=0.55$, $p=.034$. However, no significant correlation between arousal ratings of taste samples and matching notes was found, $r=0.35$, $p=.195$.

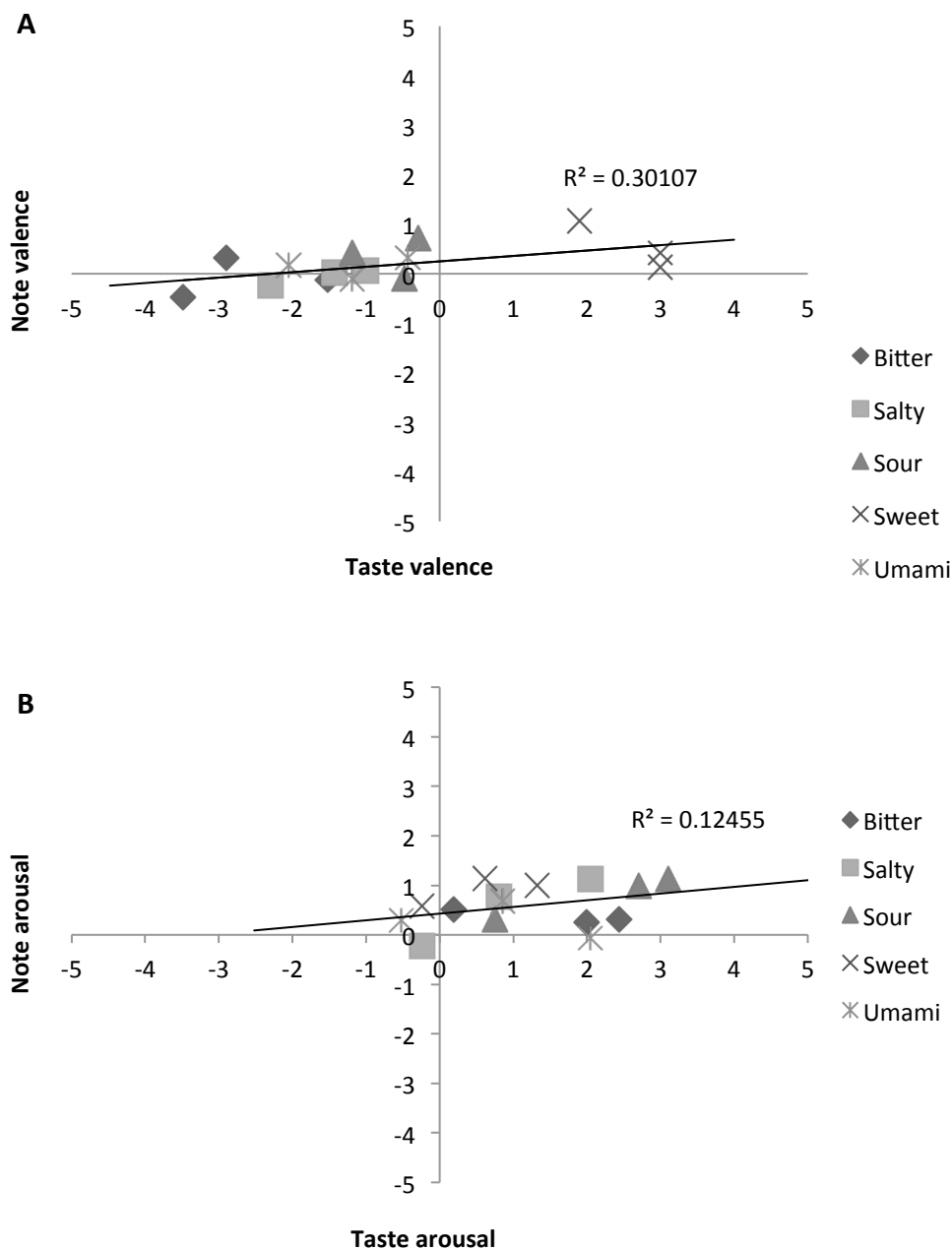


Figure 2.8. Scatterplots and correlations between the emotional ratings of the 15 taste solutions (x axis) and the emotional associations of the pitch chosen as matching best with them (y-axis), for the two emotional dimensions of valence (A) and arousal (B), in Experiment 1B.

As in Experiment 1A, differences in taste quality (bitter/salty/sour/sweet/umami) exerted a significant influence over the participants' choice of pitch and perceived taste intensity as well as their ratings of arousal and valence. With increasing taste concentration, the participants rated solutions as more intense (as expected) as well as giving higher arousal ratings and lower valence ratings (except for sweet tastes, where increasing taste concentration increased valence ratings). In addition, a weak positive correlation was observed between perceived taste intensity and pitch choice, suggesting that taste intensity does affect taste-pitch matching, but to a lesser extent than taste quality.

In Experiment 1B, the major change was to move emotion ratings to a separate block, after participants have already made their pitch selection for each tastant. This was to rule out the possibility that participants did not base their pitch selection on their emotion ratings, which was possible in Experiment 1A. Nevertheless, significant correlations between pitch choices and emotion ratings were still observed across blocks, thus showing that the relationship between pitch choice and emotion rating is genuine.

Given the additional information of emotion ratings for musical notes, it was possible to perform a more complete emotional mediation analysis by comparing emotion ratings of each taste solution with the emotion ratings of the musical note that it best matched with. This is the first time, to the best of my knowledge, where emotional responses to individual musical notes have been measured and used to analyse emotion mediation effects in crossmodal correspondences involving pitch. No significant correlation in the emotional dimension of arousal was observed, but there was a significant positive correlation in the dimension of valence. In other words, the

more pleasant the taste, the more pleasant the musical note chosen to match with the taste. It is worth noting here that this conclusion is different from Crisinel and Spence's (2012) results, where no relationship was found between the pitch that was chosen and the pleasantness of the dark chocolate that was sampled. However, it should be noted that Crisinel and Spence's analysis was limited to a single food item and the participants' emotional response to pitch was not measured.

Moving beyond pitch-taste correspondences, Experiments 2A and 2B focus on specifically designed soundtracks to match with basic tastes. The experiments assess how reliably people can match such taste-congruent soundtracks to taste words, and whether the association between soundtracks and basic tastes is also emotionally mediated.

2.4 Experiment 2A: A comparison of the effectiveness of various soundscapes in evoking specific tastes and the role of emotion mediation in soundscape-taste associations

2.4.1 Methods

2.4.1.1 Participants

100 participants (51 women, 49 men) aged between 21-62 years ($M=34.06$, $SD=9.53$) took part in the study. The participants gave their informed consent, and reported no cold or other impairment of their senses of smell, taste, or hearing. The participants were recruited from Amazon's Mechanical Turk. The experiment was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

2.4.1.2 *Auditory stimuli*

24 RMS-equalised soundtracks were used (5 bitter, 5 salty, 7 sour, and 7 sweet). The soundtracks varied originally from 30 seconds to 6 minutes in length. Since the soundtracks were generally uniform in texture, only the first 15 seconds of each soundtrack was used in order to have uniform-length auditory stimuli and to limit the overall length of the experiment. A brief description of the soundtracks is given below, categorised by composer (in alphabetical order):

Condiment Junkie: A UK sound branding agency, recently released an EP album of taste soundtracks. Excerpts of the sweet, sour, salty, and bitter soundtracks from the album were used in Experiment 2. Note that an earlier version of the sweet and bitter soundtracks had been used in a study previously reported by Crisinel et al. (2012). There, the participants rated samples of bittersweet toffee on a 7-point bitter-sweet scale while listening to the two sounds. Their results revealed that the participants rated the cinder toffee samples (which came from the same batch) significantly differently under the two auditory conditions. In the same study, the two soundtracks were pre-tested in a control experiment in order to ensure that the participants rated the soundtracks differently on a 9-point bitter-sweet scale. As expected, the sweet soundtrack ($M=6.68$, $SD=1.78$) was rated as significantly sweeter on the bitter-sweet scale than the bitter soundtrack ($M=2.97$, $SD=1.14$).

Jialing Deng and Harlin Sun: Designed a set of sweet, bitter, salty, sour, and umami soundtracks for *Synaesthetic Appetiser*, part of Deng's Masters of Arts Thesis project (June, 2015). The stated aim was to create a narrative environment of a synaesthetic world by offering augmented eating experiences through crossmodal interactions. The soundtracks were designed to evoke specific tastes, with the goal of helping those

who suffer from some form of sensory dysfunction and who might not otherwise be able to taste normally.

Evan Kassof: Designed sounds to match each of the four basic tastes. These sounds were used in a citizen science experiment as part of the Science Museum Cravings exhibit in London, UK. The participants could either access the experiment at the gallery or on-line, via the Science Museum's homepage or the Cravings exhibition information page

(http://www.sciencemuseum.org.uk/visitmuseum/Plan_your_visit/exhibitions/cravings/cravings-experiment.aspx). The sounds were composed from four basis soundtracks that, when combined in different ratios, created composite soundtracks that varied in terms of their articulation, pitch, loudness, and consonance (see Knöferle & Spence, 2012, Table 1). In the Science Museum experiment, the participants were presented with individual sounds and had to match them to a taste word (sweet, bitter, sour, salty, umami, or else indicated that they were “unsure”).

Klemens Knoeferle and Florian Käßler: Designed a set of taste soundtracks (sweet, bitter, sour, and salty) inspired by Knoeferle et al.'s (2015, Experiment 1) study in which participants had to match a number of auditory parameters (attack, discontinuity, pitch, roughness, sharpness, and speed) to basic taste words (bitter, sweet, salty, and sour). Based on these results, low-level properties of a 30 second piece of synthesised music were systematically manipulated in order to create soundtracks matching different tastes (Knoeferle et al., 2015, Experiment 2). These soundtracks were then tested online with participants from the USA and India in a matching task where all four soundtracks and four taste words were presented as a group. On average, the North American participants matched 1.75 sounds correctly while the Indian participants matched 1.38 sounds correctly (as compared to matching

1 sound correctly by chance). Importantly, both groups performed at a level that was significantly better than chance.

Bruno Mesz: A group of musicians were initially asked to improvise on a MIDI keyboard based on the taste words sweet, sour, bitter, and salty (Mesz et al., 2011). Based on those improvisations, different loudness, pitch, duration, and articulation features were extracted for each taste. Those features were then used to design an algorithm that automatically generated music of specific tastes by combining fragments of classical and popular music that matched the aforementioned auditory features (Mesz et al., 2012). Five pieces from Mesz are tested in Experiment 2A and 2B:

1. Makea (sweet soundtrack) was produced by the algorithm described in Mesz et al. (2012). The piece consists of a mosaic of musical quotations from Beethoven and Debussy, subsequently modified by hand.
2. Tango (sour soundtrack) was a collage of short tango excerpts, transposed, accelerated, and instrumented for piccolos and clarinets. This soundtrack was used in Kontukoski et al. (2015) as a sour music sample.
3. Hapan (sour soundtrack) was a combination of composed piano part in addition to algorithmically generated (Mesz et al., 2012) instrumental parts based on tango music.
4. Suolainen (salty soundtrack) was composed for marimba and un-pitched percussion.
5. Beethoven texture (sweet soundtrack) used 5 bars from the start of the second movement of Beethoven's Quartet Op. 127, reading short fragments forwards and backwards to build an 80-voice canon.

Felipe Reinoso Carvalho: In collaboration with the Institute for Psychoacoustics and Electronic Music at the University of Ghent, Reinoso Carvalho designed sweet and bitter soundtracks. The soundtracks were produced by Tim Vets @ IPEM, UGent; co-produced by Felipe Reinoso Carvalho, Sander de Keer and Tomas Serrine; and mastered by Felipe Reinoso Carvalho at sonictaste.flavours.me (2013).

The bitter and sweet soundtracks (along with a neutral soundtrack which was not used in Experiment 2A and 2B) were used in a recent chocolate-tasting study (see Reinoso Carvalho et al., 2016). These soundtracks were designed to be congruent with the taste of bitter and sweet chocolate, respectively. The soundtracks were inspired by previous work (Crisinel & Spence 2010b; Crisinel et al., 2012). The ‘bitter’ soundtrack consisted of complex overtones, low resonance filters, and static pulses as the result of a saw tooth wave function. The ‘sweet’ soundtrack involved high resonant filters with round-bubbling sounds.

The sounds were first tested in an online study by 78 participants, who rated the soundtracks as significantly different on a 6 point bitter-sweet scale. During the actual study, the participants first matched chocolate samples to the soundtracks and then tasted the samples while listening to the soundtracks.

Qian (Janice) Wang: Designed a sour soundtrack for a study on the effect of taste-congruent sounds on taste evaluation, as part of her Masters degree at the MIT Media Lab (Wang, 2013). The soundtrack was composed in Ableton Live with features of high pitch, fast tempo, and high dissonance (Mesz et al., 2011). The soundtrack consisted of notes played by synthetic instruments. The pitch of the notes ranged from C2 to C6. During the study, the participants listened to bitter, sweet, and sour soundtracks (bitter and sweet soundtracks were the same as used in Crisinel et al., 2012) while rating juice samples in terms of their bitterness, sweetness, and sourness.

The soundtrack excerpts selected for use in Experiment 2A and 2B can all be heard at <https://soundcloud.com/janicewang09/sets/taste-soundscapes-test>. See Table 2.2 for a list of all soundtracks, their usages, and any previous tests of taste associations. Note that the italicized soundtracks were not used here.

Composer/ Designer	Taste soundtracks	Usage	Previous testing
Condiment Junkie	Bitter, sweet, sour, salty, & <i>umami</i>	Featured in EP Album. Earlier versions of the bitter and sweet soundtracks were used in Crisinel et al.'s (2012) study to test whether music can bias taste ratings of bittersweet toffee	Earlier versions of the bitter and sweet soundtracks were tested by Crisinel et al. (2012) in a control study. 31 participants rated these soundtracks as significantly different on a bitter-sweet scale
Deng	Bitter, sweet, sour, salty, & <i>umami</i>	MA thesis on multisensory eating experiences	None
Kassof	Bitter, sweet, sour, & salty	Used in study at Science Museum's Cravings exhibit. The participants had to match a basic taste word to each soundtrack. The soundtracks were presented independently and not all of the participants heard all of the soundtracks	Tested as part of Science Museum exhibit. Participant's choice of taste word matches were significantly different from random. Bitter taste was mostly often matched (for bitter, sour, and salty

			soundtracks). The sweet soundtrack was most often matched to sourness.
Knoeferle	Bitter, sweet, sour, & salty	Used in Knoeferle et al.'s (2015) study where participants from India and the U.S. assigned four basic taste words to the four soundtracks presented together	Tested by Knoeferle et al. (2015). On average, U.S. participants assigned 1.75 of the 4 sounds to the correct taste word, while the Indian participants assigned 1.38 sounds correctly. Both groups performed at a level that was better than chance
Mesz	Sweet, sour, salty	The sweet Makea soundtrack was used by Mesz et al., (2012). The sour tango piece was used by Kontukoski et al. (2015). The soundtracks have also been presented in a number of multisensory performances	The sweet Makea soundtrack was used in Mesz et al. (2012), where it was tested by in a four-taste four-soundtrack matching task
Reinoso Carvalho	Bitter, sweet, <i>neutral</i>	Used in Reinoso Carvalho et al. (2015b) to study how participants matched soundscapes to different chocolates, and how soundscapes influenced the taste ratings of chocolates	A control experiment was conducted in which 78 participants rated the soundtracks as significantly different on a bitter-sweet scale (the three soundtracks were presented together and the participant could compare between them)

Wang	Sour	Used to test perceptual effects of sweet, sour, and bitter music on drinks (Wang 2013)	None
------	------	--	------

Table 2.2. A list of all soundtracks in Experiments 2A and 2B, grouped by composer/designer, detailing their usages and previous tests of taste associations.

2.4.1.3 Design and procedure

The experiment was programmed on the Xperiment experiment-design and hosting platform. Before the actual study began, the participants specified their gender, age, country of origin, and self-rated their musical expertise levels (the choices were none, amateur, intermediate, or advanced). The participants had to listen to all 24 soundtracks in a random order. After each soundtrack, the participants had to choose which basic taste (sweet, sour, bitter, or salty) best matched the sound clip that they had just heard, and rate how confident they were in having decoded the correct taste on a scale from 0-100. The presentation of taste choices was randomised for each trial.

The study lasted for approximately 10 minutes and the participants were paid \$1.20 USD for the study.

2.4.2 Results and discussion

The choices for the best-matching taste word were tallied for each soundtrack (see Figure 2.9). The soundtracks with the highest rate of matching for each taste were: for bitterness, Condiment Junkie (42% of participants matched it with bitter, and they were 1.82 times more likely to match it with bitter than with salty, the next most popular choice); for saltiness, Deng (58% of participants matched it with salty, and

were 2.76 times more likely to match it with salty than with sour, the next most popular choice); for sourness, Mesz – Tango (58% of participants matched it with sour, and were 1.76 times more likely to match it with sour than with bitter, the next most popular choice); and for sweetness, Deng (89% of participants matched it with sweet, and were 14.83 times more likely to match it with sweet than with salty, the next most popular choice). In comparison, had the participants been responding randomly, 25% of them would have matched each taste (bitter, salty, sour, and sweet) to any given soundtrack.

Averaging over all of the soundtracks that were associated with each taste, 31.4% of the participants' responses were correct for bitter soundtracks, 44.4% of participants' responses were correct for salty soundtracks, 41.7% of the participants' responses were correct for sour soundtracks, and 56.9% of participants' responses were correct for sweet soundtracks (see Figure 2.10). A chi-square test for independence was conducted in order to assess whether different taste words were chosen for different taste soundtracks, tallied over all of the soundtracks that had been generated for each taste. The results indicated that the different taste soundtracks influenced the choice of taste words [$\chi^2(9, 2400)=720.62, p<.0001$]. The strength of this effect, measured by computing Cramer's V, can be classified as medium ($V=.32$).

Soundtrack-taste matches

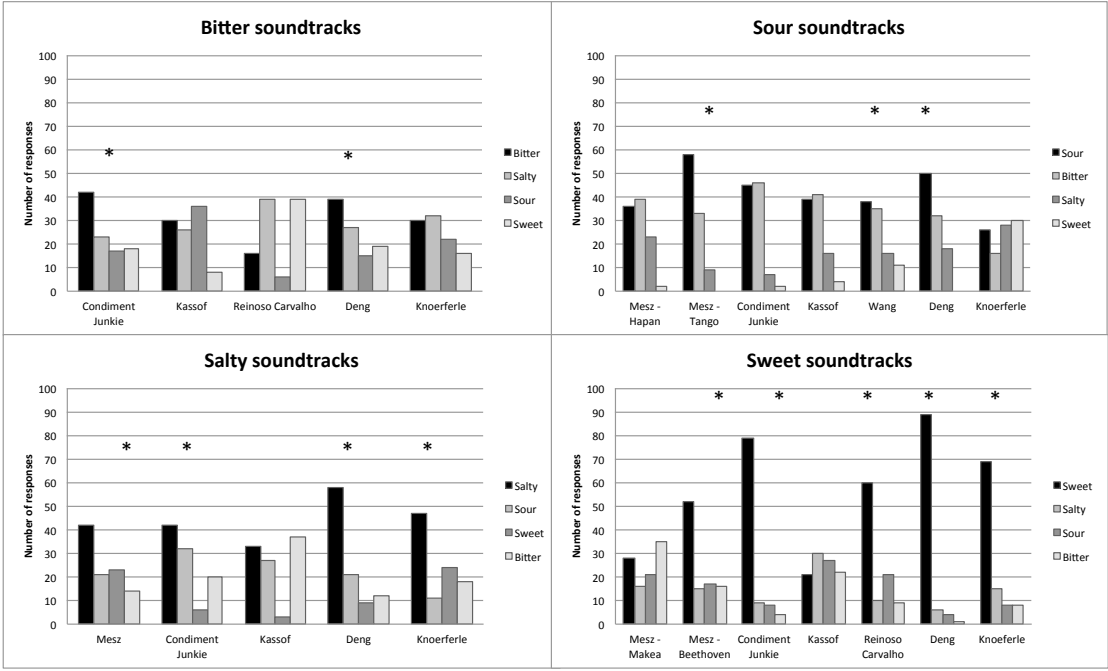


Figure 2.9. Results of participants’ soundtrack-taste matching organised by each taste that the soundtracks were designed to evoke in Experiment 2A. The chart reveals the distribution of participants’ choice of taste matches for each soundtrack, with a total of 100 responses for each soundtrack. Soundtracks with both non-random distributions of taste ratings and with the highest matching taste being the “correct” one are shown with asterisks.

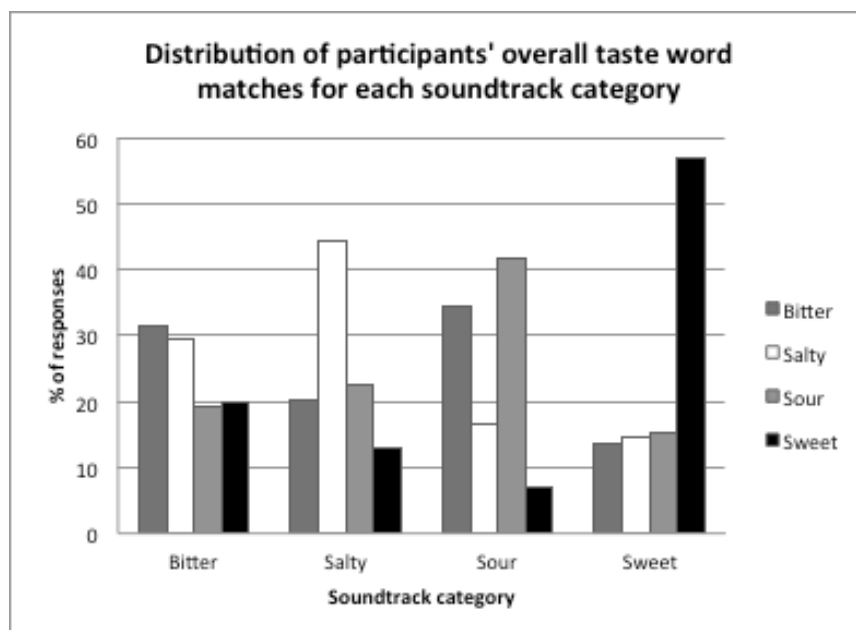


Figure 2.10. Results of participant's soundtrack-taste matching exercise, with soundtracks grouped by tastes they are designed to evoke in Experiment 2A.

A chi-square test of goodness of fit was calculated for each soundtrack to determine which of them induced a distribution of taste matches that was significantly different from chance. In fact, out of the 24 soundtracks, only 3 had non-significant preferences in the choice of taste matches (Knoeferle-bitter, Knoeferle-sour, Kassof-sweet). The soundtracks with significantly non-random taste matches were, for the bitter soundtracks: Condiment Junkie [$X^2(3, 100)=16.24, p=.001$], Kassof [$X^2(3, 100)=17.44, p=.001$], Reinoso Carvalho [$X^2(3, 100)=33.36, p<.0005$], and Deng [$X^2(3, 100)=13.44, p=.004$]; for the salty soundtracks: Mesz [$X^2(3, 100)=17.20, p=.001$], Condiment Junkie [$X^2(3, 100)=28.96, p<.0005$], Kassof [$X^2(3, 100)=27.84, p<.0005$], Deng [$X^2(3, 100)=61.20, p<.0005$], and Knoeferle [$X^2(3, 100)=29.20, p<.0005$]; for the sour soundtracks: Mesz – Hapan [$X^2(3, 100)=34.00, p<.0005$], Mesz – Tango [$X^2(3, 100)=36.02, p<.0005$], Condiment Junkie [$X^2(3, 100)=67.76, p<.0005$], Kassof [$X^2(3, 100)=38.96, p<.0005$], Wang [$X^2(3, 100)=21.84, p<.0005$],

and Deng [$X^2(3, 100)=15.44, p<.0005$]; for the sweet soundtracks: Mesz – Makea [$X^2(3, 100)=8.24, p=.04$], Mesz – Beethoven [$X^2(3, 100)=38.96, p<.0005$], Condiment Junkie [$X^2(3, 100)=156.08, p<.0005$], Reinoso Carvalho [$X^2(3, 100)=68.88, p<.0005$], Deng [$X^2(3, 100)=218.96, p<.0005$], and Knoeferle [$X^2(3, 100)=104.56, p<.0005$]. Those soundtracks with non-random distributions of responses, and where the intended taste for the soundtracks was the most chosen, are starred in Figure 2.9.

To check that there was no response bias from the participants (i.e., there was not some taste words that were simply chosen more frequently than others), the total number of responses given to each taste word was computed. There were a total of 595 bitter responses, 587 salty responses 606 sour responses, and 612 sweet responses. A chi-square test of goodness of fit was calculated and there were no distributions significantly different from chance [$X^2(3, 2400)=0.62, p=.89$].¹ In other words, there was no evidence of any response bias from the participants.

Out of 100 participants, 29 identified themselves as having no musical expertise, 51 as amateurs, 17 as intermediates, and 3 as advanced. Participants were assigned to two groups, those with no expertise (29 people) and those with some expertise (71 people). For each soundtrack, a chi-square test of independence was performed to determine whether there was an association between musical expertise and choosing the intended taste word to match the soundtracks (see Additional Materials). Did musical expertise contribute to the participants' ability to match tastes? Out of 24 soundtracks, only Mesz' sweet Makea soundtrack was found to have a different distribution of right/wrong choices depending on the self-reported musical expertise

¹ In fact, given such an even distribution of responses, one might want to consider whether the participants were perhaps implicitly equalizing their responses assuming there would be, overall, equal numbers of soundtracks of each taste (see Erlebacher & Sekuler, 1971).

of the participant. In particular, a higher percentage of those participants without musical expertise chose to match the soundtrack with sweetness than participants with musical expertise.

The confidence ratings for each soundtrack are shown in Figure 2.11. In addition to the average confidence ratings of all the participants for each soundtrack, the confidence ratings for only the “correct” choices are also included.

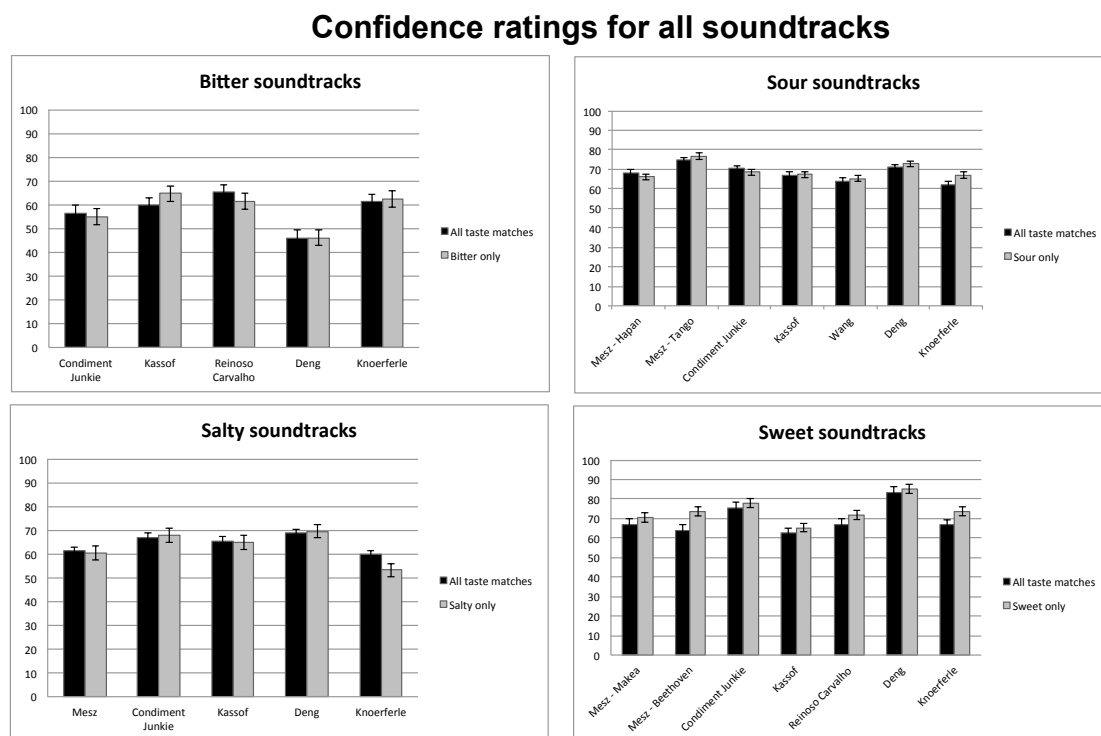


Figure 2.11. Confidence ratings of taste matches for all of the soundtracks, organised by each taste that the soundtracks were designed to evoke in Experiment 2A. Confidence ratings were given on a scale from 0-100, where 0 was the least confident and 100 the most. For each soundtrack, the average confidence ratings over all taste matches are shown in addition to the confidence levels for the “correct” taste match. The error bars represent the standard error of means.

On average, the confidence level for the bitter soundtracks was 57.80, SD=17.57, for the salty soundtracks M=64.50, SD=17.03, for the sour soundtracks M=68.18, SD=16.18, and for the sweet soundtracks M=69.53, SD=15.78 (see Figure 2.12).

ANOVA with Huynh-Feldt corrections was performed on the average of the confidence ratings, where significant differences were found between the confidence ratings ($F(2.59,297)=51.01$, $p<.0005$, $\eta^2=0.34$). Specifically, pairwise comparisons with Bonferroni corrections revealed that the confidence ratings for the bitter soundtracks were significantly lower than for all of the other soundtracks ($p<.0005$ for all), the confidence ratings for the salty soundtracks were significantly higher than for the bitter soundtracks ($p<.0005$) but lower than for the sweet ($p<.0005$) and sour ($p=.001$) soundtracks, and confidence ratings for sour and sweet soundtracks were not significantly different from each other, but both were significantly higher than for bitter and salty.

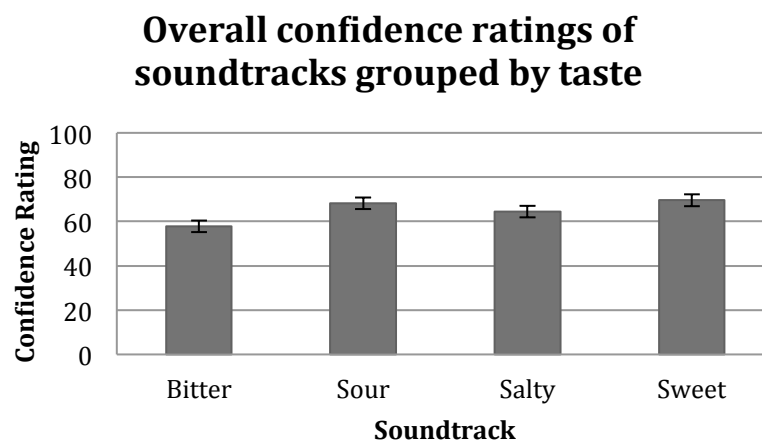


Figure 2.12. Overall confidence ratings for the various soundtracks in Experiment 2A. Participants rated their taste match choices on a scale of 0-100 for each soundtrack. The error bars represent the standard error of means.

Overall, based on the taste-matching and confidence rating data, the results of Experiment 2 demonstrate that out of 24 soundtracks [5 bitter, 5 salty, 7 sour, and 7 sweet] the sweet soundtracks most effectively evoked the desired taste, while the bitter soundtracks were the least effective. Specifically, 56.9% of the participants' responses were correct for sweet soundtracks, and confidence ratings for sweet

soundtracks were significantly higher ($M=69.53$, $SD=15.78$) than for bitter and salty soundtracks; on the other hand, 31.4% of participants' responses were correct for bitter soundtracks, and confidence ratings were significantly lower for bitter soundtracks ($M=57.80$, $SD=17.57$) than for all other soundtracks. The reason why the participants found it easiest to match soundtracks to sweetness is possibly because, out of all the tastes, people typically like sweetness most (Robin et al., 2003; also see Figure 2.6A for evidence that sweet solutions were the most distinctly liked by participants, compared to all other tastes). Perhaps there is a straightforward association here for participants with music and soundscapes that they find pleasant, whereas all the other tastes become associated with unpleasant music.

In addition, it is perhaps worth noting that sweetness is the only taste term that is also used in musical notation; i.e., “dolce”, meaning to play in a gentle, sweet style (see Fallows, s.a.). Therefore participants - especially those with musical training - might more readily associate music with sweetness (though, it should be said, it was not observed in Experiments 1A and 1B). Here it is also worth noting that the finding where sweetness somehow stands out does not only apply to sound-taste correspondences; in fact, in shape-taste matching studies, sweetness is consistently matched to round shapes whereas all of the other tastes are consistently matched to shapes that are more or less angular (Velasco et al., 2015, 2016).

Given the above-chance results of participants in matching soundtracks to their intended tastes, the question remains as to whether emotion played a role in mediating the associations between the soundtracks and tastes, and whether the degree and type (pleasantness or arousal) of emotion mediation differed for each basic taste. These questions are addressed in Experiment 2B.

2.4.3 Additional materials

Taste-matching counts from soundtrack-taste matching task in Experiment 2A, sorted by musical expertise.

Bitter soundtracks						
Experience		Condiment Junkie	Kassof	Reinoso Carvalho	Deng	Knoeferle
None	Bitter	14	7	4	10	12
	Other	15	22	25	19	17
Some	Bitter	28	23	12	29	18
	Other	43	48	59	42	53

Salty soundtracks						
Experience		Mesz Junkie	Condiment	Kassof	Deng	Knoeferle
None	Salty	11	10	8	16	17
	Other	18	19	21	13	12
Some	Salty	31	32	25	42	30
	Other	40	39	46	29	41

Sour soundtracks								
Experience		Mesz – Hapan	Mesz – Tango	Condiment Junkie	Kassof	Wang	Deng	Knoeferle

None	Sour	12	17	13	10	12	13	4
	Other	17	12	16	19	17	16	25
Some	Sour	24	41	32	29	26	37	22
	Other	47	30	39	42	45	34	49

Sweet soundtracks								
Experience		Mesz - Makea	Mesz - Beethoven	Condiment Junkie	Kassof	Reinoso Carvalho	Deng	Knoeferle
None	Sweet	14	13	24	5	15	27	19
	Other	15	16	5	24	14	2	10
Some	Sweet	14	39	55	16	45	62	50
	Other	57	32	16	55	26	9	21

2.5 Experiment 2B: A comparison of the effectiveness of various soundscapes in evoking specific tastes and the role of emotion mediation in soundscape-taste associations

2.5.1 Methods

2.5.1.1 Participants

50 participants (21 women, 29 men) aged between 20-64 years ($M=35.71$, $SD=11.30$) took part in the study. The participants gave their informed consent, and reported no cold or other impairment of their senses of smell, taste, or hearing. The

participants were recruited from Amazon's Mechanical Turk. The experiment was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

2.5.1.2 Auditory stimuli

The auditory stimuli were the same as used in Experiment 2A.

2.5.1.3 Design and procedure

The experiment was programmed on the Xperiment experiment-design and hosting platform. Before the actual study began, the participants specified their gender, age, country of origin, and self-rated their musical expertise levels (the choices were none, amateur, intermediate, or advanced). The participants had to listen to all 24 soundtracks in a random order. After each soundtrack, the participants had to rate how pleasant and how energising/exciting they found the soundtrack, each on a scale from 0-100. At the end of the test, participants were also asked how pleasant and how energising/exciting they found bitter/salty/sour/sweet-tasting foods.

The study lasted for approximately 10 minutes and the participants were paid \$1.20 USD for the study.

2.5.2 Results and discussion

The mean pleasantness and arousal ratings for each soundtrack are shown in Figure 2.13. The mean pleasantness and arousal ratings for foods of a given taste are shown in Figure 2.14.

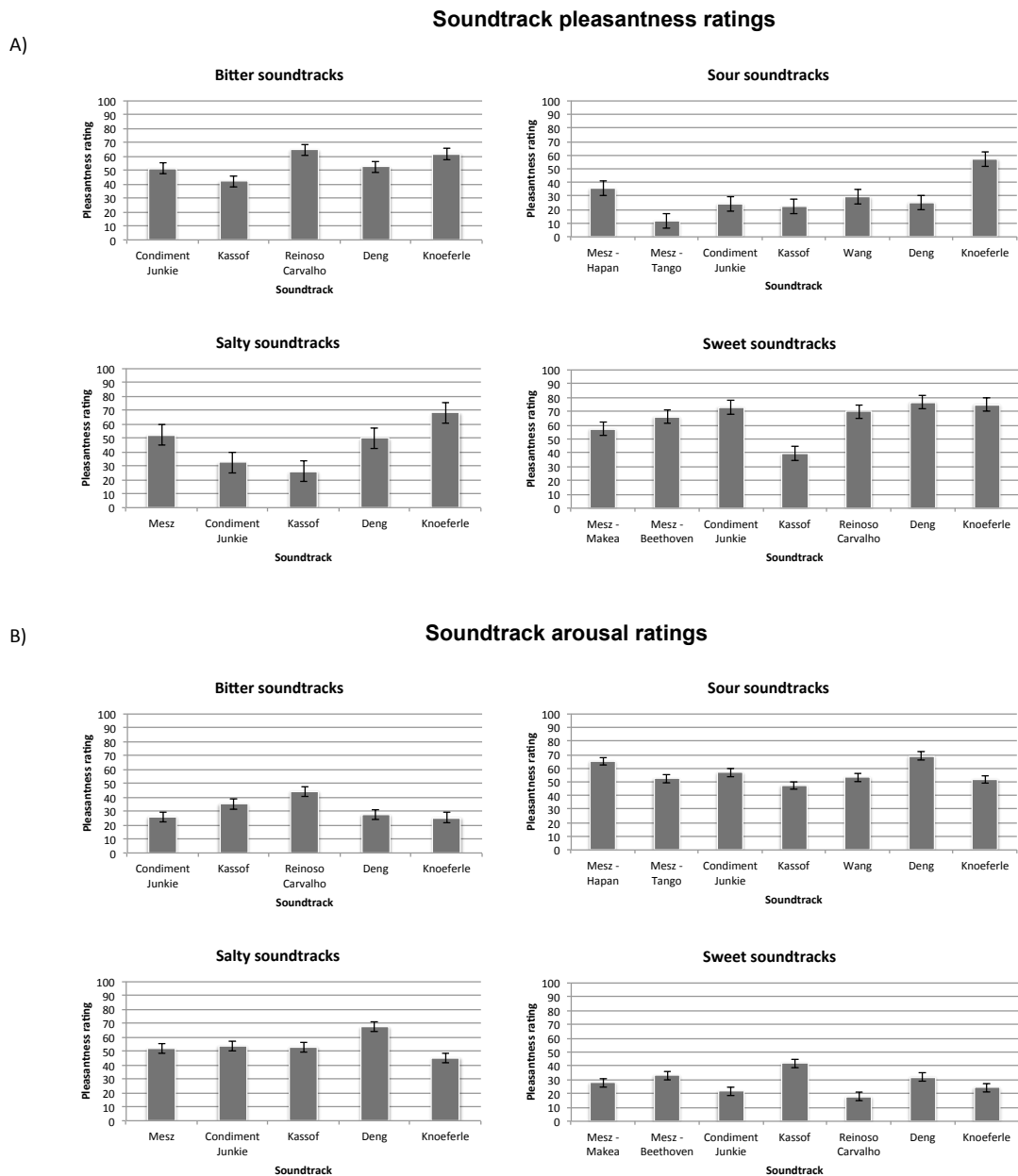


Figure 2.13. Mean pleasantness (A) and arousal (B) rating for each soundtrack, organised by each taste that the soundtracks were designed to evoke in Experiment 2B. Pleasantness and arousal are each rated on a scale of 0-100. The error bars represent the standard error of means.

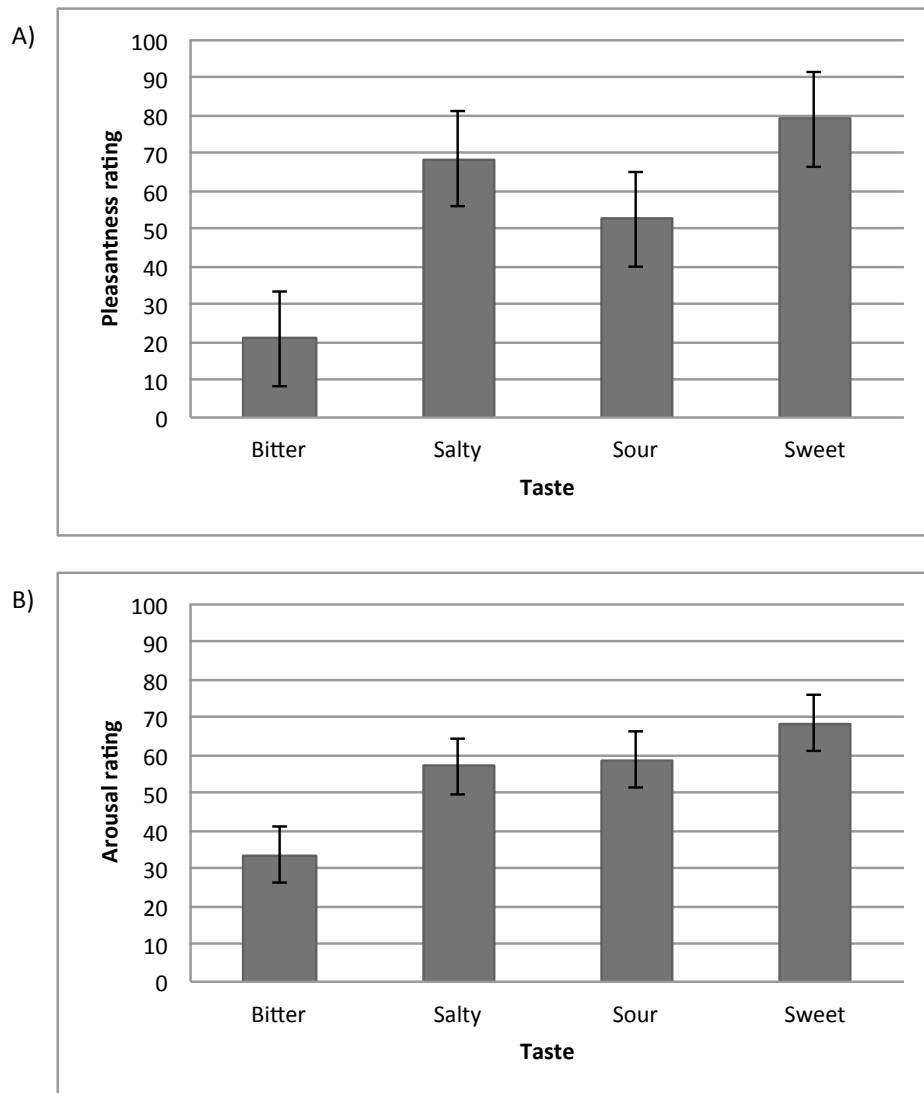


Figure 2.14. Mean pleasantness (A) and arousal (B) rating for each taste word, each on a scale of 0-100 in Experiment 2B. The error bars represent the standard error of means.

To analyse the role of pleasantness and arousal in mediating the matching between soundtracks and a given taste (T), the method documented in Schifferstein and Tanudjaja (2004) was used. For each of the $24 \times 4 = 96$ possible soundtrack-taste combinations, the absolute difference between the mean pleasantness rating of the soundtrack and the mean pleasantness rating of the given taste was calculated. Next, the correlation between this soundtrack-taste difference and the % of responses that matched the given taste to the soundtrack was computed. If the pleasantness of a taste

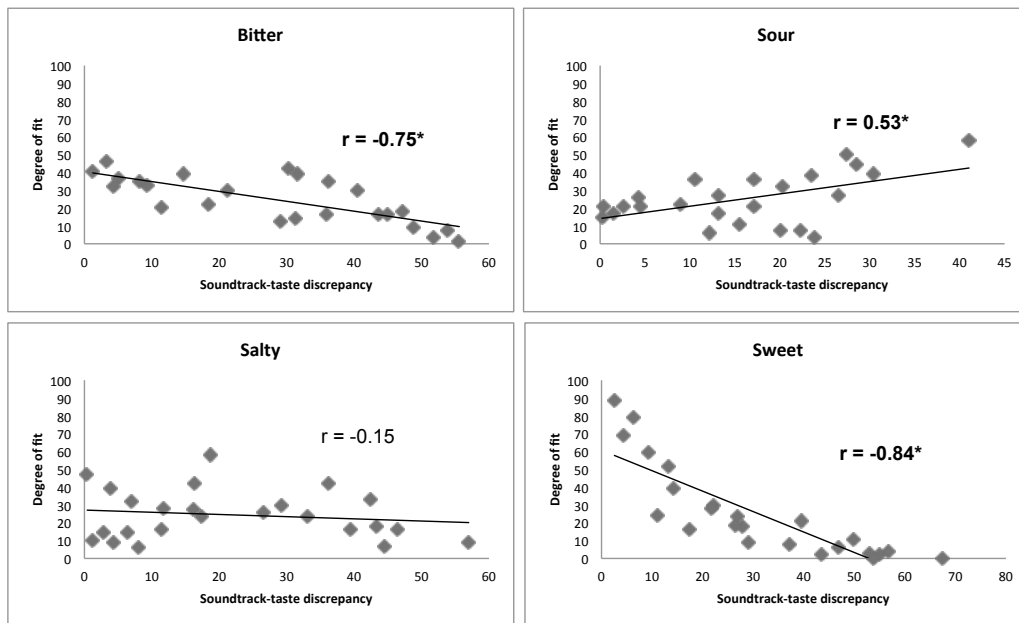
T and a soundtrack S influence their crossmodal correspondence, then one would expect the measure of pleasantness difference between taste T and soundtrack S to be negatively related to the % responses that match taste T to soundtrack S. The same procedure was repeated for arousal.

For pleasantness, Pearson correlation coefficients were -0.85 ($p < .0005$) for bitterness, -0.15 ($p = .48$) for saltiness, 0.53 ($p = .008$) for sourness, and -0.84 ($p < .0005$) for sweetness (see Figure 2.15A for plots). This suggests that pleasantness partly mediates the soundtrack-taste correspondence for bitterness and sweetness.

For arousal, Pearson correlation coefficients were of 0.19 ($p = .39$) for bitterness, -0.323 ($p = .12$) for saltiness, -0.60 ($p = .002$) for sourness, and 0.67 ($p < .0005$) for sweetness (see Figure 2.15B for plots). This suggests that arousal partly mediates the soundtrack-taste correspondence for sourness.

Furthermore, partial correlation coefficients were calculated to assess whether the effects for pleasantness and arousal are independent. The coefficient for pleasantness remained significant for soundtrack-bitter taste ($r = -0.78$, $p < .0005$) and soundtrack-sweet taste ($r = -0.72$, $p < .0005$) matches (controlling for the effect of arousal). Similarly, the coefficient for arousal ($r = -0.53$, $p = .009$) remained significant for soundtrack-sour taste matches (controlling for the effect of pleasantness).

A) **Measure of pleasantness mediation in soundtrack-taste matching, for each taste**



B) **Measure of arousal mediation in soundtrack-taste matching, for each taste**

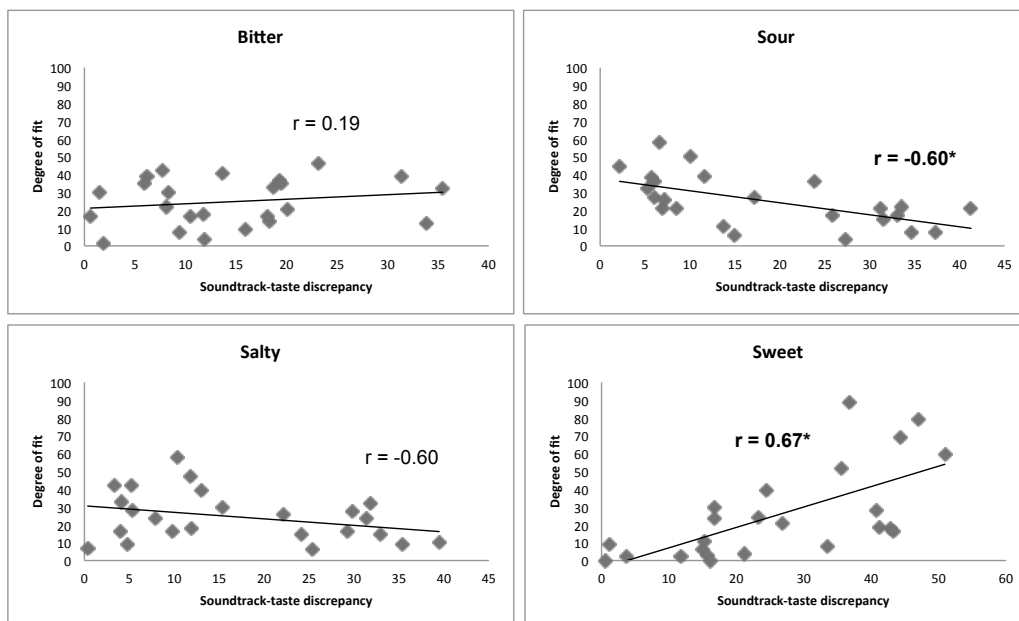


Figure 2.15. A measure of the degree of pleasantness (A) and arousal (B) mediation in soundtrack-taste matching, shown as the relationship between soundtrack-taste discrepancy from Experiment 2B (in terms of differences in pleasantness/arousal ratings of a given soundtrack-taste pair) and the degree of fit between each soundtrack and a given taste (represented by the % of responses to the soundtrack-taste matching question for a given taste from Experiment 2A) The Pearson correlation coefficient is shown on each plot along with a linear regression trendline. Asterisks after the correlation coefficient indicate statistical significance ($p < .05$).

2.6 General discussion and conclusions

In this chapter, the role of emotion in mediating crossmodal correspondences between sound and basic tastes was examined in two ways. Experiment 1 focused in depth on the relationship between basic tastes and auditory pitch, while Experiment 2 examined more generally the association between basic tastes and soundtracks especially designed to be congruent to these tastes.

The idea that pitch is associated with emotional valence as well as perceived taste intensity sheds new light on possible hypotheses behind the crossmodal matching of pitch and taste. First, it would seem that, as in colour-music and colour-odour correspondences, hedonic matching plays a role in taste-pitch correspondences. However, hedonic matching does not explain the overall positive association between perceived taste intensity and pitch, since higher taste intensity is generally associated with lower valence, but higher pitch is associated with higher valence. One other theory is a statistical learning account based on the innate orofacial gestures that infants make in response to sweet (outwards and upwards tongue positions) versus bitter tastes (outwards and downwards tongue positions) and the associated utterances (Knöferle & Spence, 2012; Scherer, 1986). With respect to intensity, one can imagine that foods with more intense flavours would lead people to make more energetic sounds, which are higher in frequency (Scherer, 1986).

Yet another theory behind pitch-taste matching is metaphorical or semantic mapping, where people may use the same language (e.g., sharp, delicate, heavy) in order to describe both pitch and taste. This theory might encompass both the valence mediation results and the intensity-pitch correlation observed thus far. For instance, sweet tastes and high pitch might both be described as delicate, with positive connotations, whereas bitter tastes and low pitch might both be described as heavy,

with negative connotations. With regard to the correlation between taste intensity and pitch, a possible semantic explanation is that both high intensity and high pitch might be described as “sharp”. An interesting follow-up study from Experiments 1A and 1B would be to repeat the pitch-matching exercise with foods that have easily identifiable basic tastes (bitter/salty/sour/sweet/umami) but differing textures. For instance, imagine two sets of taste solutions of the same concentration, one mixed with water and the other thickened with starch. If the metaphorical matching theory holds, then one would expect the tastant with a heavier body to be matched to a lower pitch than the tastants with a lighter body.

Moving beyond auditory pitch to soundtracks composed to match specific tastes, Experiment 2 showed that participants perform above chance in matching soundtracks to their intended tastes, and that pleasantness and arousal partially mediates this sound-taste association. So then, to summarise, what are the factors governing people’s associations of these soundtracks with basic tastes (taste words)? First and foremost, the studies where many of the soundtracks originated have already demonstrated associations between auditory parameters and sounds (Bronner et al., 2012; Crisinel & Spence 2009, 2010b; Mesz et al., 2011; see Knoeferle & Spence, 2012, for a review). For instance, bitterness appears to be associated with low pitched sounds (Crisinel & Spence, 2009; Knoeferle et al., 2015; Mesz et al., 2011); interestingly, Condiment Junkie’s bitter soundtrack, the most effective bitter soundtrack of all those tested, also had the lowest pitch of all the bitter soundtracks (in fact, it had the lowest pitch of all 24 of the soundtracks). It has also been shown that the sound of the piano is most closely matched with sweetness (Crisinel & Spence, 2010b); the Deng sweet soundtrack (which was matched with the word sweet by 89% of participants) makes liberal use of consonant ‘tinkling’ piano tones.

In addition, the role of emotional mediation is an important one. As shown by the results of Experiment 2B, pleasantness partly mediates soundtrack-taste matches for sweetness and bitterness, while arousal partly mediates soundtrack-taste matches for sourness. Whether intentional or not, the soundtrack designers seem to have captured some emotional aspects of tastes in their soundtracks. For instance, sweet soundtracks tend to be pleasant while bitter soundtracks tend to be unpleasant. As for sour soundtracks, they tend to be the most arousing/exciting of the group.

Lastly, some of the soundtracks – such as those by Condiment Junkie and Deng – also used semantically meaningful associations in the form of the sound of a salt shaker in the salty soundtrack (in fact, the prominent salt shaker sound in the background of the Deng soundtrack might have been the reason why it had the highest matching rate - 58% - out of all the salty soundtracks). In light of the emotion mediation results from Experiment 2B, it is interesting to note that saltiness is the only taste whose crossmodal correspondence with the soundtracks does not appear to be mediated by pleasantness or arousal. Perhaps this is why the use of straight-forward semantic associations was so effective for the salty soundtracks. On the other hand, relying on semantic associations obscures crossmodal correspondences that are yet to be discovered between sounds and saltiness. In the future, it may be useful for researchers and designers to focus their energy on a creating a salty soundtrack which does not involve salt shaker sounds!

Incidentally, it is interesting to note that for many sour soundtracks, bitter and sour constituted the majority of the responses (for instance Mesz-tango had 58 responses for sour and 33 responses for bitter, Deng's sour soundtrack had 50 responses for sour and 32 responses for bitter. See also Figures 2.9 and 2.10). Relevant here are previous findings that people tend to confuse “sour” and “bitter” more than the other tastes

(Blakeslee & Fox, 1932; Meiselman & Dzendolet, 1967; O'Mahony et al., 1979). So, for instance, O'Mahony et al. found that when participants were asked to apply taste adjectives (sweet, sour, salty, and bitter) to actual taste solutions, the most common error was calling the citric acid solution "bitter" instead of "sour". In the context of Experiment 2, the sour soundtracks might have evoked the idea or sensation of sourness, but for those participants who associate the feeling of sourness with the word "bitter", they might have chosen to match the soundtracks with bitterness instead of sourness. A follow-up experiment in which the participants are asked to match soundtracks with unlabelled taste solutions as opposed to taste words would verify whether the sour-bitter confusion contributed to many sour soundtracks being labelled as bitter.

Lastly, Experiment 2 highlighted an example of the role that musical expertise can play in matching music with tastes. For Mesz's sweet Makea piece, those with no musical experience were significantly more likely to match it to sweetness than those with musical experience (for whom, bitterness was the most common choice). This was potentially because Makea has dissonant chords and high pitched piano instrumentation; as musical novices tend to focus on timbre while those with musical experience tend to focus on melody and harmony (Wolpert 1990), perhaps novices matched the high pitched piano sounds to sweetness, while more experienced listeners matched the dissonant chords to bitterness. This is an interesting example of an instance where musical experience hindered, rather than helped, participants' performance on a crossmodal matching task; Previously, Walker (1987) found that musical training was the biggest factor in correctly answering a sound and visual metaphor task. More recently, Wolter and her colleagues (2016) found that only pianists exhibited a compatibility effect between implied pitch height and horizontal

space (left-right), and deduced therefore that the experience of playing the piano led to such a learned association. These examples highlight the role of individual differences in crossmodal correspondences (see Chapter 9 for a more in-depth discussion) and highlight the importance for future compositions to have consistent musical features at all levels of music cognition.

Experiment 2 represents the first step in creating such sonic seasoning, the aim being to find soundtracks that people can reliably match to tastes and which can alter the taste of a dish. From the soundtracks studied here, the earlier versions of bitter and sweet soundtracks by Condiment Junkie have been shown to significantly change people's ratings of toffee on a bitter-sweet scale (Crisinel et al., 2012). In addition, the bitter and sweet soundtracks generated by Reinoso Carvalho were used in a chocolate tasting study in which the sweet soundtrack significantly changed people's rating of bitter chocolate on a bitter-sweet scale (Reinoso Carvalho et al., 2015a, b). In fact, many experiments in this thesis uses the most effective soundtracks found in Experiment 2, such as Deng's sweet soundtrack (Experiments 6, 8, 9, 10, 13, and 14), Condiment Junkie's bitter soundtrack (Experiments 6, 8, 9, and 14) or Mesz' sour tango soundtrack (Experiments 12 and 13). Building on the relationships between sound and tastes demonstrated in this chapter, I will examine any perceptual effects crossmodally congruent soundtracks may have over the taste of real foodstuffs in the next chapter.

3 Examples of the auditory modulation of taste/flavour

3.1 Introduction

Crisinel et al. (2012) first demonstrated that, over-and-above any crossmodal associations (or correspondences) between sounds and taste words (and tastes), auditory stimuli could have an effect on taste/flavour evaluation. In their study, participants were given samples of bittersweet toffee to evaluate while listening to one of two specially-created soundscapes. Toffee sampled while listening to the higher-pitched soundscape was rated as tasting significantly sweeter than while listening to the lower-pitched soundscape. Since then, large-scale experimental studies conducted with both whisky (Velasco et al., 2013a) and wine (North, 2012; Spence et al., 2013, 2014a) have demonstrated that what people hear in the background, namely specifically chosen/composed music (or soundscapes), can significantly alter their perception (or at the very least their ratings) of a variety of taste/flavour characteristics, such as the woodiness of a whisky and the freshness/fruitiness of wine.

In this chapter, I will describe several studies showcasing how auditory stimuli influence food/beverage evaluation in a variety of ways. Auditory stimuli range from melodies with different harmonic consonance (Experiment 3), through to live classical music (Experiment 4), and the manipulation of auditory articulation (Experiment 5). The targeted food/beverage characteristics range from the basic tastes of sweetness/sourness (Experiment 3), through to wine flavour attributes (Experiment 4), and the mouth-feel of creaminess (Experiment 5).

Experiment 3 focuses on the role of musical harmony – an important attribute of music that has both psychophysical and neurological underpinnings – on sweet/sour evaluation. Harmony refers to the combination of two or more frequencies occurring simultaneously (Zentner & Kagan, 1998). This combination can be classified as either consonant or dissonant depending on the feeling that it invokes in the listener – while consonance suggests a feeling of stability and repose, dissonance often suggests a feeling of tension (Turek, 1976). Western music has concrete rules for which intervals – that is, which differences between the frequencies of two simultaneous notes – are consonant or dissonant. For example, semitone intervals in the count of 12 (octave), 7 (fifth), 4 (major third), and 3 (minor third) are considered consonant. By contrast, intervals of 1 (minor second) or 5 (augmented fourth) semitones are considered dissonant (Turek, 1976). Harmony, at least in terms of consonance and dissonance, is undoubtedly a salient musical attribute, and one that even human infants are sensitive to (Zentner & Kagan, 1998).

From an acoustics point of view, consonance and dissonance can be explained by the harmonic relationship between frequency components in a sound (Helmholtz, 1954). When frequency components are close, but not identical, to each other, their interaction produces fluctuations in the amplitude of the overall waveform that are perceived as beats (below 20Hz) or roughness (from 20-250Hz; Fishman et al., 2001). Beats and roughness are perceived as unpleasant and are attributed to the perceived dissonance of musical intervals and chords (Helmholtz, 1954). On the other hand, those chords that are consonant have frequency spectra that are harmonically related (they are all multiples of a common fundamental frequency; McDermott et al., 2010). Neurologically-speaking, the processing of consonance and dissonance may be deeply ingrained in low-level sensory processing. Indeed, recordings from brainstem

responses to consonant and dissonant musical intervals have shown that the strength of responses generally followed the traditional western music ordering of consonance, with more pronounced responses observed for more consonant musical relationships (Bidelman & Krishnan, 2009). In addition, neural correlates of dissonance have been found in the primary auditory cortex of both monkeys and humans (Fishman et al., 2001). Brain scans of participants listening to music having various degrees of consonance and dissonance have highlighted increased activity in the paralimbic and neocortical regions responsible for processing pleasant/unpleasant emotional states (Blood et al., 1999).

Hedonic matching is one of the mechanisms that have been proposed to underpin at least certain crossmodal correspondences (e.g., Knöferle & Spence, 2012; Palmer et al., 2013, Schifferstein & Tanudjaja, 2004; Velasco et al., 2015). The idea here is that people associate disparate stimuli from different sensory modalities with each other based on their pleasantness, even if they have never experienced the stimuli together before. For instance, it may be natural to expect that consonant musical intervals, which are more pleasant to the ear, would be matched with more pleasant tastes, whereas less pleasant dissonant intervals will be matched with less pleasant tastes instead.

In Experiment 3, I chose to test sourness as the basic taste that matches with, and may be affected by, musical dissonance. In terms of crossmodal correspondences, Bronner et al. (2012) reported that highly consonant music was consistently mapped to the flavour of vanilla (which is generally accompanied by a sweet taste in the western diet; e.g., see Stevenson et al., 2004), whereas low consonance music was mapped on to citrus flavours (which are generally accompanied by sour tastes) instead. In addition, by analysing music that had been composed on the basis of the four basic

taste words (e.g., sweet, sour, bitter, and salty), Mesz et al. (2011) were able to show that chord consonance was associated with sweetness whereas chord dissonance was associated with sourness.

In Experiment 3, consonant and dissonant musical stimuli were used in order to study the effect of manipulating the harmony of short musical selections on the evaluation of taste. The musical stimuli were created with different melodies and instrumentations in order to ensure that the experiment was focused on the high level concept of harmony, rather than on the specificities of a particular melody or instrument (see Crisinel & Spence, 2010b). The participants in Experiment 3A tasted samples of fruit juice while listening to four musical stimuli consisting of consonant and dissonant versions of two different melodies. The participants in Experiment 3B had to listen to four musical stimuli consisting of consonant and dissonant versions of the same melody in different instrumentation (piano and trumpet). In both Experiment 3A and 3B, the participants had to rate the taste of the juice as well as rate the pleasantness of the music and the juice, after listening to each short piece of music.

Experiment 4 tackled the problem of music-wine congruency and its potential downstream effects on wine evaluation. For instance, Spence et al. (2013) looked more specifically at the problem of matching music to a selection of fine wines. In the initial stages of their study, these researchers assessed how well their participants considered that a set of eight pieces of classical music matched a predetermined set of four wines (chosen to display quite different characteristics), and extracted one particularly well-matched piece of music for each wine. Next, they compared the ratings of those wines that were tasted while listening to the aforementioned matching music versus the ratings of the same wine when tasted in silence. Overall, the

participants rated the wines as tasting significantly sweeter, and more liked, when tasted while listening to the putatively congruent music.

However, as yet, researchers have not examined whether there are any overarching effects on sensory ratings with congruent or incongruent music that go beyond the influence of a single piece of music on a single glass (that is, type) of wine. For instance, according to the theory of processing fluency (see Labroo et al., 2008; Winkielman et al., 2003), one might predict that congruent music and food would lead to increased ratings of food pleasantness; When the music and food are congruent, participants can perhaps more easily evaluate the sensory properties of the food, and consequently may find it more pleasant. By the same token, eating food while listening to incongruent music should lead to reduced pleasantness ratings.

To explicitly assess the role of music congruency on wine ratings, Experiment 4 was designed such that half of the participants listened to putatively more congruent music while tasting two different wines, while the other participants listened to music that was putatively less congruent. It is worth noting that in the field of sound-odour correspondences research, the studies that have been published to date have demonstrated that odours tend to be rated as more pleasant while people are listening to congruent sounds than while listening to incongruent sounds (e.g., Seo & Hummel, 2011; Seo et al., 2014). Aroma-sound congruency in Seo and Hummel's study involved the aroma of a food item (either coffee or potato chips) and the sound associated with consuming the same food (someone sipping a cup of coffee, or munching on potato chips). Meanwhile, Seo et al. used the aromas of food items and musical excerpts commonly associated with those sounds (cinnamon aroma with Christmas carols or coffee aromas with a coffee advertising jingle). Admittedly, these examples of congruence and incongruence are much more literal (and not crossmodal

per se, as discussed in Section 1.1), and commonly encountered in everyday life, than the more abstract pairings of music with wine.

Furthermore, the participants in Experiment 4 rated a number of attributes commonly assessed in wine (Fielden, 2009), including length (of flavour remaining in the mouth after swallowing), richness, and complexity. As music is temporal in nature, it was particularly interested to test whether the tempo of the music would affect the perceived length of flavour remaining in the mouth after swallowing.

Finally, Experiment 5 went beyond examining perceptual effects of music on flavour. Moving forward, there is now a growing interest in determining whether sound can also influence people's perception of tactile-based attributes as well (Spence, 2015c). For instance, can the presentation of appropriate sounds (that are not necessarily related to eating/drinking) make food/drinks appear more/less crispy, crunchy, creamy, and/or carbonated?

In Experiment 5, it was hypothesized that specific soundtracks might affect the perceived texture of chocolate, in particular its creaminess. Here, it is important to mention that previous similar research has assessed the various different ways in which the perceived texture of food can be associated – and potentially altered – by the different combinations of sensory stimuli. For instance, round shapes tend to be associated with creaminess (Yorkston & Menon, 2004). Furthermore, differences in the texture of a food's surface can apparently also alter its perceived sourness (Slocombe et al., 2016). Previous research has also demonstrated that sweeter chocolates are usually associated with rounder shapes, whereas more bitter chocolates are more commonly matched with angular shapes instead (Gallace et al., 2011; Ngo et al., 2011; also see Bremner et al., 2013, and Spence & Deroy, 2012, for overviews).

In Experiment 5, the participants tasted and rated the same chocolate twice (without knowing that the chocolates were identical), each time under the influence of one of two soundtracks. The soundtracks were produced to evoke either creaminess or roughness (in this case, roughness has being defined as the opposite of creaminess). First, the bouba-kiki effect (also known as the “maluma-takete” effect) was taken into consideration as a starting point for soundtrack production. People tend to associate round/smooth visual/auditory cues with “bouba”-like words, whereas sharp/rough stimuli may be naturally associated with more “kiki”-like words (Köhler, 1929, and 1947; Bremner et al., 2013). With this in mind, one might associate purer waveforms with smoothness (bouba/maluma) and more complex waveforms with roughness (kiki/takete). Eitan and Rothschild (2010) also provided some potential musical guidance here. These researchers addressed how musical parameters, such as pitch, loudness, timbre, may affect auditory-tactile metaphorical mappings. They found, for example, that a flute’s simpler sound wave was rated as smoother than the more complex sound of a violin. These research findings guided the production of the soundtracks used in Experiment 5.

3.2 Experiment 3: Assessing the influence of consonant and dissonant music on taste perception

3.2.1 Methods

3.2.1.1 Participants

The participants were recruited at the Food Matters Live conference at London ExCeL, November 18-20th, 2014. The experiment was approved by the Central University Research Ethics Committee of Oxford University. The participants gave their informed consent and reported no impairment of their senses of hearing and

taste. 39 participants (28 females, aged 19-65 years, $M=34.15$, $SD=12.50$) took part in Experiment 3A. 32 participants (20 females, aged 18-60 years, $M=32.09$, $SD=11.57$) took part in Experiment 3B. The experiment lasted for approximately 4 minutes.

3.2.1.2 *Auditory stimuli*

Consonant and dissonant versions of two short melodies were created (melody B was just an inversion, in time, of melody A, in order to ensure that the melodies did not differ in terms of their pitch range). The melodies and harmonisations were based on stimuli used in a study on infant perception of consonance and dissonance in music reported by Zentner and Kagan (1998). The consonant version was harmonised with major and minor thirds (three and four half-steps, respectively), while the dissonant version was harmonised with minor second intervals (one half-step). The music scores (see Figure 3.1) were created using the MuseScore software and exported to midi form. The musical stimuli for Experiment 3A were produced with GarageBand's Steinway Grand Piano plugin. In Experiment 3B, consonant and dissonant versions of melody A were produced with Garage Band's Trumpet plugin as well as the aforementioned piano plugin. The sounds can be downloaded from <https://soundcloud.com/janicewang09/sets/harmony>.



Figure 3.1. Melody A harmonised with consonant intervals (top) and dissonant intervals (bottom), used in Experiment 3.

A control experiment was conducted in order to check that the consonant and dissonant musical pieces would indeed be associated with sweet and sour sounds, respectively. 30 participants (18 female, aged 18-52 years, $M=30.07$, $SD=7.92$) had to rate the soundtracks on a 10-point sour-sweet scale (where 1=very sour and 10=very sweet) in an online test. The sour-sweet ratings of the consonant ($M=6.78$, $SD=1.54$) and dissonant soundtracks ($M=3.27$, $SD=1.34$) differed significantly in the expected direction, $t(59)=12.18$, $p<.001$.

3.2.1.3 Gustatory stimuli

A juice mixture was prepared for use in the two experiments. In Experiment 3A, a blend of 2:2:1 ratio Tesco brand orange juice, grapefruit juice, and apple juice was used. Meanwhile, a blend of 1:1 Welch's white grape, apple, and pear juice and Welch's white grape and peach juice was used in Experiment 3B. Juice mixtures were used instead of a single juice in order to minimize the possibility that the participants would guess the nature of the juice sample and hence perhaps feel overconfident giving their responses, especially given the within-participants nature of the experimental design. The juice mixture was switched in the second experiment because many of the participants in Experiment 3A had assumed that the samples were all just orange juice based on the colour. The juice mixture used in Experiment 3B had a clear, pale yellow colour to discourage participants from jumping to any conclusions about the nature of the juice. All of the juices were pulp-free, but there was no independent assessment regarding the texture of the juice mixtures between Experiments 3A and 3B.

3.2.1.4 Design and procedure

Both of the experiments were hosted on the LimeSurvey online survey platform. The participants were seated in front of a computer and were instructed to put on a pair of headphones. They were given a cup of water to rinse their mouths before the start of the experiment and between trials. Four cups were placed in front of the computer, labelled 1-4. After consenting to take part in the study, the participants were presented with the four auditory stimuli, one after the other. With each sound, the participants were instructed to start drinking once the sound file had started playing. After drinking, the participants rated: 1) how much they liked the music; 2) how much they liked the drink and; 3) how they rated the taste of the drink on a sour-sweet scale. All three of the response scales had 10 points. The participants were instructed to rinse their mouths out with water between each trial. The order of presentation of the four pieces of music and the order of each set of questions were randomised for each participant. Experiments 3A and 3B had the same protocol and only differed in terms of the auditory and gustatory stimuli that were presented.

3.2.2 Results and discussion

Experiment 3A: A two-way repeated measures analysis of variance (ANOVA) was performed on the data using harmony type (consonant or dissonant) and melody type as factors. Analysis of the data revealed a significant main effect of harmony type on sour-sweet ratings ($F(1,38)=8.939$, $p=.005$, $\eta^2=.190$). Specifically, the participants rated the fruit juice as tasting sourer while listening to the dissonant music ($M=4.09$, $SD=1.63$) than while listening to the consonant music ($M=4.87$, $SD=1.96$). There was no significant effect of melody type and no interaction between harmony and melody type. Post-hoc pairwise comparisons with Bonferroni correction revealed that

participants' sour-sweet ratings for consonant melody A was significantly different from dissonant melody B ($p=.021$).

Harmony type also affected both ratings of music pleasantness ($F(1,38)=58.227$, $p<.001$, $\eta^2=.605$) and of flavour pleasantness ($F(1,38)=22.608$, $p<.001$, $\eta^2=.373$). The participants found the consonant music to be more pleasant and liked the juice samples significantly more when it was played. Post-hoc analysis revealed that each consonant melody was significantly different from each dissonant melody for both music and flavour pleasantness ratings ($p<.001$).

The correlations between music pleasantness ratings and sour-sweet ratings ($r_{156}=.289$, $p<.01$), between sour-sweet ratings and flavour pleasantness ratings ($r_{156}=.734$, $p<.01$), and between music pleasantness ratings and flavour pleasantness ratings ($r_{156}=.451$, $p<.01$) were also significant.

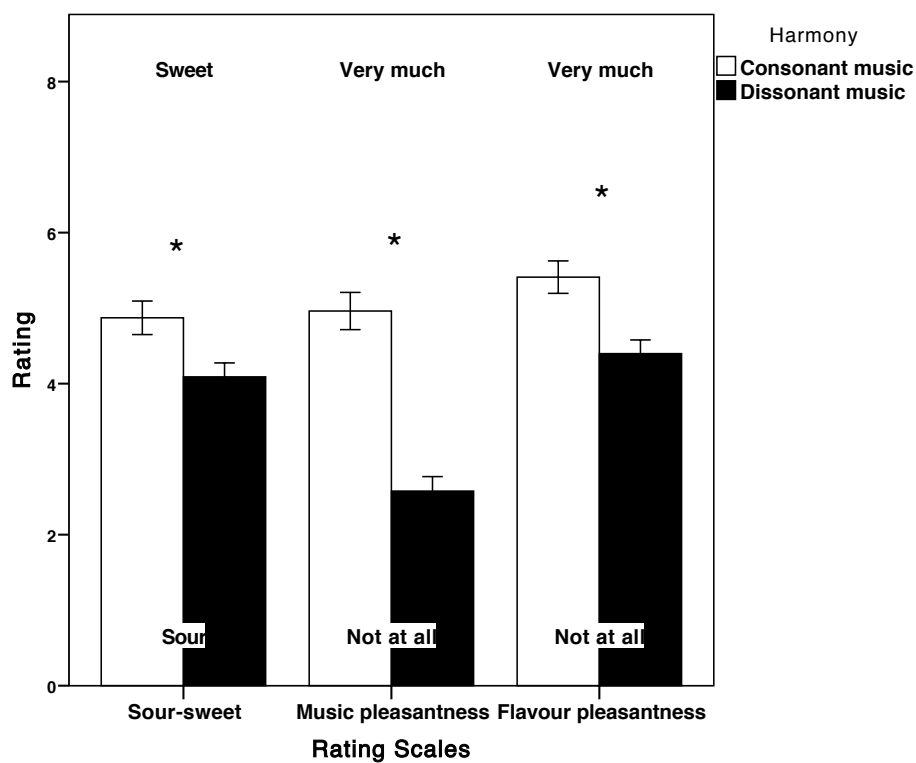


Figure 3.2. Mean ratings of sour-sweet, music pleasantness, and flavour pleasantness in Experiment 3A, for the consonant and dissonant music conditions. The error bars represent the standard error of the means. Asterisks mark significant differences ($p < .05$) between the consonant and dissonant music conditions.

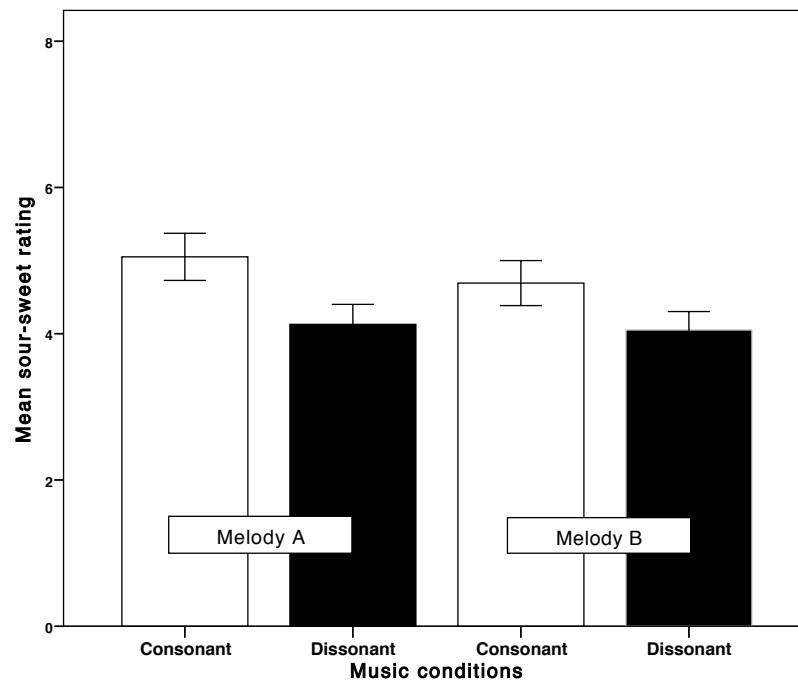


Figure 3.3. Mean ratings on the sour-sweet scale (1=very sour, 10=very sweet) for all of the conditions in Experiment 3A. The error bars represent the standard error of the means. Consonant melody A and dissonant melody B were significantly different according to post-hoc pairwise comparisons with Bonferroni correction.

Experiment 3B: A two-way repeated measures ANOVA was performed on the data using harmony type (consonant or dissonant) and instrument type as factors. The analysis revealed a significant main effect of harmony type on sour-sweet ratings ($F(1,31)=11.536$, $p=.002$, $\eta^2=.271$). The participants rated the fruit juice as tasting

significantly sweeter while listening to consonant music stimuli ($M=6.25$, $SD=1.94$) than while listening to the dissonant music ($M=5.40$, $SD=2.04$). There was no significant effect of instrument type, nor any interaction between harmony type and instrument type. Post-hoc pairwise comparisons with Bonferroni corrections revealed that the sour-sweet ratings for the dissonant piano music was significantly different from the consonant trumpet ($p=.010$).

Harmony type ($F(1,31)=62.640$, $p<.001$, $\eta^2=.669$) had a significant effect on music pleasantness. Once again, the participants liked the consonant musical stimuli more than the dissonant ones. Instrument type also exerted a significant effect on music pleasantness ($F(1,31)=4.529$, $p=.041$, $\eta^2=.127$), with the participants liking the piano stimuli more than the trumpet stimuli. (Note that a similar result was also observed in Crisinel et al.'s, 2010b, study, where participants found synthesised piano sounds more pleasant than brass). Post-hoc analysis revealed each consonant melody was significantly rated as more musically pleasant than each dissonant melody regardless of instrument type.

Participants' ratings of flavour pleasantness were also affected by harmony type, $F(1,31)=10.183$, $p=.003$, $\eta^2=.247$. The participants liked the fruit juice more while listening to the consonant musical stimuli than while listening to the dissonant stimuli. Post-hoc analysis revealed that flavour pleasantness under the consonant piano stimuli condition was significantly higher than under the dissonant trumpet stimuli ($p=.020$). Furthermore, flavour pleasantness was significantly higher while listening to the consonant than while listening to the dissonant trumpet stimuli ($p=.005$).

The correlations between ratings of music pleasantness and ratings on the sour-sweet scale ($r_{128}=.358$, $p<.01$), between sour-sweet ratings and flavour pleasantness ratings

($r_{128}=.457$, $p<.01$), and between music pleasantness ratings and flavour pleasantness ratings ($r_{128}=.506$, $p<.01$) were all significant.

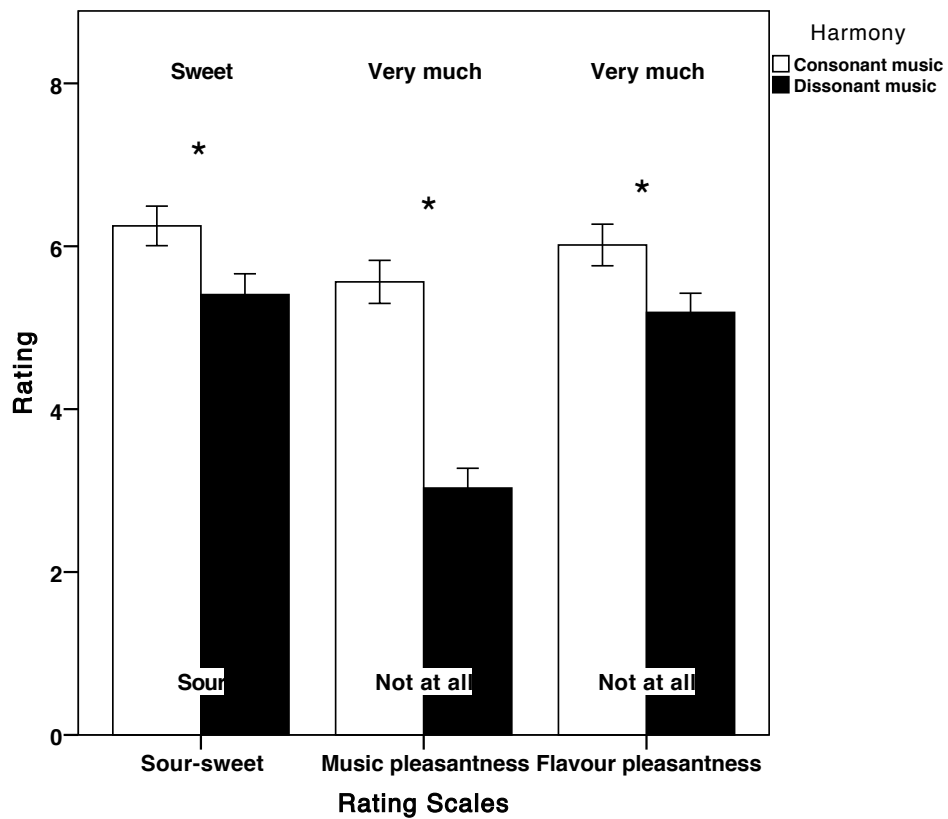


Figure 3.4. Mean ratings of sour-sweet, music pleasantness, and flavour pleasantness in Experiment 3B, for both consonant and dissonant music conditions. The error bars represent the standard error of the means. Asterisks indicate significant differences ($p<.05$) between the consonant and dissonant music conditions.

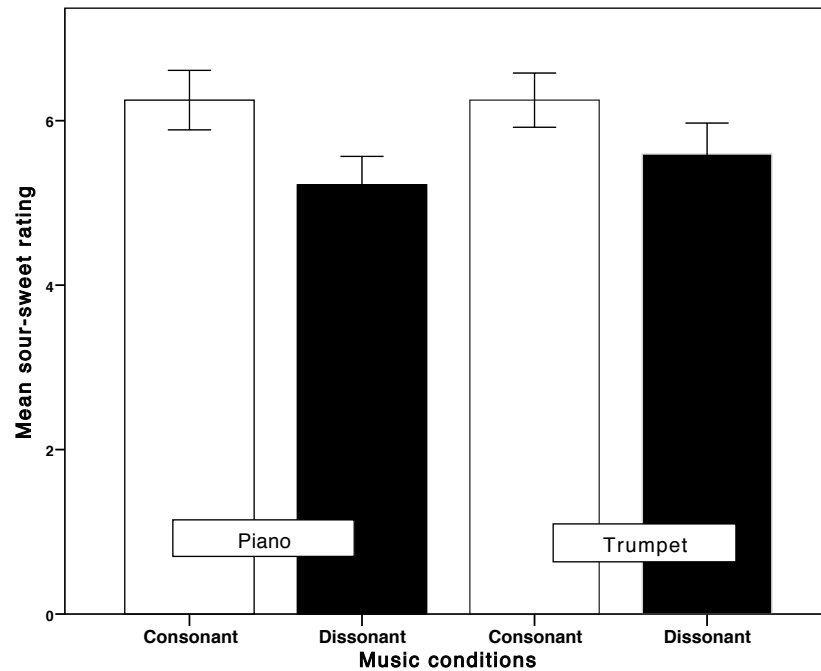


Figure 3.5. Mean ratings on the sour-sweet scale (1=very sour, 10=very sweet) for all conditions in Experiment 3B. Error bars represent the standard error of the means. Consonant trumpet and dissonant piano were significantly different in post-hoc pairwise comparisons with Bonferroni correction.

The results of Experiment 3 therefore demonstrate that music of different harmonic content is associated with different basic tastes; in the pre-test, the participants associated consonance with sweetness and dissonance with sourness. Furthermore, the results also demonstrate that music of different harmonic content can alter taste and pleasantness ratings for fruit juice. More specifically, a mixed fruit juice sample was rated as tasting significantly sweeter when the participants listened to consonant music while its sourness was rated higher while listening to dissonant music, independent of melody or instrumentation.

These results build on previous research showing that background sounds can induce perceptual sweet/bitter taste differences by manipulating its pitch and timbre (Crisinel et al., 2012), or that people will reliably mix juices with significantly different amounts of sugar/acid content to match different pieces of instrumental music (Kontukoski et al., 2015). Experiment 3 also demonstrates that differences in taste ratings can be induced by music independent of any changes in pitch or timbre, since all the music stimuli used were in the same pitch range, and the timbre of the instrument did not have a significant effect on participants' taste ratings. As in Crisinel et al.'s (2012) study, the magnitude of the change in participants' taste ratings was in the region of 10%.

As predicted, the participants rated the fruit juice tasted while listening to the consonant music selections as being more sweet/less sour than the same juice tasted while listening to the dissonant music. The fact that the participants associated musical consonance with greater sweetness, both of which are usually considered pleasant (see Blood et al., 1999, for neurological evidence that consonant/dissonant music elicits pleasant/unpleasant emotional states), is in line with the hedonic matching hypothesis. According to this account, crossmodal correspondences may be formed because two stimuli are similar in terms of their pleasantness (Knöferle & Spence, 2012; Palmer et al. 2013; Velasco et al., 2015). It is, however, important to note that one other possible origin for the consonance-taste correspondence may be mediated by a link with oral-somatosensation. For instance, the auditory sensation of roughness associated with dissonant chords may be matched with the 'sharp' mouthfeel of sour tastes; conversely, the smoothness of consonant musical chords may be associated with the 'smooth' mouthfeel of sweet foods (see Obrist et al., 2014; see also Knöferle et al., 2015, where participants matched sweet taste words

with low auditory roughness and sour taste words with significantly higher auditory roughness).

Another key question is the putative mechanism by which this correspondence might have influenced participants' responses. One suggestion is that the correspondence might have exerted a direct influence on participants' taste perception. In line with the oral-somatosensory origin of the crossmodal correspondence proposed in the introduction (Section 3.1), the auditory roughness of dissonant music might have emphasised the puckering sensation of sourness in the mouth just as the smoothness of consonant music might have brought out the sweetness in the juice. Another possibility is that the high-pitched and dissonant music might have somehow triggered greater saliva flow in the mouth. Since eating (or even looking at) sour foods can make the mouth water (Hagenmuller et al., 2014; Pangborn, 1968; Pangborn et al., 1979; see Spence, 2011b, for a review), people might have misattributed their experience of increased saliva flow (from listening to the dissonant music) to the sourness of the juice samples. This physiological influence theory is explored in more depth in Chapter 8.

Alternatively, however, given that expectations are known to influence people's sensory and hedonic evaluations (Deliza & MacFie, 1996; see Piqueras-Fiszman & Spence, 2015, for a review), it might be suggested that the music stimuli set up some kind of expectations, conscious or otherwise, in the minds of the participants about the drinks that they were about to taste. This is especially relevant since participants would have heard the music for a few seconds before they started to taste each sample. For instance, the pleasurable consonant music could have resulted in a halo effect, whereby the participants would have expected the juice to taste more pleasant (i.e., sweeter). Alternatively, the consonant music could have primed participants to

expect a sweet taste before they consumed the juice. This expectations-based effect could have manifested itself either as a change in sensory evaluations or as a change in self-report. The expectations theory will be explored in Chapter 5.

A third interpretation comes from a consideration of the transfer of value. It is well-known that music can prime certain emotions in the mind of the listener and that these may then be transferred to other sensory modalities. For instance, it has been shown that music-elicited emotion can influence people's performance in those tasks that involve the processing of visual emotional stimuli (Logeswaran & Bhattacharya, 2009). In this case, perhaps any positive feelings induced by consonant music (and equally, any negative emotions induced by the dissonant music) may simply have been transferred to the juice rating task, thus resulting in higher reported pleasantness (and sweetness) ratings for the juice. Some evidence for this was found in both the significant correlations between music pleasantness and flavour pleasantness, and between music pleasantness and sour-sweet ratings.

Another way of discriminating between these various accounts might be to try and separate the hedonic or tactile origins of the crossmodal correspondence. One might repeat the current study with a pool of participants who don't find sweetness especially pleasant, perhaps those who are 6-*n*-propylthiouracil (PROP) supertasters (Peterson et al., 1999; Yeomans et al., 2007). The hedonic matching account assumes sweetness is matched with consonant music because most people find them both pleasant. For these participants, however, sweetness would not be pleasant. If consonant music still increases their sweetness rating compared to dissonant music, then one should probably look beyond hedonic matching as the origin of this correspondence. Taster status as a potential source of individual differences is further explored in Chapter 9.

What is not clear from the results of Experiment 3 is whether dissonance was truly associated with sourness, or merely with “not sweet”. Following-up with separate rating scales for putatively pleasant tastes (sweet) and unpleasant tastes (sour, bitter) should help to clarify this point. A separate rating for bitterness would be especially useful as a replication of Shankar and Spence’s (2008) bittersweet chocolate study where no significant effect of consonance/dissonance was found sweet/bitter ratings (see Spence et al., 2010).

Here, it is perhaps also worth noting that the participants in Experiment 3 were not asked about their levels of musical expertise. Walker (1987) showed that musical expertise might be the single most important factor in the choice of sound-shape matching for participants of different ages, cultural backgrounds, and musical expertise; so it is possible that different levels of musical expertise could potentially also have contributed to any differences in terms of how the participants associated harmony with tastes. This is especially relevant in the light of research showing that background in classical musical training increases preference for consonant music and decreases preference for dissonant music (e.g., Dellacherie et al., 2011; McDermott et al., 2010). In addition, Experiment 1 revealed a difference in how people matched a particular soundtrack (Mesz’s Makea) to either sweetness or bitterness, depending on their musical expertise.

Finally, this area of research is ripe for cross-cultural studies with populations whose musical traditions differ from the western music model (see Knöferle et al., 2015, for a cross-cultural study on music-taste matching); for those individuals who do not find minor second intervals unpleasant, will dissonance map onto other tastes?

The results of Experiment 3 therefore demonstrate the role of musical harmony in taste association and evaluation. Further research is, however, needed in order to

pinpoint the origin and mechanism of this intriguing crossmodal correspondence. Next, Experiment 4 moves beyond simple manipulations of consonance/dissonance to assess whether classical music can influence the wine tasting experience.

3.3 Experiment 4: Assessing the effect of musical congruency on wine tasting in a live performance setting

3.3.1 Methods

3.3.1.1 Participants

From the 80 participants who attended a ‘Cheesemas’ event held at Somerville College, Oxford University on November 28th, 2014, 64 participants returned their rating sheets to the experimenters after the event. Because the experiment was conducted at a public event, the participants did not sign a standard consent form; however, the purpose of the experiment and procedure was clearly explained to them. The participants were also informed that they did not need to complete the questionnaire should they not want to, and that they could stop responding at any stage. Information concerning age and gender was not collected, although only those over 18 years of age were eligible to attend the event. The experiment was approved by the Central University Research Ethics Committee of Oxford University.

3.3.1.2 Auditory stimuli

Two pieces of classical music were performed by Lucia Brighenti and Irene Ortega Albaladejo from the Royal Academy of Music, London. The first was Debussy’s *Jardin Sous la Pluie*, at roughly 150 beats per minute, a virtuosic piano solo with many fast passages in a high pitch range. This piece was chosen to match the white

wine, since high tempo and pitch have been shown to be associated with a sour taste and citrus flavours (Bronner et al., 2012; Mesz et al., 2011). The second piece of music was Rachmaninoff's *Vocalise*, a piano and cello duet played in a slow tempo (roughly 80 beats per minute). This piece was chosen to match the red wine, since legato articulation and a consonant melody have both been shown in prior research to be matched with sweet tastes and full body (Bronner et al., 2012; Mesz et al., 2011).

3.3.1.3 Wines

Two wines were chosen for the event. The white wine was a Marcel Martin Sauvignon Blanc 2013, from the Loire Valley in France. It has grass, citrus, and gooseberry notes, light body, and high acidity. The red wine was a Para Dos Malbec 2013, from Mendoza, Argentina. It has black fruit, oak, and vanilla notes, medium body, medium acidity, and soft tannins. Note that the wines were selected to be very different from one another, in order to facilitate the musical matching.

3.3.1.4 Design and procedure

Each participant was given a glass of the white and red wine. The musicians performed the Debussy piece first, followed by the Rachmaninoff¹. One group of participants was instructed to taste the white wine while listening to Debussy and the red wine while listening to Rachmaninoff (therefore always tasting the wine with putatively better matching music), the other group tasted the red wine while listening to Debussy and the white wine while listening to Rachmaninoff (therefore always tasting the wine with the music that matched less well). For each song/wine pairing, the participants were instructed to fill out a rating form (see Figure 3.6) with scales

¹ Due to the nature of the public event, it was not possible to counterbalance the order in which the two pieces of music were presented.

ratings for all the scales were converted from the range of -50mm to 50 mm to a score of 0-100. In order to compare the effect of different musical conditions on the same wine, independent samples t-tests were performed on the ratings made by the participants from both groups for the same wine (see Figures 3.7A and B). In addition, mixed measures analyses of variance (ANOVAs) were performed in order to assess any overall effect of the music and music-wine congruence (and interactions thereof) on the ratings of each attribute.

For the white wine, the participants rated the Debussy ($M=53.47$, $SD=21.10$) as a significantly better match than the Rachmaninoff ($M=40.58$, $SD=25.90$), $t(61)=2.18$, $p=.033$. In addition, the white wine was rated as significantly more acidic while listening to Debussy ($M=69.20$, $SD=16.79$) than while listening to Rachmaninoff ($M=47.69$, $SD=17.50$), $t(62)=5.01$, $p<.001$. However, no significant differences were observed in the ratings of fruitiness, richness, complexity, length, or pleasantness.

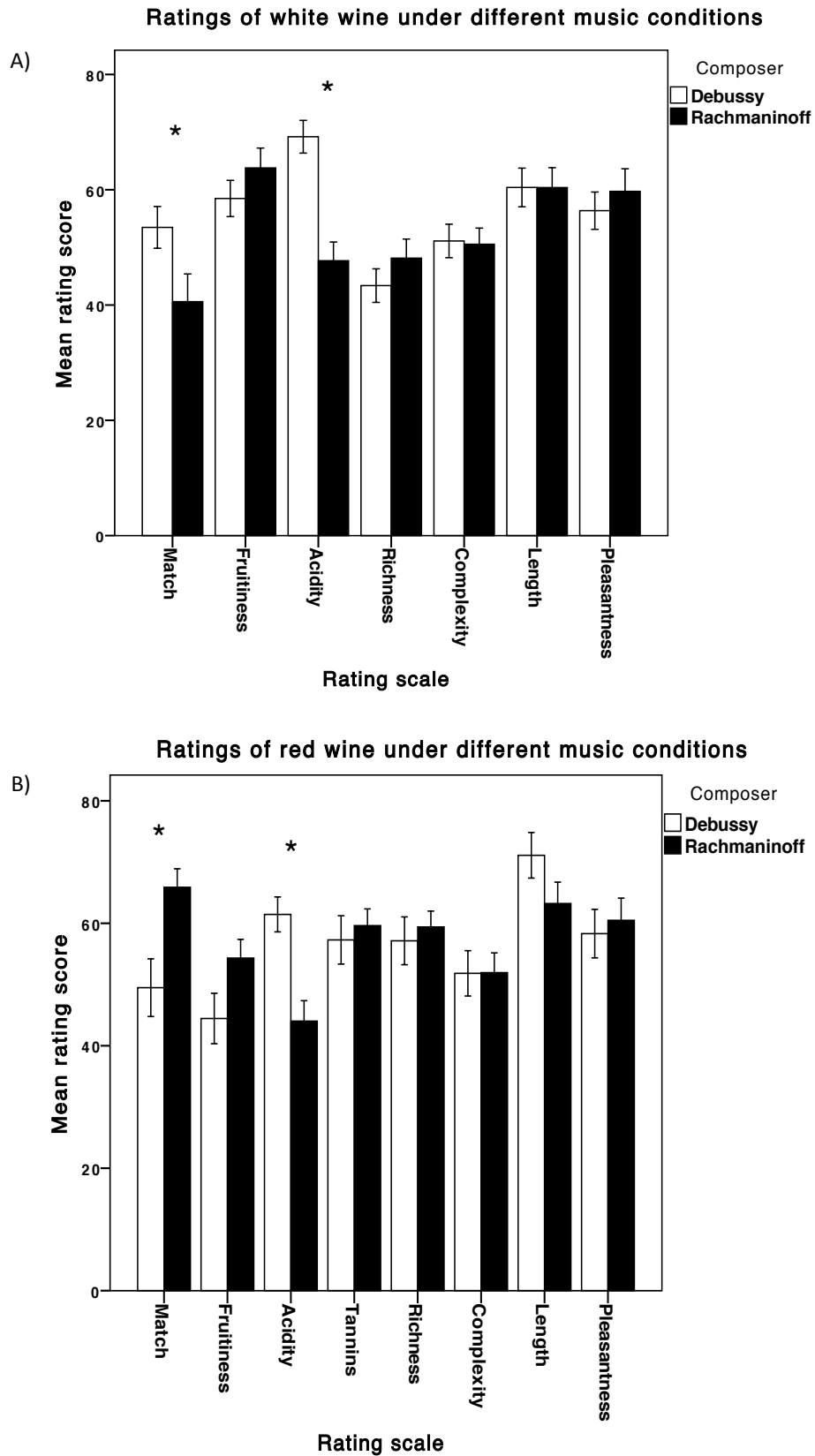


Figure 3.7. Mean rating scores of wine-music match, fruitiness, acidity, richness, complexity, length, and pleasantness for the white wine (A) and red wine (B), under both music conditions in

Experiment 4. Note that each scale was 100 mm long, with 0 corresponding to the midpoint of the scale. Scale ratings for all the scales were converted from the range of -50mm to 50 mm to a score of 0-100, as shown on the y-axis. The error bars represent the standard error of the means. Asterisks mark significant differences ($p < .05$) between music conditions.

As expected, an independent samples t-test revealed that the participants rated the Rachmaninoff ($M=65.88$, $SD=17.60$) as a significantly better match for the red wine than the Debussy ($M=49.48$, $SD=25.36$), $t(48.7)=-2.93$, $p=.005$. Additionally, the red wine was rated as significantly less acidic while listening to Rachmaninoff ($M=44.03$, $SD=19.70$) than while listening to Debussy ($M=61.45$, $SD=22.16$, $t(62)=3.88$, $p < .01$). There were, however, no significant differences in the ratings of fruitiness, tannins, richness, complexity, length, or pleasantness. That said, a borderline significant trend towards a difference in fruitiness is perhaps worth noting, with the red wine being rated as more fruity while listening to Rachmaninoff ($M=54.31$, $SD=18.06$) than while listening to Debussy ($M=44.49$, $SD=22.16$, $t(53.86)=-1.93$, $p=.059$).

Pearson correlations between the different rating attributes were examined (see Table 3.1). A significant positive correlation was observed between pleasantness and wine-music match ($r_{126}=.175$, $p < .05$), suggesting that people associate increasing congruency between wine and music with increasing pleasantness. Significant positive correlations were also observed between pleasantness and fruitiness, richness, complexity, and length; between length and both richness and complexity; between complexity and richness; between richness and acidity. In addition, there was a significant negative correlation between richness and acidity. The positive correlations between length, richness, and complexity can all perhaps be explained by

the fact that these terms are all used as indicators of quality in the world of wine (e.g., see Fielden, 2009).

1.	Match	Fruitiness	Acidity	Richness	Complexity	Length	Pleasantness
Match	1	.162	-.130	.193*	.121	.168	.175*
Fruitiness		1	-.072	.015	.109	.139	.247**
Acidity			1	-.231**	-.150	.007	-.123
Richness				1	.562**	.333**	.460**
Complexity					1	.390**	.428**
Length						1	.433**
Pleasantness							1

Table 3.1. Pearson correlation coefficients between wine ratings for both red and white wines. *

indicates significant correlations at $p < .05$, ** indicate significant correlations at $p < .01$.

Further data analysis was conducted via a two-way mixed measures ANOVA, with music (Debussy vs. Rachmaninoff) as the within-participants factor and congruence (whether the music matched the wine or not) as the between-participants factor, for the attributes of fruitiness, acidity, richness, complexity, length, and pleasantness. (Tannin was not considered since it was only applicable to the red wine.) The main effects of music and wine-music congruence were analysed as well as any interaction effects between music and wine-music congruence.

3.3.2.1 *Fruitiness*

The analysis revealed a significant main effect of music on fruitiness ratings ($F(1,62)=5.684$, $p=.020$, $\eta^2=.084$). Specifically, the wine tasted while listening to Rachmaninoff ($M=59.05$, $SD=2.29$) was rated as significantly more fruity than when listening to Debussy ($M=51.47$, $SD=2.55$), regardless of the type of wine (white or red). In addition, in both music conditions, the white wine was rated as significantly more fruity than the red wine ($p=.008$ while listening to Debussy, $p=.043$ while listening to Rachmaninoff).

3.3.2.2 *Acidity*

There was a significant main effect of music on ratings of acidity ($F(1,62)=40.543$, $p<.001$, $\eta^2=.395$). Specifically, the wine tasted while listening to Debussy ($M=65.32$, $SD=2.03$) was rated as significantly more acidic than the wine tasted while listening to Rachmaninoff ($M=45.86$, $SD=2.35$), regardless of the type of wine (red or white) that was being tasted.

3.3.2.3 *Richness*

There was a significant interaction between music and music-wine congruence, $F(1,62)=18.697$, $p<.001$, $\eta^2=.232$. In both music conditions, the red wine was rated as richer than the white ($p=.005$ while listening to Debussy, $p=.008$ while listening to Rachmaninoff).

3.3.2.4 *Length*

There was a significant interaction between music and music-wine congruence, $F(1,62)=4.075$, $p=.048$, $\eta^2=.062$). Post-hoc analysis with Bonferroni corrections

revealed that during the performance of Debussy, the length of the red wine was rated as significantly longer than the white wine ($p=.036$).

3.3.2.5 *Complexity and pleasantness*

None of the factors were significant for complexity or pleasantness ratings.

Overall, there were no significant differences in the ratings between the group which listened to more congruent music while drinking wine and the group which listened to less congruent music while drinking wine, except for the rating of how much the atmosphere matched the wine.

In general, the results of Experiment 4 clearly demonstrate that the participants rated specific pieces of music as constituting a better match for each wine, and that the music exerted a significant effect on the perceived acidity and fruitiness of the wine. The participants' ratings of wine-music matching justified the musical pre-selections, thus supplying additional evidence in support of Spence et al.'s (2013) findings that social drinkers often concur when it comes to matching wines to specific pieces of music. Just as for the audiovisual correspondences that have been observed between music and colour (Palmer et al. 2013, 2016), one could, perhaps, think of this as a crossmodal correspondence between music and wine rather than participants merely choosing wines that match what they happen to be listening to on a whim.

For each wine, there were significant differences between ratings of fruitiness and acidity under the two music conditions. Both red and white wines tasted while listening to the Rachmaninoff piece were rated as being significantly fruitier than when tasted while listening to the Debussy piece. In addition, ratings of acidity were significantly higher (on average by 20%) for both wines while Debussy was played

than while the participants listened to Rachmaninoff (see Table 3.2 for a summary of the results from all of the music and wine studies published to date).

Study	Music	Effect on wine	Matching wine
North, 2012	<i>Carmina Burana</i> - Orff	Powerful and heavy	
	<i>Waltz of the Flowers</i> (from <i>The Nutcracker</i>)	Subtle and refined	
	<i>Just Can't Get Enough</i> – Nouvelle Vague	Zingy and refreshing	
	<i>Slow Breakdown</i> – Michael Brook	Mellow and soft	
Spence et al., 2013	Mozart's Flute Quartet No. 1 in D major, K 285 – Movement 1	More enjoyable with music than silence	Domaine Didier Dagnueneau, Pouilly Fumé Silex 2010
	Ravel's String Quartet in F major – Movement 1		Domaine Ponsot, Clos de la Roche 2009
	Tchaikovsky's String Quartet No 1 in D major – Movement 2	More enjoyable with music than silence	Château Margaux 2004
Spence et al., 2014a	Mozart's Flute Quartet No. 1 in D major, K 285 – Movement 1	More enjoyable than <i>Suvitunnelma</i>	Tattinger Brut Réserve
	Viljami Nittykoki's <i>Suvitunnelma</i>		Chateau Carsin Cuvée Noire 2010

Experiment 4	Debussy's <i>Jardins Sous la Pluie</i>	Higher acidity than <i>Vocalise</i>	Marcel Martin Sauvignon Blanc 2013
	Rachmaninoff's <i>Vocalise</i>	Higher fruitiness than <i>Jardins Sous la Pluie</i>	Para Dos Malbec 2013

Table 3.2. Summary of findings from music and wine studies published to date.

Individual rating correlations revealed that there was a significant association between music/wine matching and pleasantness. This supports the hypothesis outlined in the introduction (see Section 3.1) that wines would taste more pleasant while listening to music that is congruent with the wine. However, the positive correlation between music/wine congruency and wine pleasantness is a weak one ($r_{126}=.175$, $p<.05$), in comparison to the correlations reported in Seo and Hummel's (2011) study between sound/odour congruency and odour pleasantness ($r_{88}=.41$, $p<.001$). It is worth noting that while the congruent pairings from Seo and Hummel are likely to be commonly encountered in everyday life (for example, coffee aroma and the sound of drinking coffee, or cinnamon aromas and Christmas music), none of the participants are likely to have consciously encountered the same wine/music combinations before. Congruence based on statistical encounters such as those from Seo and Hummel may be more strongly formed, and thus have stronger effects on processing fluency (Labroo et al., 2008; Winkielman et al., 2003), and hence pleasantness ratings.

For both wines, the participants gave higher average pleasantness ratings (although not significantly so) while listening to the Rachmaninoff piece. Perhaps this was because they enjoyed the Rachmaninoff selection more than the Debussy, although this cannot be verified since ratings of music pleasantness were not collected at the

event². Here, it is also worth noting that in Spence et al.'s (2014b) study, where each wine was rated while listening to two pieces of music, one by Mozart and one by Finnish composer Niittykoski, the participants uniformly ranked the wine tasted while listening to Mozart as more pleasant, regardless of how well it matched the specific wine.

In hindsight, another piece of information that might have been helpful to collect from the participants was their wine tasting expertise. For, according to Deliza et al. (1996), those individuals who are less confident in their sensory abilities generally tend to be influenced more by peripheral information surrounding what they are evaluating. As wine can be considered as a notably complex flavour stimulus (e.g., Smith, 2007), one might have expected the participants' tasting expertise to play a role in how much the music modified their perception of the wine. According to the prediction outlined here, those participants who are experienced wine tasters, and hence who are presumably more confident in their ratings of wine, may perhaps be less influenced by the sound (or any other product-extrinsic) stimuli. The notion of wine tasting expertise as a source of individual differences in the degree to which they experience sonic seasoning effects will be explored in Chapter 9.

On a related note, ratings involving more specific wine terminology - such as length, complexity, and richness - did not differ significantly between the two musical conditions. In contrast, acidity and fruitiness, which were significantly different under different musical conditions, are terms that are certainly more approachable, and used on a more day-to-day basis by the social drinker. It is possible that wine experts, who

In a separate online experiment, 14 participants rated the two pieces of music on a 10-point scale in terms of their pleasantness, complexity, and richness. Mean pleasantness ratings for the Rachmaninoff piece was somewhat higher than for the Debussy piece, but this difference was not significant ($M=7.43$, $SD=1.22$, $M=6.79$, $SD=2.19$, $t(13)=0.93$, $p=.370$).

have more experience dealing with wine-specific terms than the casual social drinker, would give more consistent ratings.

Furthermore, the lack of any difference on ratings of the wine's richness may be explained by background distractions during the experiment. First, it is important to note that while tasting dry wines, the perception of richness is based on the alcohol content (Fielden 2009). In two studies looking at the effect of auditory distraction on the perception of alcohol, Stafford et al. (2012, 2013) found that, when the distractor conditions involved music, shadowing (listening to and repeating news stories), and music plus shadowing, it was the latter condition that resulted in impaired discrimination of alcohol strength. Since the participants in Experiment 4 were free to converse during the musical performance, it is possible that setting inadvertently mimicked the music and shadowing condition in Stafford et al.'s study, thus lessening the participants' ability to judge the alcohol content of the wine (and thereby its richness). Furthermore, the live performance nature of the music possibly demanded more attention from the participants than pre-recorded music played back over headphones, so it could have additionally distracted the participants and thus impacted their ability to distinguish alcohol strength.

The Rachmaninoff piece was played at a slower tempo (roughly 150 beats per minute) than the Debussy piece (roughly 80 beats per minute). That said, the music did not have a significant effect on the ratings of the length of the wine. This did not support the hypothesis that longer musical phrases, as a result of slower tempo, might be associated with longer perceived length of the wine³. Neither were results from

³ A small follow-up study (N=23) was conducted with the music condition counterbalanced across participants (though this time the participants listened to pre-recorded versions of the two musical selections). A significant main effect of music on length was observed, $F(1,19)=5.96$, $p=.025$, $\eta^2=.239$. That is, those wines tasted while listening to Rachmaninoff ($M=6.44s$, $SD=0.35$) were rated as significantly longer than when listening to Debussy ($M=4.99s$, $SD=0.55$).

Experiment 4 in line with Droit-Volet et al.'s (2013) finding that music with a faster tempo leads to longer perceived time duration (i.e. a contrast effect). As mentioned above, the length of flavour remaining in the mouth after swallowing (unlike acidity or fruitiness) is a wine-specific term that social drinkers are unlikely to think about on a daily basis, therefore internal variation in length ratings is likely high. I will return to address the question of wine length perception in Experiment 13, using expert participants and a within-participants experimental design.

Nevertheless, significant differences in fruitiness and acidity ratings were observed; specifically, the Debussy piece was associated with high acidity and low fruitiness ratings, whereas the Rachmaninoff piece was associated with low acidity and high fruitiness ratings. How could these surprising associations between music and taste/flavour possibly be explained? Spence and Deroy (2013b) put forward two possible mechanisms regarding how what we hear might influence what we taste, or at least what we report tasting. The first mechanism works by linguistic or conceptual matching, whereby the matching sound and taste share a common descriptor (Deroy et al., 2013; Spence & Deroy, 2013c). For instance, in North's (2012) study, hearing music that was "powerful and heavy" also increased the "powerful and heavy" rating of the wine that the participants happened to be tasting at the time. In contrast to North's results, the different musical conditions did not result in different wine ratings on metaphorical scales such as "richness" and "complexity", which could be used to describe both music and wine⁴.

⁴ More tellingly, participants in a separate online study (N=14) did, in fact, rate the Debussy piece as significantly more complex than the Rachmaninoff piece (on a 10-point scale, M=8.00, SD=1.41, M=5.71, SD=2.02, $t(13)=8.60$, $p<.01$), even though no differences in complexity ratings of the wines were observed in Experiment 4.

A second mechanism proposed by Spence and Deroy (2013b) addresses the possibly low-level influences of crossmodal correspondences on perception. It has been shown previously that crossmodal correspondences, a notion that captures people's tendency to match attributes from stimuli in different senses, can lead to behavioural effects that influence participants' performance on a variety of detection and multisensory integration tasks (Deroy et al., 2013; Spence & Deroy, 2013a). Crossmodal correspondences can be explained by associative learning, arising from statistical co-occurrences in the environment, or by amodal, mediated, and transitive mappings across modalities (Connolly, 2014; Deroy et al., 2013). The association that was observed between the Debussy piece and acidity could possibly be explained via the crossmodal correspondence between high pitch and sourness (Mesz et al., 2011) or between fast tempo and sourness (Bronner et al., 2012). Similarly, the association between the Rachmaninoff piece and fruitiness could be explained by a correspondence between legato articulation and sweetness (Bronner et al., 2012; Mesz et al., 2011) or between slow tempo and sweetness (Bronner et al., 2012), under the transitive property that sweetness corresponds with fruitiness by associative learning⁵. Alternatively, however, the associations between music and flavour documented in Experiment 4 could also result from their emotional similarities (for the hedonic matching hypothesis, see Knöferle & Spence, 2012; and Deroy et al., 2013). For instance, it is possible that the Rachmaninoff piece was associated with fruitiness because the participants found both to be pleasant, whereas the Debussy piece was associated with acidity because they were both rated as less pleasant. Furthermore, in terms of the psychoacoustical roughness of the two pieces, more dissonant chords

⁵ Kontukoski et al. (2015) demonstrated that when asked to mix juices to match with certain pieces of music, participants reliably mixed sweeter drinks when listening to sweet music and more sour drinks when listening to sour music. The sweet music in question had slow tempo and consonant harmonics. The sour music in question had very high pitch and fast tempo.

occur in Debussy than in Rachmaninoff. As dissonance is associated with unpleasantness, this further supports the hypothesis that participants may have found the Debussy piece more unpleasant than the Rachmaninoff. More encouragingly for the hedonic matching hypothesis, studies on crossmodal correspondences have already shown the role of emotion in mediating correspondences between colour and music (Barbieri et al., 2007; Bresin, 2005; Palmer et al., 2013) and between colour and aroma (e.g., Schifferstein & Tanudaja, 2004).

On the subject of colour, one needs to consider that the participants were given the red and white wines in clear glasses. Therefore, it is likely that the colour of the wine itself may have entered into the crossmodal matching process (to be fair, though, even if the participants were given black tasting glasses, they could, of course, still have imagined the colour of the wine as they tasted). For instance, light colours are associated with fast tempo, and dark colours with slow tempo (Palmer et al., 2013). Perhaps the white wine was matched with the faster Debussy piece based on its pale straw colour, and the red wine was matched with the slower Rachmaninoff piece based on its deep ruby hue. In order to control for the role of colour, Experiment 13A and 13B – which also use wine as experimental stimuli – examine pairs of white wines from similar grape varieties and regions.

Finally, order effects cannot be ruled out for these observed associations between music and flavours. Due to the public nature of the event, the music conditions could not easily be counterbalanced across participants. The Debussy piece was performed first, then the Rachmaninoff. Therefore, it is possible that the perception of acidity decreases with exposure (hence the Debussy piece, which played first, was associated with higher acidity than the Rachmaninoff piece). On the other hand, perception of fruitiness might increase with time (hence perhaps explaining why the Rachmaninoff

piece was associated with rated as more fruity than the Debussy piece). Further studies with counterbalanced order of music presentation will be needed to examine this possible explanation⁶.

In general, the results of Experiment 4 demonstrate the existence of crossmodal matches between music and wine as well as the impact of music on modulating the wine drinking experience. Further testing will, however, be needed to reveal the specific mechanisms behind these associations between music and taste/flavours. Moving forward, Experiment 5 assesses whether musical attributes can be associated with the mouthfeel of foods as well as their taste/flavour.

3.4 Experiment 5: Music modulates the perceived creaminess, sweetness, and bitterness of chocolate

3.4.1 Methods

3.4.1.1 Participants

116 participants (65 females and 51 males; mean age=35.11 years, SD=14.49) took part in the experiment, after giving their informed consent. They reported that they did not have a cold or any other known impairment of their sense of smell, taste, or hearing at the time of the study. The participants were informed that they would taste chocolates while sometimes listening to different pieces of music. The experiment lasted for approximately 10 minutes.

⁶ A small follow-up study (N=23) was conducted with counterbalanced music conditions. This revealed a near-significant main effect of music on fruitiness ratings, as in the original study, $F(1,19)=3.289$, $p=.086$, $\eta^2=.148$. Specifically, the wine tasted while listening to Rachmaninoff (M=67.14, SD=2.33) was rated as more fruity than when listening to Debussy (M=56.55, SD=4.34) regardless of the type of wine. There were not, however, any effect of music on ratings of acidity. This might imply that Rachmaninoff somehow increases fruitiness ratings over and above any order effects, whereas it has a less robust effect on acidity.

3.4.1.2 *Auditory stimuli*

Two soundtracks were prepared for this experiment, one corresponding to smoothness/creaminess, and the other to roughness. Along with the bouba-kiki effect (Köhler, 1929, 1947), the relationship between touch and sound highlighted by Eitan and Rothschild (2010) acted as a starting point for the production of the soundtracks. Soft/smooth sounds are usually correlated with long-consonant-legato notes. By contrast, hard/rough sounds are most likely represented by short-dissonant-staccato notes. For example, in Eitan and Rothschild's (2010) study, higher – and louder – pitches/notes were rated as rougher/harder. Moreover, the sound of the violin was rated as rougher/harder and drier than the sound of the flute. That being said, the first soundtrack (produced to be congruent with creaminess, namely the 'creamy soundtrack') consisted of a loop-ascending scale of consonant-long flute notes, mixed with large hall reverberation. The second soundtrack (namely the 'rough soundtrack', was intended to have an opposite effect from creamy soundtrack) consisted of a loop-ascending scale of three blended dissonant-dry pizzicato short violin lines.

Both soundtracks had approximately the same pitch range, and both lasted for approximately 1 minute. They were mastered to have similar dynamics and loudness ($Leq_{1min}=70 \pm 3$ dBA). Note that due to the fact that the rough soundtrack has three melodic lines playing together, with one of those melodies in a higher pitch, it is possible that this soundtrack may have been perceived as higher in pitch, when compared to the creamy soundtrack. The soundtracks can be accessed via the following link: <http://tinyurl.com/creaminess-chocolate> (retrieved in October, 2016).

Initially, a pre-test was conducted in order to verify that naïve listeners would indeed associate each of the soundtracks with the intended texture. Sixty-five people (36 female, 29 male; Mean age=31.63 years, SD 16.46) took part in this pre-test. Here,

the goal was to make the rating scales as comprehensive as possible for naïve listeners, in order to complete the task of evaluating soundtracks in terms of tasting attributes. First, sweetness and bitterness were used since they are usually opposites in terms of valence – as compared to, for example, sourness and bitterness (Reinoso Carvalho et al., 2016; Salgado-Montejo et al., 2015; Yarmolinsky et al., 2009). Second, due to the fact that the opposite of creaminess may have more than one interpretation (i.e., watery, rough, lumpy, etc.), it was decided to structure the scale using roughness as the opposite of creaminess. That being said, the pre-test included two bipolar dimensional scales, one creamy-rough, and another bitter-sweet. This decision was made with the intention of providing an objective way of evaluating two soundtracks that were produced to have opposite perceptual effects on the texture of chocolate.

Each participant listened to both soundtracks and rated them on a 7-point bitter-to-sweet scale ('1'=Very bitter, '4'=Balanced, '7'=Very sweet), and on a 7-point rough-to-creamy scale ('1'=Very rough, '4'=Balanced, '7'=Very creamy). A significant difference between the ratings of the soundtracks was reported (ANOVA, $F(2, 127)=33.62$, $p<.005$, $\eta^2=.35$). The results of the pre-test revealed that the creamy soundtrack was rated as significantly creamier (Mean creamy soundtrack=4.95, $SE=0.17$; Mean rough soundtrack=3.00, $SE=0.17$, $p<.005$) and sweeter (Mean creamy soundtrack=4.49, $SE=0.17$, Mean rough soundtrack=3.20 $SE=0.17$, $p<.005$), than the rough soundtrack. In summary, the participants were able to classify both soundtracks as expected. Figure 3.8 shows the aforementioned ratings.

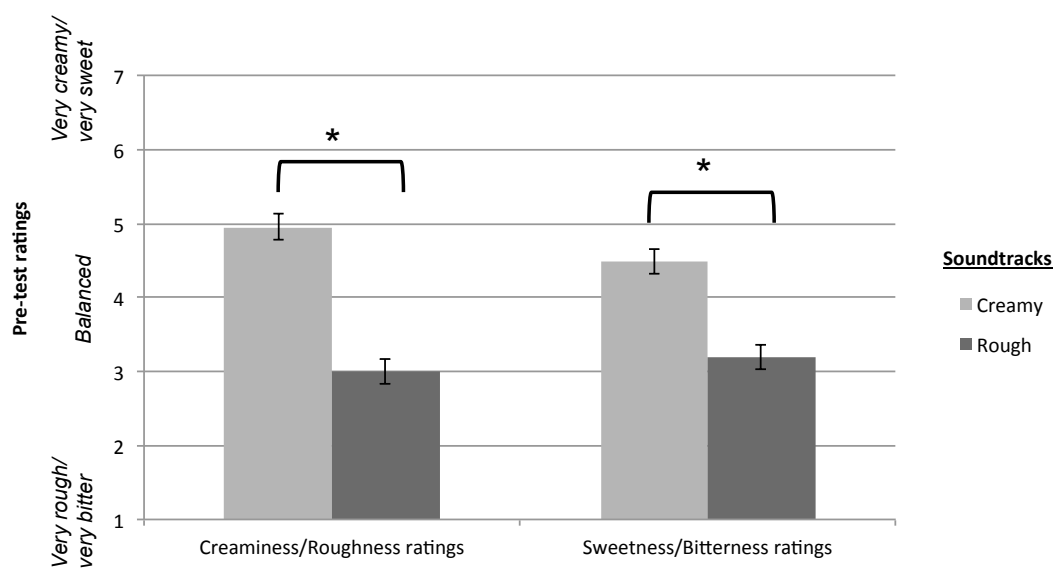


Figure 3.8. Mean values of the pre-test ratings (based on 7-point scale) for Experiment 5. Error bars indicate standard error (SE). Asterisk (“*”) indicate a significant difference at $p < .005$.

3.4.1.3 Gustatory stimuli

In order to test the effect of the sound stimuli on different types of chocolates, two chocolate formulas were chosen. While designing these chocolate samples, it was discovered that the only chocolate formulas that wouldn’t have significant changes in colour would be the ones that do not include milk. It was important to keep the colour of the chocolate samples as similar as possible as to not influence participants’ responses (see Spence, 2015e, for a review). Therefore, it was decided to use only cocoa-based formulas. However, prior the definitive choices of cocoa percentages, pilot studies were performed in order to determine which combination of cacao would be appropriate to use for the experiences. These pilots were developed along with professional chocolatiers, and included several different formulas. Finally, the chosen formulas had 71% and 80% cocoa content (both milk-free chocolate formulas, with the following basic ingredients: cocoa mass, sugar, cocoa butter and natural vanilla flavour). Moreover, each formula was presented in two different molds (see Figure 3.9, top). In total, four different chocolate types were available, one for each group of

participants (see Figure 3.9 bottom). The chocolates were developed at The Chocolate Line factory in Bruges, under the supervision of the award-winning Belgian chocolatier Dominique Persoone (www.thechocolateline.be).

Note that all of the experimental chocolate samples had the same dark brown colour, and similar volume (approximately 2.0 cm³).

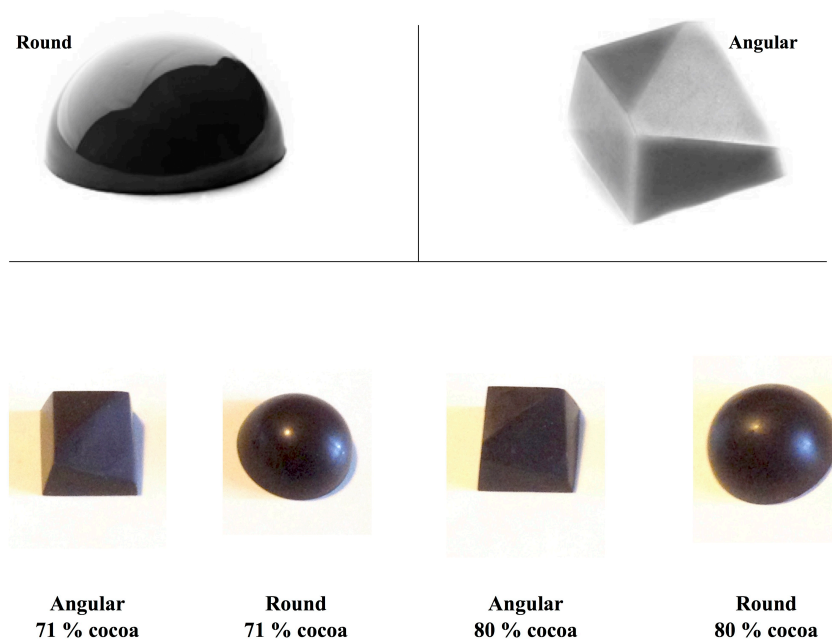


Figure 3.9. Round (top-left) and angular (top-right) shapes of the chocolates used in Experiment 5. Each group tasted one type of chocolate (bottom). All of them had the same colour, and each shape was prepared with the 71%, and 80% cocoa chocolate formulas.

3.4.1.4 Design and procedure

The study was approved by the Social Ethics Committee at KU Leuven – SMEC (Protocol G2016 03 519). Different participants tasted and rated two identical chocolates in two trials, each time listening to one of the two soundtracks (all ratings based on 7-point scales; see supplementary material for complete questionnaire). The independent variables for each experiment were sound condition (within-participants)

and chocolate type (between-participants). The dependent variables were the ratings that the participants made for each trial. The soundtracks were presented in a counterbalanced order across participants. The order of presentation of the questions was fully randomized as well.

The ninth floor of the Musical Instruments Museum Brussels (MIM) was chosen as the site for the experiments. Due to its independent location inside of the museum, being located between the museum's restaurant on the top floor and the rest of the exhibitions below, it was possible to have a well-controlled experimental environment during experimental hours. Four rectangular tables were placed in the experimental area, one for each experiment, with two computers on each table. The natural light present in the experimental area was sufficient to provide a more 'intimate' ambience. Therefore, artificial light was kept to a minimum.

Each participant was seated in front of a computer screen. Each participant had three chocolates, a glass of tap water, a pair of headphones, a computer mouse, and a keyboard to interact with the survey. The calibration of the reproduction system was set to a comfortable – but at the same time immersive – listening level of $Leq_{1min}=70 \pm 3$ dB (corresponding to 50% of the volume of the existent sound system). The soundtracks were presented over SONY MDRZX310 headphones. Note that the participants were not able to hear the sounds from the other participants' headphones.

The survey consisted of an electronic form containing three main steps. In the first step of the survey, the participants were instructed to read and accept the conditions of the informed consent before entering their personal details. They were instructed to drink water before eating each one of the experimental chocolates. Prior to eating, the participants were also instructed not to chew the chocolates, but to let them melt in

their mouth. This instruction was included in order to help standardizing the way that all the participants experienced the texture of the experimental chocolates.

In a second step, the participants had to taste a small drop of bitter chocolate, as a covariant (such chocolate drop was part of an industrial batch of ‘Callebaut Dark Callets’, recipe 70-30-38, with 70.5% cocoa). Here, they rated how much they liked it, and how sweet, bitter, and creamy they thought that it was. In this part of the experiment, the participants tasted and rated the chocolate without any sound.

In the third and final step, the participants were randomly assigned to one of four groups. This assignment defined which of the four available chocolate types (71% angular, 71% round, 80% angular, or 80% round) they would taste (see Figure 3.9). Here, they had to taste and rate the same chocolate twice, each time listening to one of the two soundtracks. Both chocolates were numbered. Hence, the participants were instructed to eat chocolate Number 1 first, while listening to the first soundtrack, and then chocolate Number 2, while listening to the second soundtrack. After tasting each chocolate, they rated how much they liked it, how sweet, bitter, and creamy they thought it was. They also rated how much they liked each soundtrack, and how much they thought it matched the taste of the chocolate (all ratings based on individual 7-point scales, with ‘1’ being ‘Not at all’, ‘4’ ‘Neutral’ and ‘7’ ‘Very much’; see supplementary material for complete questionnaire).

Together with the written guidelines concerning the experiment, at least one supervisor was present during the experiment in order to provide guidance and support. Upon finishing the experiment, the participants were instructed to leave the room without discussing any details with the next group of participants. The experiment lasted for around 10 minutes.

3.4.2 Results and discussion

Multivariate tests: Chocolate type did not have a significant effect on the participants' ratings ($F(18,327)=1.46$, $p=.103$, $\eta^2=.074$), but soundtrack condition did ($F(6,107)=6.26$, $p<.005$, $\eta^2=.260$). More specifically, participants reported that the chocolates tasted creamier while listening to the creamy soundtrack, as compared to the rough soundtrack ($p=.002$; Mean creamy soundtrack=4.07, $SE=0.14$; Mean rough soundtrack=3.67, $SE=0.14$). The participants also reported that the chocolates tasted sweeter while listening to the creamy soundtrack ($p=.004$; Mean creamy soundtrack=3.77, $SE=0.13$; Mean rough soundtrack=3.39, $SE=0.12$), and that the chocolates tasted more bitter while listening to the rough soundtrack ($p=.010$; Mean creamy soundtrack=3.92, $SE=0.12$; Mean rough soundtrack=4.23, $SE=0.12$).

Moreover, the participants reported having liked the creamy soundtrack significantly more than the rough soundtrack ($p<.005$; Mean creamy soundtrack=3.64, $SE=0.16$; Mean rough soundtrack=2.53, $SE=0.13$). When comparing how well they thought the soundtracks matched the taste of the chocolates, a trend suggested that the creamy soundtrack might have been a better match than the rough soundtrack ($p=.077$; Mean creamy soundtrack=3.74, $SE=0.15$; Mean rough soundtrack=3.36, $SE=0.15$). Finally, no significant differences were found in terms of participants' enjoyment of the chocolates when comparing the two soundtrack ratings ($p=.16$; Mean creamy soundtrack=4.49, $SE=0.12$; Mean rough soundtrack=4.32, $SE=0.13$; see Figure 3.10).

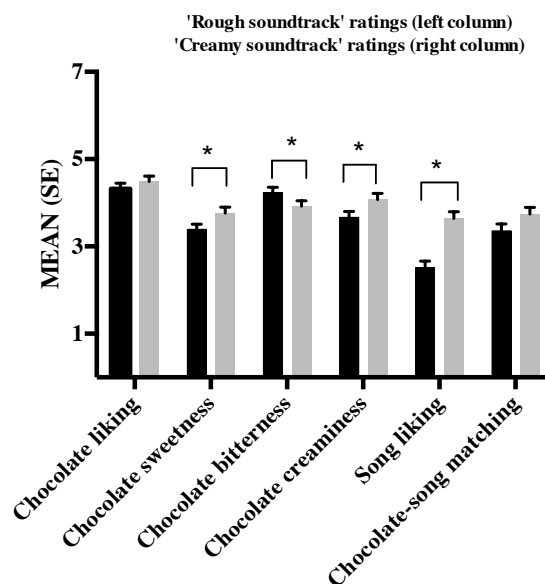


Figure 3.10. Participants' mean ratings (based on 7-point scale) in Experiment 5. For each attribute, the left column (black) corresponds to rough soundtrack ratings, and the right column (grey) to creamy soundtrack ratings. Error bars indicate standard error. Asterisks "*" indicate a significant difference at $p=.01$, between the rough and creamy soundtracks ratings.

The participants were further subdivided in two groups – those who liked the creamy soundtrack more than the rough soundtrack ($N=71$), and the rest ($N=38$). Such a grouping – namely 'soundtrack preference' – was included as an independent variable as part of the main analysis. The results revealed a significant interaction between soundtrack condition and soundtrack preference ($F(6,109)=47.89$, $p<.005$, Pillai's Trace=.98), in particular for chocolate liking ($F(1,114)=12.51$, $p=.01$, $\eta^2=.10$), chocolate-soundtrack match ($F(1,114)=32.44$, $p<.005$, $\eta^2=.22$), and creaminess ($F(1,114)=5.02$, $p=.027$, $\eta^2=.04$). For chocolate liking and chocolate-soundtrack match, the participants tended to give a higher rating to whichever soundtrack they preferred. However, for creaminess ratings, Bonferroni-corrected post-hoc testing revealed that only the group that preferred the creamy soundtrack reported higher creaminess ratings while listening to the creamy soundtrack ($p<.005$). By contrast, there was no significant interaction effect of soundtrack condition and soundtrack

preference on sweetness ($F(1,114)=0.41$, n.s.) or bitterness ratings ($F(1,114)=0.65$, n.s.).

Correlations: Table 3.3 shows the calculated correlations. Sweetness and creaminess ratings were positively correlated with chocolate liking, whereas bitterness ratings were negatively correlated with the three aforementioned attributes. Moreover, chocolate liking and creaminess were positively correlated with soundtrack liking and chocolate-soundtrack matching. Finally, soundtrack liking and chocolate-soundtrack matching were positively correlated. In summary, there was a positive relationship between soundtrack liking, chocolate liking, and chocolate sweetness/creaminess ratings.

2. /	Chocolate Liking	Chocolate sweetness	Chocolate bitterness	Chocolate creaminess	Soundtrack liking	Chocolate-soundtrack match
Chocolate liking	1	.329	-.140	.360	.256	.251
Chocolate sweetness	.329	1	-.296	.424	.168	.054
Chocolate bitterness	-.140	-.296	1	-.252	-.048	.076
Chocolate creaminess	.360	.424	-.252	1	.232	.189
Soundtrack liking	.256	.168	-.048	.232	1	.525
Chocolate-soundtrack match	.251	.054	.076	.189	.525	1

Table 3.3. Pearson correlation coefficients between participants' ratings in Experiment 5. Bold indicates significant correlations at the .05 level.

In Experiment 5, two soundtracks were produced with the aim of modulating the perceived creaminess of chocolate. The first soundtrack was produced to be congruent with creaminess, and the second with roughness. Note that both soundtracks were compared and validated by means of a pre-test. In total, four chocolate samples were produced, with a combination of two shapes (round/angular) and two formulas (71% and 80% cocoa). The participants were subdivided into four groups, one corresponding to each of the available chocolate types. In each group, the participants tasted and rated the same chocolate twice, each time under the influence of one of the soundtracks.

The results revealed that the soundtracks had the predicted effect on the perceived creaminess of the chocolates (see Figure 3.10). In particular, the creamy soundtrack significantly elevated ratings on creaminess, when compared to the effects of the rough soundtrack (that potentially decreased the perceived creaminess). In addition, there was a direct relationship between ratings of sweetness and creaminess. Table 3.3 reveals that creaminess and sweetness ratings were positively correlated, whereas creaminess was negatively correlated with bitterness. These correlations also highlight the fact that creaminess and sweetness are positively correlated with chocolate liking (see Table 3.3). One possible explanation for these correlations is that there may have been a general confound in the mind of the participants between creaminess and sweetness ratings. In particular, sweetness was perhaps used as a proxy for creaminess. Two analogous cases have been reported previously with alcoholic beverages, where those who are generally poor at estimating alcohol

content, may use taste cues as a substitute. For instance, Stafford et al. (2012) reported that ratings of alcohol content was correlated with bitterness ratings in vodka when participants tasted a variety of vodka-juice mixtures at different alcohol levels. On top of that, it seems that high-impact flavour may be used as a proxy for alcohol content as well, such as hoppiness/bitterness in the case of beer (i.e., Reinoso Carvalho et al., 2016; see also Harrar et al., 2013).

It is also important to highlight the fact that, when producing the creamy soundtrack, some musical attributes that were here considered as congruent with creaminess, are also parameters that are usually correlated with sweetness. Similarly, musical attributes that here were considered to correspond to roughness are also commonly correlated with bitterness. For example, consonance, legato articulation, and low discontinuity – musical parameters that are used in the creamy soundtrack – are all musical parameters that were previously reported as being congruent with sweetness. Here the above-mentioned parameters were used to be congruent with creaminess as well. On the other hand, higher discontinuity and dissonance (in this case harmonic) are parameters that were previously reported as congruent with bitterness, and were here used as incongruent with creaminess (see Knöferle & Spence, 2012; and Knoeferle et al., 2015, for a review on musical and psychoacoustic parameters and their correspondent congruency with basic taste attributes). That being said, it would be plausible that these soundtracks could also have had an modifying effect on the perceived texture of the chocolates while, in parallel, potentially having a perceptual effect on the chocolate's sweetness and bitterness.

In general, the participants liked the creamy soundtrack significantly more than the rough soundtrack. On the basis of this result, it could be presumed that the greater enjoyment of the creamy soundtrack could have enhanced chocolate liking (Kantono

et al., 2015, 2016; see Cheskin, 1972, and Spence, 2016b, on the notion of sensation transference), which then heightened the perceived creaminess of the chocolate (shown in the correlations present in Table 3.3). A further subdivision of the data (splitting the participants by soundtrack preference) revealed that only those who preferred the creamy soundtrack rated the chocolate as creamier while listening to the creamy soundtrack, thus implying a role of sensation transference in modulating participants' responses. It is equally important to note that no similar interaction effect was observed for either sweetness or bitterness ratings.

Furthermore, differences in people's liking for the soundtrack did not affect their overall enjoyment of the chocolates⁷ (see Figure 3.10). Previous similar studies have reported that music tends to have an effect in the hedonic and perceptual ratings on food/beverages multisensory tasting experiences, with sound enhancing the enjoyment of food and drinks (cf. Kantono et al., 2016; Reinoso Carvalho et al., 2015a, b, c; Spence et al., 2013; Wang & Spence, 2015b).

Experiment 5 demonstrates that sounds can, in some cases at least, have a perceptual effect on food without altering its hedonic experience, regardless of the fact that people might prefer one sound stimulus over the other (cf. Wang & Spence, 2015a). The results revealed that the soundtracks that were produced specifically for Experiment 5 could be considered as a reliable baseline for the production of other soundtracks, to be used in future similar assessments.

Nevertheless, there are a few limitations that should be mentioned here and which deserve to be assessed in future work. Principally, with these results it is difficult to conclude whether there is only one, or perhaps several mechanisms underlying these

⁷ On the other hand, there were no individual correlations between soundtrack liking and chocolate liking, which is similar to the findings from Experiment 4.

sound-chocolate associations. It would appear that there are a number of explicit crossmodal sound-flavour correspondences, driven mainly by the salient musical attributes of each soundtrack. However, since there is a clear correlation between soundtrack and chocolate liking, it could be argued that the results of Experiment 5 hinge on some form of sensation transference effect rather than reflecting a ‘true’ crossmodal correspondence, at least when it comes to the creaminess ratings. Still, most of the musical attributes used in these exercises were chosen on the basis of contrast (think of consonant versus dissonant harmonies, reverberant versus dry ambiances, and so on). That being said, a plausible assumption would be that assessments such as this one would most likely be under the constant subjective preference of each participant, especially when working with those individuals lacking of specific musical training. For instance, most people prefer listening to consonant harmonies over dissonant ones, and so on.

Moreover, it is also worth highlighting the fact that the soundtracks produced for this experiment are simple in terms of their musical composition. A similar exercise could use more complex sound stimuli (i.e., with more instrumental layers and/or more sound effects). This way it would be possible to assess the potential of, for example, using popular music formats in order to modulate the perceived creaminess of chocolate (just think of all the music, e.g., advertising jingles, that are not produced with any thought given to sound-taste correspondences, but which could nevertheless still have a perceptual effect on people’s perception of food and beverages).

3.5 General discussion and conclusions

Several different examples of auditory modulation of the eating/drinking experience have been demonstrated in this chapter. Experiment 3 revealed that melodies with different harmonic structures are associated with different basic tastes, and can influence taste and pleasantness ratings of fruit juice. Experiment 4 reinforced the notion that people can make reliable matches between specific pieces of music and specific wines, and further demonstrated that participants' ratings of fruitiness and acidity for the wines differed significantly based on the music they heard. Finally, Experiment 5 showed that soundtracks produced to correspond with creaminess and roughness can influence the perceived creaminess, as well as sweetness and bitterness, of chocolates.

The results of Experiments 3-5 naturally invite a key question, one that the majority of the rest of this thesis will work to address – what are the putative mechanisms by which sound-flavour correspondences might influence people's experience of eating and drinking? As the discussions for each experiment revealed, there are a variety of mechanisms that could potentially be at work. In the next chapter, I will begin by addressing perhaps the most straightforward mechanism – do the variety of soundtracks actually have a genuine perceptual effect? Or are participants just changing their ratings to, for instance, satisfy the experimenters?

The concern with demand effects is highlighted by the fact that the majority of the studies that have been published to date have assessed the perceptual effects of sound on taste perception (i.e., on specifically gustatory attributes; Crisinel et al., 2012, Spence et al., 2013, 2014a, b; Velasco et al., 2013a) utilized a within-participants experimental design, sometimes with the participants being aware that the taste stimuli were actually the same in the different auditory conditions (see Spence et al.,

2013, 2014a, b; Velasco et al., 2013a). Experiment 3 and 5 certainly fall into this category. It is worth noting that in order to avoid any possibility that the participants might come to realise, or believe, that they were sampling the same drink across different conditions, Experiment 4 used a crossover design in which each participant rated two distinctively different wines.

One further point of interest raised by the results from Experiments 3-5 is the role of individual differences. For example, Experiment 3 brings up the possibility of supertaster status and sweetness liking as factors influencing the extent to which consonant/dissonant soundtracks might modify food evaluation, while Experiment 4 mentions wine tasting expertise as a similar factor. As a result, Chapter 9 is devoted to covering individual differences that might moderate the effect of crossmodal modulation.

4 Response bias as a possible mechanism underlying the auditory modulation of taste/flavour

4.1 Introduction

“Sonic seasoning”, the idea that sound can be used to alter people’s taste perception, is becoming an increasingly popular term both in the academic literature (Reinoso Carvalho et al., 2015a; Spence 2015; Spence & Wang, 2015a, b; Wang et al., 2015) but also in the mainstream media (e.g., Basu, 2016; Fleming, 2014; Knapton, 2014). Previous research has shown that music is associated with basic taste words (for a review, see Knöferle & Spence, 2012). Crisinel et al. (2012) first demonstrated that beyond any crossmodal associations between sounds and taste words, auditory stimuli could also affect people’s taste evaluations. The participants in their study were given samples of bittersweet cinder toffee to evaluate while listening to one of two soundtracks that had been specifically composed to correspond to either sweet or bitter tastes. The participants rated the cinder toffee samples higher on the sweet-bitter scale (i.e., more sweet and less bitter) while listening to the sweet soundtrack than while listening to the bitter soundtrack. In Chapter 3, a variety of studies demonstrated the effect of specially designed soundscapes on the taste and even mouthfeel attributes of fruit juices (Experiment 3), wines (Experiment 4), and chocolates (Experiment 5).

What remains unclear, however, is the mechanism (or, more likely, mechanisms) underlying such modulatory effects. One of the most fundamental questions is whether such “perceptual” effects really are perceptual, or whether instead they reflect biased self-reported ratings (Litt & Shiv, 2012). In other words, it is not clear

whether the soundscapes actually modify participants' sensory experience, or whether instead the participants are motivated to evaluate the food/beverages differently due to some other reasons unrelated to their actual perceptual experience (generally referred to in this thesis as "response bias"), such as wishing to conform to what they perceive to be the experimenter's expectations. In a similar context, Lee and colleagues (2006) set out to determine whether external conceptual information about a product only affects people's preferences, or whether it can alter their actual gustatory experiences. They demonstrated that the timing of information regarding the content of the "MIT Brew" – a beer secretly adulterated with a few drops of balsamic vinegar – affected participants' preference for the mystery beverage. More specifically, participants' preference for the beer was decreased only when the disclosure of the secret information preceded the tasting, but not if the disclosure took place after tasting. The authors inferred from this that the disclosure influenced the real-time experience of drinking itself, rather than altering participants' retrospective interpretation of the experience. If the extrinsic information (the addition of balsamic vinegar) had acted to bias participants' responses ("balsamic vinegar in beer, gross!"), then participants' beer preference should have been similar regardless of when the information was presented.

In a similar vein, the goal of Experiment 6 is to uncover whether taste-congruent soundtracks can influence the actual experience of tasting. The timing of when the sonic stimuli were presented to the participants was manipulated in the experiment; If sound truly changes the perceptual experience of taste, then one would expect there to be a greater difference in taste ratings between sweet/bitter soundtrack conditions when the soundtracks are presented while the participants are tasting the chocolate samples, as compared to when the soundtracks are presented after tasting. However, if

sound only biases participants' self-report, then it should not matter whether the sonic stimuli are presented during, or immediately after, tasting; one would expect to see differences in taste ratings in both scenarios.

4.2 Experiment 6: The effect of auditory stimuli timing on sound-taste interactions

4.2.1 Methods

4.2.1.1 Participants

A total of 113 participants (76 women, 35 men, 2 unreported) aged 18-51 years ($M=24.66$, $SD=4.81$) took part in the study. The data collection was spread out over two sessions. Participants were recruited via BI Norwegian Business School's participant recruitment platform. All of the participants gave their informed consent to take part in the study. None of the participants reported a cold nor any other known impairment of their sense of smell, taste, or hearing at the time of the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

4.2.1.2 Auditory stimuli

Wang et al.'s (2015) study compared and ranked 24 different soundtracks that had previously been designed to be associated with taste attributes (comparison based on ratings made on basic tastes scales). The sweet and bitter soundtracks with the highest number of matches in Wang et al.'s (2015) study were chosen for Experiment 6. The sweet soundtrack (chosen by 89 out of 100 participants), was developed by Jialing Deng and Harlin Sun as a soundtrack for Synaesthetic Appetiser, part of Deng's Masters of Arts Thesis project (June, 2015). The bitter soundtrack (chosen by 42 out

of 100 participants) was the one used by Crisinel et al. (2012) in their sound-taste modulation study. Both soundtracks were matched higher than if the participants had matched the tastes to soundtracks by chance (25%). The soundtracks were edited to last approximately 30 seconds each, and root-mean-square (RMS) equalised (Bharitkar & Kyriakakais, 2006). They can be heard at the following link: <https://soundcloud.com/janicewang09/sets/timing>.

4.2.1.3 Gustatory stimuli

70% Lindt chocolate was used for the study for its fairly complex taste, ambiguous sweet-bitter balance, and commercial availability. Each sample consisted of approximately 3g of chocolate served in a small clear plastic cup.

4.2.1.4 Design and procedure

The study was designed with timing condition (sound during or after tasting) as a between-participants factor and soundtrack type (bitter or sweet) as a within-participants factor. Each participant was randomly assigned to one timing condition, completed four trials, and heard each soundtrack twice.

The experimental sessions comprised of up to ten participants at a time. Each participant was seated in front of a computer screen in an experimental cubicle, isolated from other cubicles by opaque plastic separators. No two participants sat immediately adjacent to one another during the experimental sessions. The experiment was programmed on the Qualtrics online survey platform and participants responded by using the mouse to click or drag the indicator on the continuous rating scales. Each participant was given four chocolates on labelled plates (A, B, C, and D) as well as tap water and crackers to cleanse their palates. Participants were randomly assigned to either the during-tasting (N=56) or after-tasting timing condition (N=57).

Participants in the during-tasting condition were instructed to start tasting the chocolate once they heard an auditory cue. They were then instructed to cleanse their palate with water and crackers before moving onto the next page and evaluating the chocolate's enjoyment and its taste on a bitter-sweet scale. Participants in the after-tasting condition tasted the chocolate sample in silence, cleansed their palates, then moved onto the evaluation page, where they were instructed to start the evaluation as soon as they heard an auditory cue. (Participants cleansed their palates right after tasting to ensure that the soundtracks during the evaluation stage did not act on any residual tastes in the mouth). The order of the soundtracks and the order of the rating scales were randomised.

Finally, the participants listened to each soundtrack again and rated it on scales of 1-7 for valence (unpleasant-pleasant) and arousal (calming-exciting). Participants also rated how much each soundtrack matched each of four basic tastes (sweet, bitter, sour, and salty).

The experiment lasted for around 20 minutes and participants were paid 50 NOK (approximately 5 GBP) for their time.

4.2.2 Results and discussion

Bittersweet and liking ratings were first averaged over the two identical soundtrack trials for each participant. Separate RM-ANOVAs were conducted for bittersweet ratings and for liking ratings as the dependent variables, with timing as the between-participant factor and soundtrack type as the within-participant factor.

The mean values of the participants' ratings for the chocolates are shown in Figures 4.1 and 4.2. For bitter-sweet ratings, the RM-ANOVA test revealed that there was a significant main effect of soundtrack ($F(1,111)=6.80$, $p=.010$, partial $\eta^2=0.06$), where

the chocolates experienced with the sweet soundtrack (either during or after tasting) were rated as sweeter ($M=42.96$, $SE=1.61$) than the chocolates experienced with the bitter soundtrack ($M=39.45$, $SE=1.66$, $p=0.01$). More importantly, there was a significant interaction between timing condition and soundtrack type ($F(2,90)=3.81$, $p=.026$, partial $\eta^2=0.08$). Specifically, those participants who heard the soundtrack while tasting evaluated the chocolates as significantly sweeter when the sweet soundtrack was playing ($M=46.03$, $SE=2.29$) than when the bitter soundtrack was playing ($M=39.37$, $SE=2.35$, $p=.001$). Music had no such effect on taste for participants who heard the soundtrack after tasting, that is, during evaluation ($p=.85$).

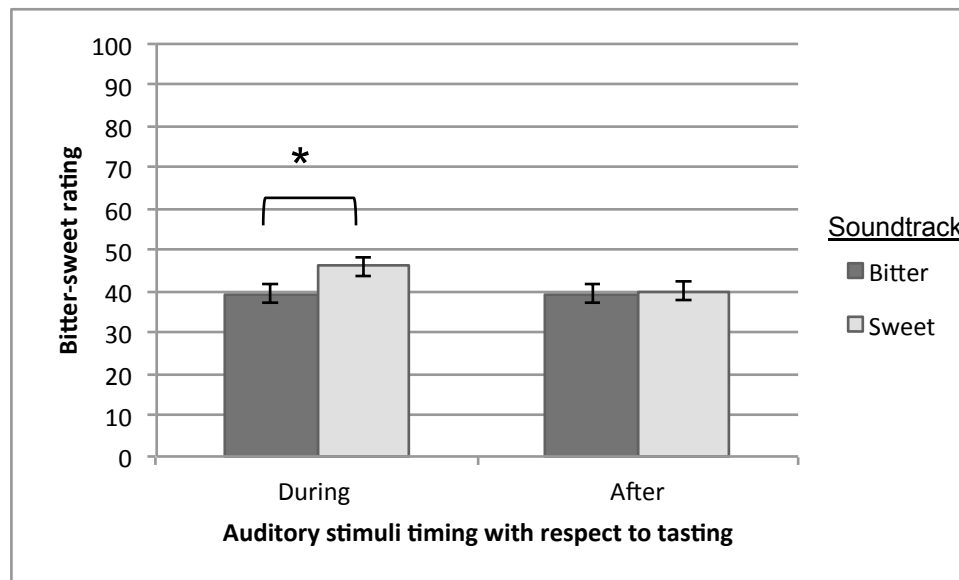


Figure 4.1. Mean values of bitter-sweet ratings for chocolates, grouped by auditory stimulus timing, in Experiment 6. Error bars indicate standard error. Asterisk ‘*’, indicates statistical significance at $p<.05$.

In terms of chocolate liking, there was a significant main effect of timing condition ($F(1,111)=5.71$, $p=.02$, partial $\eta^2=0.05$), where the chocolates were liked more when the soundtracks were played during tasting ($M=5.28$, $SE=0.19$), than after tasting ($M=4.62$, $SE=0.19$, $p=.02$). Furthermore, there was a trending interaction effect

between timing condition and soundtrack type, ($F(1,111)=2.91$, $p=.09$, partial $\eta^2=.026$), whereby participants who heard the soundtracks while evaluating the chocolates preferred the chocolates while listening to the sweet soundtrack ($M=5.44$, $SE=0.20$) as compared to when listening to the bitter soundtrack ($M=5.12$, $SE=0.22$, $p=.05$). However, no differences were observed for participants who heard the soundtracks after tasting ($p=.67$). In addition, there was a significant overall positive correlation between bittersweet and liking ratings (Pearson $r_{226}=0.39$, $p<.0005$).

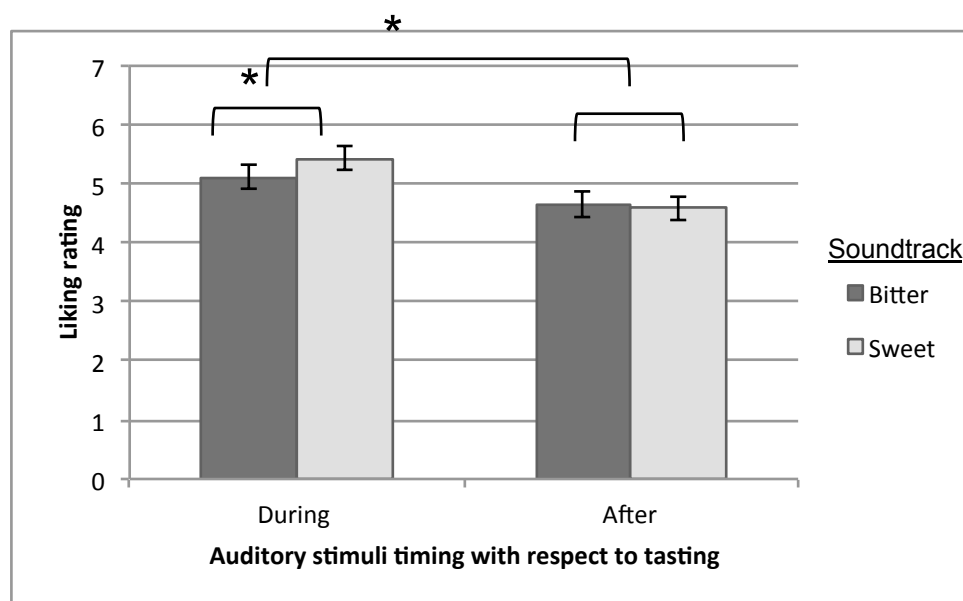


Figure 4.2. Mean values of liking ratings for chocolates in Experiment 6, grouped by timing of the auditory stimuli. Error bars indicate standard error. Asterisk ‘*’, indicates statistical significance at $p<.05$.

Assessing the emotion ratings of the soundtracks, participants rated the sweet soundtrack as significantly more pleasant than the bitter soundtrack ($M_{\text{sweet}}=5.12$, $SE_{\text{sweet}}=0.16$, $M_{\text{bitter}}=3.02$, $SE_{\text{bitter}}=0.17$, $p<.0005$). There was no significant difference in terms of arousal ratings between the two soundtracks ($M_{\text{bitter}}=3.50$, $SE_{\text{bitter}}=0.16$, $M_{\text{sweet}}=3.49$, $SE_{\text{sweet}}=0.17$, $p=.95$). In terms of the participants’ sound-taste association

ratings for each soundtrack (see Figure 4.3), the bitter soundtrack was rated as matching bitterness significantly better than the other tastes ($p < .0005$ for all comparisons), and the sweet soundtrack was rated as matching with sweetness significantly better than with the other tastes ($p < .0005$ for all comparisons).

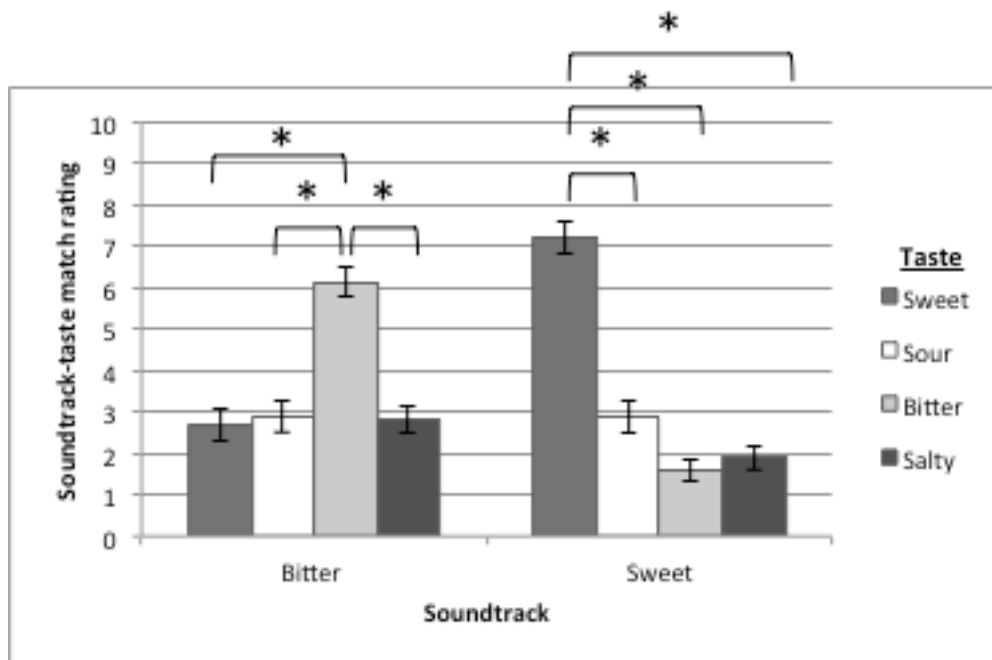


Figure 4.3. Mean values of soundtrack-taste match ratings for the sweet and bitter soundtrack used in Experiment 6. Participants rated how well each soundtrack matched the four basic tastes (sweet, sour, bitter, and salty) on a scale of 1-10. Error bars indicate standard error. Asterisk “*”, indicates statistical significance at $p < .05$.

The results of Experiment 6 demonstrate that the timing of sonic stimuli does indeed make a difference to the modulatory effects of sound on taste perception. The sweet soundtrack enhanced the sweetness of chocolates as compared to the bitter soundtrack only when the soundtracks were played *during* the tasting process. When the soundtracks were played only during the rating process, post-tasting, no such modulatory effects were observed.

The results of Experiment 6 provide evidence that the crossmodal modulation of taste by audition, first reported by Crisinel and her colleagues (2012), involves a genuine interaction between the soundtrack and the tasting experience that goes beyond altering people's retrospective interpretation of the experience. In the study reported by Crisinel et al (2010), the soundtracks were played while participants tasted toffee samples, therefore the observed effect on taste ratings could have been perceptual and/or attributed to demand effects. In Experiment 6, participants tasted chocolate samples in one of two conditions; the first was identical to the aforementioned Crisinel et al. (2012) study. In the second condition, where participants heard soundtracks only after they tasted the chocolate samples, any effect of the soundtracks on taste ratings could not be perceptual in origin. The fact that the modulatory effect was only observed when sound was heard during the tasting, and not after tasting, provides evidence that the soundtracks influenced the real-time experience of eating. To further verify that the soundtracks had an effect on the actual perceptual experience, a plausible future study might involve running a functional Magnetic Resonance Imaging (fMRI) experiment where participants listen to a variety of soundtracks (taste congruent and otherwise) inside the brain scanner. If the sweet soundtrack were found to enhance activity in the sweet gustatory field in the primary taste cortex (Chen et al., 2011; Peng et al., 2015; Reiter et al., 2015) relative to, say, white noise, then there would be evidence for a truly perceptual influence (see Spence, 2016c; Woods et al., 2011).

A further question, then, is which mechanisms might explain the crossmodal modulation effect observed in Experiment 6. According to one account, the sound might focus a person's attention on a specific (i.e., corresponding) taste. In support of this account, the results indeed showed that participants associated the bitter and

sweet soundtracks with their intended tastes. This would also explain why the modulatory effect was only observed while the soundtracks were played during the tasting phase; once the participants had finished tasting, the soundtracks ceased to exert any influence since participants did not have taste stimuli to attend to. Attention will be explored as a mechanism for crossmodal modulation in Chapter 6.

Another possibility here is that the listener might transfer their feelings about the sound to the taste stimuli (Cheskin, 1972); therefore, if participants find the soundtrack pleasant, they might also find the taste stimuli more pleasant than they would otherwise have done. Consequently, they may perceive those tastes they happen to like more, more strongly. Support for this theory is also seen in Experiment 6, whereby participants preferred the sweet soundtrack more than the bitter soundtrack, and the chocolates were liked more in the sweet than in the bitter soundtrack condition. Enhanced sweetness ratings under the sweet soundtrack condition can therefore perhaps be explained by the positive correlation between chocolate liking and sweetness rating.

To follow-up, I assessed whether participants' valence ratings of the soundtracks affected the taste modulatory effect observed in Experiment 6. Repeating the same analysis on taste ratings while introducing the difference in soundtrack valence ratings ($\text{valence}_{\text{sweet}} - \text{valence}_{\text{bitter}}$) as a covariate revealed that there was still a significant interaction effect of soundtrack and auditory stimuli timing even after controlling for the effect of soundtrack liking, $F(1,110)=5.46$, $p=.02$, partial $\eta^2=.05$. Therefore, sensation transference, while relevant, is not the only mechanism at work here.

On a separate note, the observation that the chocolates were liked more in the sound-during-tasting condition as compared to in the music-post-tasting condition might be of interest to those in the food service industry. It echoes the results of Spence et al.

(2013), where wines tasted while listening to a matching piece of music were liked more than wines tasted in silence.

In summary, the results of Experiment 6 suggest that the auditory-taste modulation effect observed here and in previous research (e.g., Crisinel et al., 2012; Reinoso Carvalho et al., 2015a, b, 2016; Wang & Spence, 2016) involves participants' actual perceptual experience. Furthermore, it likely involves a host of different mechanisms beyond simply sensation transference. In Chapters 5-8, I will examine these possible mechanisms – such as the role of expectations (Chapter 5), attention (Chapter 6), emotion (Chapter 7), and direct physiological effects (Chapter 8).

5 Sensory expectation as a possible mechanism underlying the auditory modulation of taste/flavour

5.1 Introduction

A growing list of crossmodal correspondences have now been demonstrated between sonic properties and different flavour attributes; for example, high pitch is associated with sweet and sour basic tastes as well as vanilla flavouring, and low pitch is associated with bitter taste and coffee aromas and flavours (e.g., Crisinel & Spence, 2009, 2010b; Mesz et al., 2011; see Knöferle & Spence, 2012, for a review). Beyond the basic tastes, however, there are more complex flavours and tactile sensations involved in eating and drinking (Auvray & Spence, 2008; Reinoso Carvalho et al., 2017; Stevenson, 2009). Experiment 7 is concerned with the trigeminal sensation of spiciness/piquancy¹, a burning or warming sensation triggered by the activation of capsaicin receptors in the mouth (Caterina et al., 1997). To date, no crossmodal correspondences have been documented between sounds and spiciness, but there is recent evidence of correspondences involving spiciness and attributes from the other senses². Visually, spiciness is associated with the colour red. For instance, Shermer and Levitan (2014) demonstrated that the intensity (saturation) of red colouring of a salsa affected its perceived spiciness. Going one step further, Tu et al. (2016) recently demonstrated that even the colour of the plate on which a food is served can change both expected and actual spiciness ratings. According to Shermer and Levitan, the red colouring used in their study to modify the colour of the salsa samples generated

¹ Within the scope of the present paper, the term spiciness is used interchangeably with the term piquancy to refer to the hot/burning sensation from chili peppers.

² There is also evidence that the smell of pepper, an approximate olfactory analog to the trigeminal spicy sensation, is matched with angular, rather than rounded, shapes (Seo et al., 2010).

sensory expectations of spiciness, which, in turn, enhanced the perceived spiciness of the salsas.

If visual stimuli can generate sensory expectations, it stands to reason that auditory stimuli may be able to do the same. When it comes to sounds and taste, the research that has been published so far has revealed that soundtracks that are crossmodally congruent with basic tastes could enhance those tastes in food and drinks (Crisinel et al., 2012; Reinoso Carvalho et al., 2015a, b, 2016; Wang & Spence, 2016). However, there has been no research published to date with the goal of delineating the possible mechanisms that underlie the auditory modulation of taste. One hypothesis, to be examined in this chapter, is that auditory stimuli corresponding to a specific taste/flavour might also enhance the said taste/flavour by way of generating expectations. In other words, if a specific soundscape can act to enhance expectations for a specific taste, then hearing the soundscape could lead people to involuntarily adjust their sensory perceptions to try and conform to their initial taste expectations (Lelièvre et al., 2009; see Piqueras-Fiszman & Spence, 2015, for a review). Therefore, the aim of Experiments 7 and 8 is to verify whether spicy (Experiment 7) or sweet/bitter (Experiment 8) soundtracks might enhance spiciness/sweetness/bitterness ratings by acting on participants' sensory expectations.

In Experiment 7A, an online study was conducted in order to explore the extent to which individuals consistently associate different musical parameters with spiciness. A spicy soundscape was then composed, based on the findings from Experiment 7A, and was used in Experiments 7B - D in order to determine whether sound can influence participants' expected and actual taste ratings for real food. Experiment 8 follows-up by directly examining the role of soundtrack timing on auditory taste

modulation, especially in the case where the soundtrack is heard only before food is consumed.

5.2 Experiment 7A: Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack

5.2.1 Methods

5.2.1.1 Participants

44 participants (27 women, 17 men) aged between 25 to 65 years of age ($M=38.47$, $SD=9.81$) took part in the study. The participants gave their informed consent, and reported no hearing impairments. The participants were recruited from mailing lists. The experiment was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

5.2.1.2 Auditory stimuli

Short sound clips were composed by iV (an audio branding consultancy) to reflect different variations of specific musical parameters. Short musical segments (2-10 seconds) were composed with three levels each for articulation (staccato or legato), distortion, tempo, pitch height, complexity, length of attack, length of decay, length of both attack and decay, and harmony. Musical segments were also composed with two levels of difference for modality (major/minor). Three additional questions were included to explore cultural associations reflected in a series of ambient, classical, and percussion music samples. Three ambient samples were presented: the first was bass heavy and rhythmical, the second was relaxing and high pitched, and the third was sharp and sporadic with stuttering drum beats. The classical music samples included

Mozart's violin concerto No 3 Movement 1 (with a regular 4/4 rhythm) and Saint-Saëns' violin concerto No 3 Movement 3 (featuring a high-pitched and tempestuous string solo). Percussion samples included a Native North American drumming pattern and a Brazilian samba percussion pattern. All-in-all, there were 36 music segments. The segments can all be heard at <https://soundcloud.com/janicewang09/sets/sound-of-spiciness-online-test>.

5.2.1.3 Design and procedure

The experiment was programmed on the Qualtrics online survey platform. Before the study began, the participants had to correctly answer an acoustically-presented question to ensure sound playback was functional and to allow them to adjust the volume to a comfortable listening level.

To take part, the participants answered 13 questions, one for each of the musical attributes tested. For each question, the participants had to choose the sound clip that best matched spicy foods (see Figure 5.1 for question format). In the test itself, it was further clarified that spicy meant piquant foods like chili pepper or hot sauce. For the attributes of articulation, distortion, tempo, pitch height, complexity, length of attack, length of decay, length of attack and decay, and ambient music, there were 3 sound clips to choose from. For the attributes of modality, classical music, and percussion music, there were 2 sound clips to choose from. Each sound clip was labelled with a random 3-digit number. For each question, the order in which the sound clips were presented was randomised.

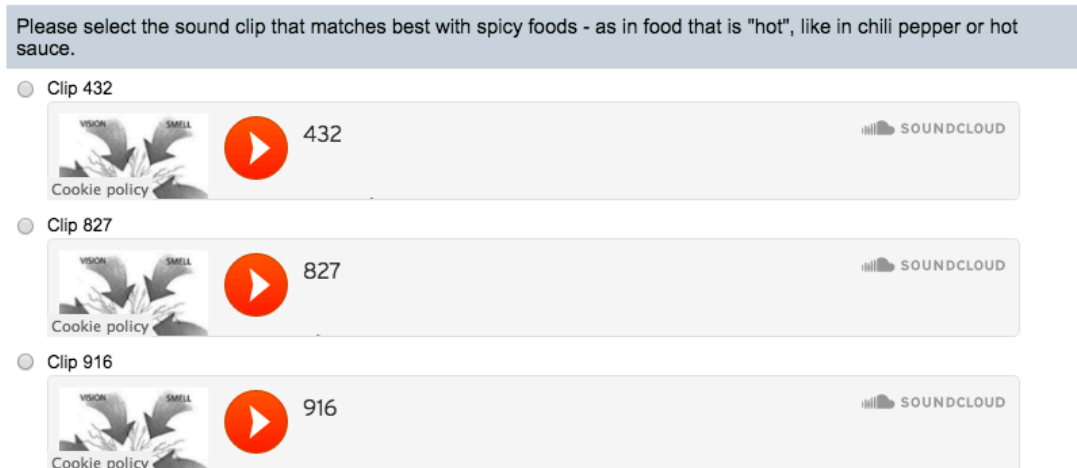


Figure 5.1. Sample question from Experiment 7A.

After the trials, the participants were asked how many years of musical training they had, their enjoyment of spicy food (on a scale from 1 - hate it, to 7 – love it), and the frequency with which they ate spicy food (never, less than once a month, once a month, 2-3 times a month, once a week, 2-3 times a week, daily).

The study lasted for approximately 15 minutes.

5.2.2 Results and discussion

A chi-square test of goodness of fit was calculated for each musical attribute to determine which of them induced a distribution of spicy matches that was significantly different from chance (see Table 5.1). The auditory features with non-random distributions were distortion ($X^2(2,44)=16.41$, $p<.0005$), tempo ($X^2(2,44)=13.68$, $p=.0011$), and pitch ($X^2(2,44)=11.23$, $p=.0036$), where the sound clip with the most distortion, fastest tempo, and highest pitch, respectively, was most often chosen as the one best matching spiciness. Furthermore, the two faster tempo ambient music clips, the classical music clip featuring gypsy violins, and the percussion clip featuring samba drums, were matched more often with spiciness than the slower tempo pieces.

				X ²	p value
	<i>Legato</i>	<i>Medium</i>	<i>Staccato</i>		
Articulation	12	11	21	4.14	.13
	<i>Least distorted</i>	<i>Middle</i>	<i>Most distorted</i>		
Distortion	6	11	27	16.41	.00030
	<i>Slowest tempo</i>	<i>Middle</i>	<i>Fastest tempo</i>		
Tempo	7	11	26	13.68	.0011
	<i>Lowest pitch</i>	<i>Middle</i>	<i>Highest pitch</i>		
Pitch	8	11	25	11.23	.0036
	<i>Minor</i>	<i>Major</i>			
Mode	17	27		2.27	.13
	<i>Least complex</i>	<i>Middle</i>	<i>Most complex</i>		
Complexity	15	10	19	2.77	.25
	<i>Shortest</i>	<i>Middle</i>	<i>Longest</i>		
Attack	16	13	15	0.32	.85
	<i>Shortest</i>	<i>Middle</i>	<i>Longest</i>		
Decay	14	17	13	0.59	.74

	<i>Shortest</i>	<i>Middle</i>	<i>Longest</i>		
Attack + Decay	21	8	15	5.77	.056
	<i>Most consonant</i>	<i>Middle</i>	<i>Most dissonant</i>		
Harmony	17	13	14	0.59	.74
	<i>Energetic</i>	<i>Relaxing</i>	<i>Nervous</i>		
Ambient music	18	3	23	14.77	.00060
	<i>Mozart violin concerto</i>	<i>Saint Saens violin concerto</i>			
Classical music	5	39		26.27	<.0005
	<i>Native American drumming</i>	<i>Samba drumming</i>			
Percussion music	4	40		29.46	<.0005

Table 5.1. Results of participants' soundtrack-spiciness matches in Experiment 7A, for all of the auditory parameters. The chart reveals the distribution of participants' choice of soundtrack that best evoked spiciness for each auditory parameter, with a total of 44 responses for each parameter/question. Features with non-random distributions of answers, according to a chi-square test of goodness of fit and corrected by Bonferroni correction ($\alpha=.0038$), are shown in bold.

For each soundtrack, a chi-square test of independence was performed to determine whether there was an association between spicy liking and choice of spicy sounds (see Table 5.2). The participants were split into two approximately equally-sized groups according to their spicy liking ratings. As the median rating was 6, all those who rated spicy liking as 6 out of 7 or below were placed into one group (N=24), while all those who rated their spicy liking as 7 out of 7 (N=20) were put into another group. When significance values were Bonferroni corrected to avoid inflated Type 1 error, there were no significant differences due to spicy liking.

					X ²	p value
Articulation	Liking < 7	8	4	12	2.24	.33
	Liking = 7	4	7	9		
Distortion	Liking < 7	2	10	12	8.07	.018
	Liking = 7	4	1	15		
Tempo	Liking < 7	3	8	13	2.07	.36
	Liking = 7	4	3	13		
Pitch	Liking < 7	4	8	12	1.97	.37
	Liking = 7	4	3	13		
Mode	Liking < 7	13	11		5.37	.020
	Liking = 7	4	16			
Complexity	Liking < 7	7	3	14	5.61	.060
	Liking = 7	8	7	5		

Attack	Liking < 7	9	4	11	5.12	.077
	Liking = 7	7	9	4		
Decay	Liking < 7	8	10	6	0.53	.77
	Liking = 7	6	7	7		
Attack + Decay	Liking < 7	12	4	8	0.13	.94
	Liking = 7	9	4	7		
Harmony	Liking < 7	8	6	10	2.36	.31
	Liking = 7	9	7	4		
Ambient music	Liking < 7	9	1	14	1.07	.59
	Liking = 7	9	2	9		
Classical music	Liking < 7	3	21		0.068	.79
	Liking = 7	2	18			
Percussion music	Liking < 7	1	23		0.58	.44
	Liking = 7	2	18			

Table 5.2. Results of participants' soundtrack-spiciness matches in Experiment 7A, divided into those who like spicy foods on a scale of 6 or less out of 7 (N=24) and those who like spicy foods on a scale of 7 out of 7 (N=20). Chi squared tests of independence revealed that there were no auditory features where participants had different choices based on spicy liking (alpha level.0038, corrected for multiple tests).

Overall, the musical features that were associated most strongly with spiciness were high pitch, fast tempo, and a distorted timbre. The cultural musical selections were all fast-paced as well. Since these are all traits that are associated with high arousal,

perhaps the association comes from the sensation of consuming spicy food, which includes elevated metabolism (Watanabe et al., 1987) and increased energy expenditure (Janssens et al., 2013).

The results of Experiment 7A were used to compose a “spicy soundtrack”. If people consistently associate musical features in the soundtrack with spiciness, might the soundtrack also enhance the perceived spiciness of real foods? As in Levitan and Shermer’s (2014) study, mentioned earlier, the spicy soundtrack was expected to enhance participants’ expectations of spiciness and, along with it, their actual perception of the taste of the food itself (also see Tu et al., 2016, Experiment 3, for evidence of the mediating role of spiciness expectations on the influence of plate colour on perceived spiciness). In Experiment 7B, I set out to test this hypothesis empirically.

5.3 Experiment 7B: Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack

5.3.1 Methods

5.3.1.1 Participants

180 participants (103 women, 77 men) aged between 18-81 years ($M=41.4$, $SD=11.9$) took part in the study. The participants gave their informed consent, and reported no hearing impairments. The participants were recruited from mailing lists. The experiment was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

5.3.1.2 Auditory stimuli

A spicy soundtrack was composed by iV Audio Branding based on the attributes found to be significantly associated with spiciness from Experiment 7A. A sweet soundtrack was also composed by iV, with high pitch, legato articulation, and consonant harmony (Crisinel & Spence, 2010b; Knöferle & Spence, 2012; Mesz et al., 2011). The sweet and spicy soundtracks were validated in a separate online test (N=110) where the participants listened to each individual soundtrack and were asked to rate, on a scale of 0-10, how much it matched with sweet, spicy, sour, and salty tastes (see Figure 5.2 for average ratings). Both the spicy and sweet soundtracks were matched significantly more frequently to spicy and sweet tastes, respectively, than all other options ($p < .0005$ for all comparisons).

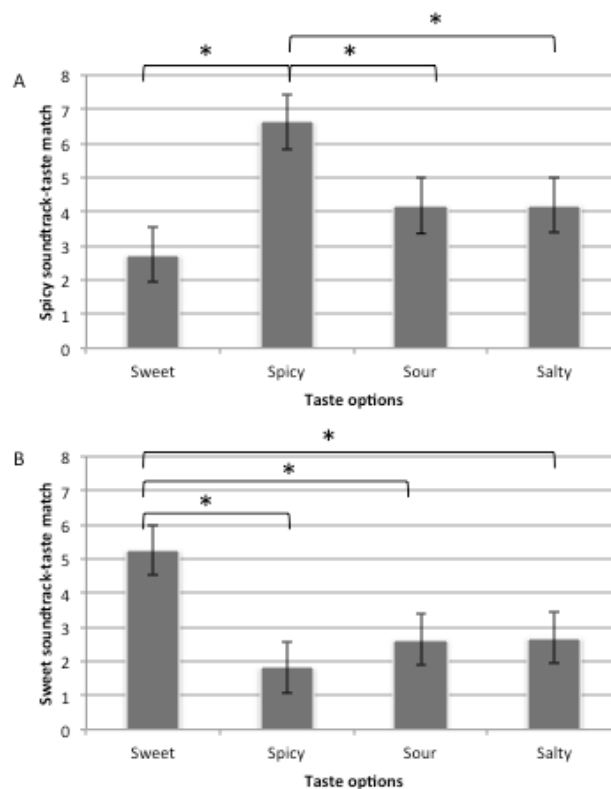


Figure 5.2. Results of participants' ratings of sound-taste matching (on a scale from 0-10) in Experiment 7B for A) the spicy soundtrack and B) the sweet soundtrack composed by iV Audio

Branding. Error bars indicate the standard error of the means. * indicate statistical significant comparisons ($p < .05$).

A 10-minute segment of white noise was produced as well. All sounds can be streamed at <http://ivaudiobranding.com/spicysounds/>.

5.3.1.3 *Gustatory stimuli*

A dish was created by chef Deb Paquette of Etch restaurant in Nashville, Tennessee, USA that contained both sweet and spicy components (see Figure 5.3). The dish consisted of a piece of ancho chili dusted butternut squash tempura with brie, pear, mushrooms and salad leaves, topped with a sour-spicy dressing, spicy pepitas, and spiced chocolate sauce.



Figure 5.3. The dish served at Etch restaurant during Experiment 7B. The dish itself is a novel creation that is likely to be unfamiliar to diners.

5.3.1.4 Design and procedure

The experiment was conducted at Etch, a restaurant in Nashville, Tennessee, USA (<http://www.etchrestaurant.com/>). The study was repeated four times, at 11AM, 12PM, 1PM, and 2PM. The participants were seated at tables of four, and there were three tables of each sound condition (spicy, sweet, white noise, and silence). Participants at tables with sounds were provided with earbud-style headphones while those in the silent condition table were given earplugs. Soundtracks were played from headphones at approximately 80 dB(A).

Before the actual study began, the participants specified their gender, age, years of musical training, and liking for spiciness (from 0 – hate it, to 10 – love it). First, all participants tasted a pre-prepared sauce and rated its sweetness and spiciness on scales of 0 (no taste) to 10 (most intense imaginable). Afterwards, they were instructed to rinse their mouths out with water.

Next, the dish was served and the participants were instructed to look at the dish while listening to the soundtrack (or just looking, for those in the silent condition) for 30 seconds. The participants then rated how they expected the dish to taste, in terms of its sweetness, spiciness, flavour intensity, as well as how much they expected to like the dish. (All ratings on a scale from 0-10.)

Finally, the participants were instructed to listen to the soundtrack or put on earplugs while eating the dish. After 5 minutes, the participants once again rated the dish on the same four scales, this time based on their actual experience.

The study lasted for approximately 30 minutes.

5.3.2 Results and discussion

The average expected ratings for the dish are shown in Figure 5.4. A MANOVA with spiciness, sweetness, flavour intensity, and liking ratings as measures – and with sound condition (white noise, silence, sweet, spicy) as the factor – revealed a significant main effect of sound condition ($F(12,522)=2.02$, $p=0.02$, Pillai's Trace=0.13). Follow-up ANOVAs revealed the effect of sound conditions on expected spiciness ($F(3,175)=7.90$, $\eta^2=0.12$, $p<.0005$), but not on expected sweetness ($F(3,175)<1$, n.s.), flavour intensity ($F(3,175)=1.28$, $p=0.28$), or liking ($F(3,175)<1$, n.s.). Pairwise comparisons with Bonferroni corrections revealed that expected spiciness ratings were higher in the spicy soundtrack condition than any other sound condition ($p\leq.001$ for all comparisons).

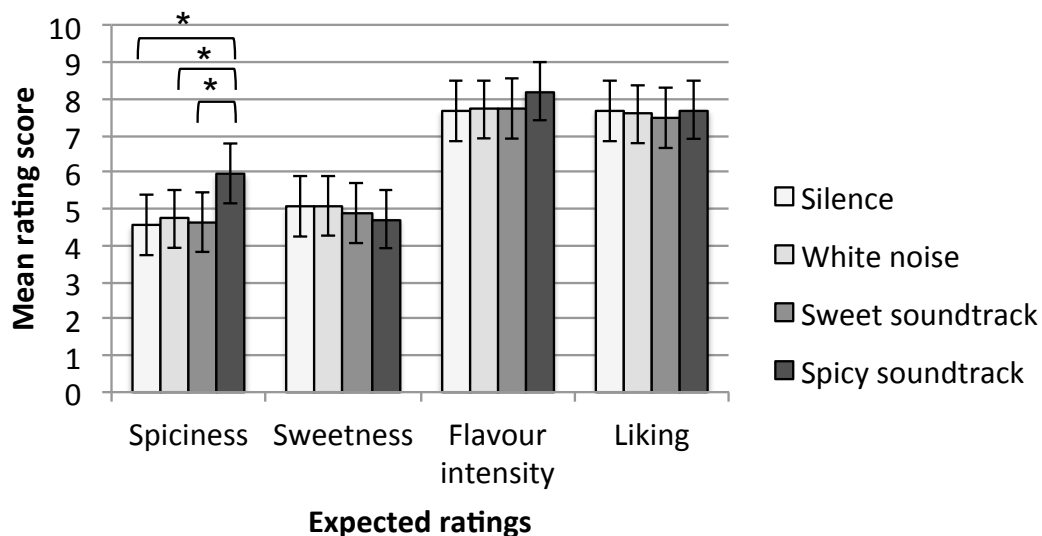


Figure 5.4. Participants' mean ratings of the expected taste of the dish in Experiment 7B, in all four sound conditions. Higher ratings scores reflect higher perceived levels of spiciness and sweetness, as well as greater flavour intensity and liking. Error bars represent the standard error of the means. Asterisks denote statistical significance (* $p<.05$).

The average actual ratings of the dish are shown in Figure 5.5. A similar MANOVA test as above was conducted with actual ratings in place of expected ratings, but the sound condition did not have a significant effect ($F(12, 522) < 1$, n.s.).

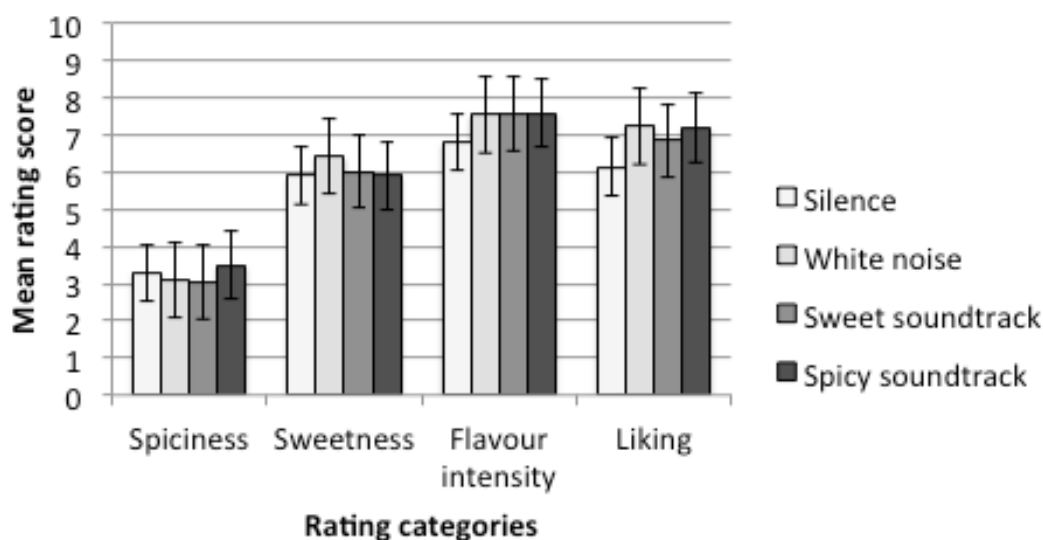


Figure 5.5. Results of participants' ratings of the actual taste of the dish in Experiment 7B, in all four sound conditions. Higher scores reflect higher perceived levels of spiciness and sweetness, as well as greater flavour intensity and liking. Error bars represent the standard error of the means.

Pearson correlations showed that liking for the dish was correlated with different taste ratings of the dish, depending on the sound condition. Food liking was positively correlated with spiciness in the spicy soundtrack condition ($r=0.33$, $p=.025$) but not with sweetness. Inversely, in the sweet soundtrack condition, food liking was positively correlated with sweetness ($r=0.31$, $p=.033$) but not with spiciness.

Age was positively correlated with spiciness ($r=0.28$, $p<.0005$) and negatively correlated with sweetness ratings ($r=-0.26$, $p<.0005$). The negative correlation with sweetness ratings is in line with findings that taste perception is dulled with age (Bartoshuk et al., 1986). It is possible that sensitivity to spiciness, however, increases with age, or rather that, compared to younger people, the older generation of North Americans may not be as used to eating spicy food. Certainly, the popularization of

ethnic restaurants, such as Thai, Mexican, and Indian, in the United States, began in the 1980s (Gvion & Trostler, 2008).

On average, liking for the sound conditions (on a scale of 1-10) were 5.62 (SD=2.52) for the spicy soundtrack, 5.83 (SD=2.07) for the sweet soundtrack, and 2.38 (SD=2.07) for white noise. A univariate ANOVA revealed a significant difference in liking for the soundtrack ($F(2,137)=33.83$, $p<.0005$), with the white noise liked significantly less than either of the taste soundtracks ($p<.0005$). However, there were no significant correlations between soundtrack liking and any of the food-related ratings.

The results of Experiment 7B revealed that the spicy soundtrack increased the expected spiciness of the dish significantly as compared to all of the other sound conditions (including silence), but there were no similar effects when it came to actual spiciness ratings. One possible explanation for this result is that the dish used in Experiment 7B was simply not spicy enough. Intriguingly, in Shermer and Levitan's (2014) study, the effect of colour on salsa was only evident in the spicier version (an average of 4.4 on a 7 point scale); there was no effect of colour on the piquancy of the milder salsa (an average of 3.0 on a 7 point scale). More importantly, both Shermer and Levitan, and Woods et al. (2010, 2011) demonstrated that people's differing expectations do not necessarily translate into an actual change in perception/ratings if people's expectations turn out to be too far from reality³ (although Cardello & Sawyer (1992, Experiment 2) do report an example where positive disconfirmation about a pomegranate juice's bitterness levels resulted in lower bitterness ratings as compared

³ As observed by Anderson (1973) and Olson and Dover (1976), contrast effects, where large differences between expectations and reality result in ratings in the opposite direction from what's expected (and from what the ratings might have been without any prior expectations), may only occur infrequently.

to a control condition⁴). In Experiment 7B, the spicy soundtrack led to expectations of spiciness (average 6 out of 10), but the actual level of spiciness (3 out of 10) was incongruent from what was expected, and so the resulting percept was not altered in favour of the expectations.

Multiple models have been put forward to account for the effect of the disconfirmation of consumer expectation. The theory most used by food science researchers is the theory of assimilation/contrast (Cardello, 2007; Deliza & MacFie, 1996; Piqueras-Fiszman & Spence, 2015), which predicts that consumers are more likely to adjust their perception of the product in the direction of their expectations for small discrepancies but are more likely to shift their ratings in the opposite direction from their expectations for large differences. It follows that perhaps, had the dish been spicier and closer to the expected level of heat (on average 6 out of 10), a difference between the spicy soundtrack and other sound conditions would have been observed. To validate whether this model can explain the results from Experiment 7B, Experiment 7C was conducted using medium spicy salsa as the taste stimuli.

In addition, the soundtracks might have acted on participants' ratings of the dish indirectly, as a result of the soundtracks' different degrees of pleasantness. In other words, more pleasant music might have made the dish appear more pleasant (see Cheskin, 1972, for just such a theory of sensation transference). If this were to be the case, then one would expect to see positive correlations between music pleasantness and dish pleasantness ratings. There were not, however, any significant correlations between music pleasantness and expected ($r_{137}=0.07$, $p=0.41$) or actual ($r_{137}=0.11$, $p=0.19$) liking of the dish. Therefore, the results suggest that some mechanism other

⁴ It is possible, even though highly unlikely, that the contrast effect could have been due to an imbalance in the distribution of supertasters in the different between-participant groups.

than sensation transference is at work in terms of the soundtracks' influence on participants' ratings.

5.4 Experiment 7C: Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack

5.4.1 Methods

5.4.1.1 Participants

18 participants (9 women, 9 men) aged between 21-61 years ($M=35.33$, $SD=9.91$) took part in the study. The participants gave their informed consent, and reported no hearing impairments. The participants were recruited from mailing lists and social media. The experiment was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

5.4.1.2 Auditory stimuli

The same sound stimuli were used as in Experiment 7B (spicy soundtrack, sweet soundtrack, white noise, and silence). 60-second excerpts were made by taking the first 60 seconds of audio from the sound stimuli.

5.4.1.3 Gustatory stimuli

A medium spicy salsa mixture was created by combining approximately 700g of Red Gold brand mild salsa and 1500g of Krogers store brand medium spicy salsa, then adding 25 drops of PureCap capsaicin extract. The capsaicin extract was added incrementally until the spiciness level was judged by a small pilot study group ($N=3$) to be around 5/10. The salsa mixture was blended to ensure a smooth and uniform

texture. The salsa was served in approximately 10 mL portions in 89 mL (3 oz) opaque white cups with a spoon for each sample.

To help verify the theory that sound can affect taste ratings only when the expected and actual tastes are similar, an online control study (N=110) was conducted in which the participants evaluated the expected spiciness, intensity, and liking of an image of a bowl of salsa while listening to the same spicy soundtrack used in Experiment 7B. The mean expected ratings of the salsa image was, for spiciness (M=6.96, SD=2.35), flavour intensity (M=6.88, SD=1.88), and liking (M=5.88, SD=2.62).

5.4.1.4 Design and procedure

The experiment was conducted at the Curb Center at Vanderbilt University in Nashville, Tennessee, USA. The participants were seated at tables with ear-bud headphones, a paper questionnaire, four salsa samples (arranged in a line from top to bottom), a cup of water, and saltine crackers.

Before the actual study began, the participants specified their gender, age, and liking for spiciness (from 0 – hate it, to 10 – love it). Next, the participants were instructed to taste the four samples, one at a time, in order from top to bottom, when cued by the experimenter. Each sample was tasted while listening to one of four sound conditions (played at approximately 80 dB(A), for the non-silent conditions.) In order to make sure that the participants fully appreciated the effect of the salsa's heat, they were asked to wait at least 10 seconds after tasting before making their ratings of the salsa's flavour intensity, pleasantness, and spiciness (all on scales of 0-10). Each trial lasted for 60 seconds, and participants were given a 60 second break in between trials to cleanse their palates with water and crackers. The order of the soundtracks was

determined using a Latin square design; participants were divided into four groups, where each group heard the soundtrack in a specific order.

The study lasted for approximately 15 minutes. The participants were debriefed afterwards in a presentation.

5.4.2 Results and discussion

The average ratings for the salsa under the different sound conditions are shown in Figure 5.6. Repeated-measured ANOVAs were conducted on participants' spiciness, flavour intensity, and liking ratings with sound condition as the factors. Sound condition had a significant effect on spiciness ratings ($F(5,31)=9.51$, $p<.0005$, $\eta^2=0.36$) and flavour intensity ratings ($F(3,51)=5.04$, $p=.004$, $\eta^2=0.23$). Pairwise comparisons with Bonferroni corrections revealed that spiciness ratings were significantly higher during the spicy soundtrack condition ($M=6.89$, $SD=1.18$) than in the sweet soundtrack condition ($M=4.94$, $SD=1.66$, $p=.002$), white noise condition ($M=5.56$, $SD=1.42$, $p=.016$), and silent condition ($M=4.94$, $SD=2.01$, $p=.001$). In addition, flavour intensity ratings were significantly higher during the spicy soundtrack condition ($M=7.00$, $SD=0.84$) than in the silent condition ($M=5.72$, $SD=1.57$, $p=.021$).

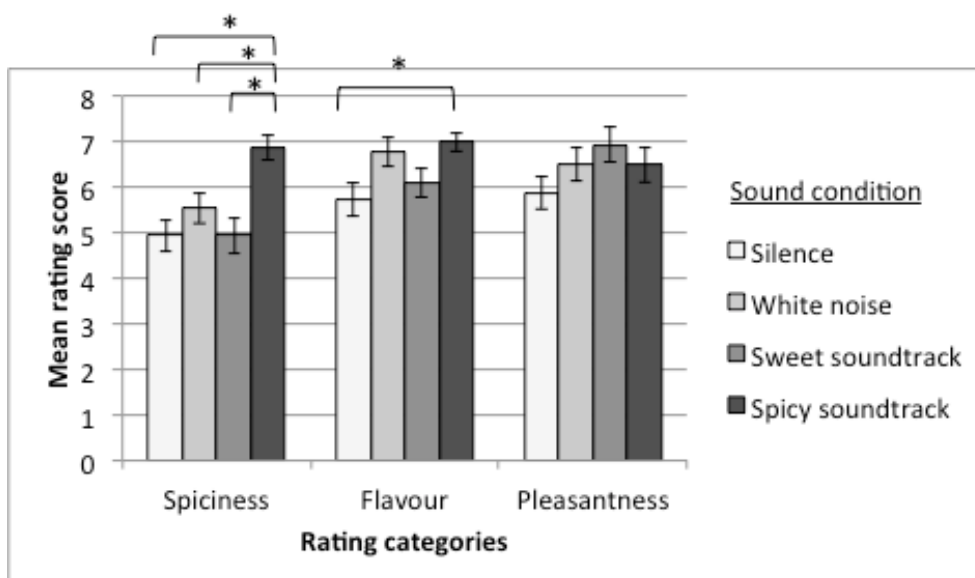


Figure 5.6. Results of participants' ratings of the taste of salsa in Experiment 7C, in all four sound conditions. Error bars represent the standard error of the means. Asterisks denote statistical significance ($p < .05$).

The results of Experiment 7C revealed that the spiciness ratings of the salsa samples were significantly higher while the participants listened to the spicy soundtrack than to any of the other sound conditions. This result demonstrates that the perception of spiciness/piquancy in a food can be enhanced simply by playing a crossmodally corresponding soundtrack, similar to the results for bitter/sweet tastes first documented by Crisinel et al. (2012). One significant difference between Experiments 7B and 7C was that the taste stimuli were spicier in the latter experiment, with a mean rating of 5.58 across all sound conditions (as compared to 3.23 in Experiment 7B). More importantly, the mean spiciness rating under the spicy soundtrack condition was 6.89, similar to the expected spiciness from the control study ($M=6.96$), when participants evaluated an image of salsa while listening to the spicy soundtrack. This provides support for the hypothesis that the spicy soundtrack only modifies taste evaluations if the participants' expectations are similar enough to reality.

Interestingly, there were not any similar enhancement effects of the spicy soundtrack in terms of food pleasantness, even though the majority of participants reported that they enjoyed eating spicy foods ($M=7.83$ out of 10, $SD=2.48$).

One issue to bear in mind with the interpretation of Experiment 7C is that there was not a balanced Latin Square design, and that, with 18 participants and 4 soundtrack sequences, there were not equal numbers of participants for each sequence. Given the relatively small number of participants ($N=18$) and the uneven balancing, it is possible that these results could have been subject to carryover effects of capsaicin.

In order to confirm the theory that the spicy soundtrack influences spiciness perception via enhancing participants' expectations, experiment 7D was conducted using salsa of two spiciness levels – mild and hot. Furthermore, in order to account for the carryover effects of capsaicin (Green 1991), a Williams design Latin square was used in the study design.

5.5 Experiment 7D: Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack

5.5.1 Methods

5.5.1.1 Participants

40 participants (24 women, 16 men) aged between 21-49 years ($M=22.58$, $SD=6.53$) took part in the study. The participants gave their informed consent, and reported no hearing impairments. The participants were recruited from the Oxford Psychology Research Participant Database and the Experimental Psychology Research

Participation Scheme. The experiment was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

5.5.1.2 Auditory stimuli

Since Experiment 7C showed that only the spicy soundtrack enhanced perceived spiciness compared to all of the other sound conditions, the only auditory stimuli in Experiment 7D were the spicy soundtrack and silence (as a control). The 60 second spicy soundtrack was the same as in Experiment 7C. For the silent condition, a soundtrack was made with the words “start” and “stop” spoken at the beginning and the end, separated by 60 seconds of silence in between.

5.5.1.3 Gustatory stimuli

Salsa samples of two different heat levels were used, Doritos Mild Salsa and Doritos Hot Salsa. The same brand was used to maximise consistency in terms of taste between the two different levels of spiciness. The salsa samples were blended to ensure a smooth and uniform texture. The salsa was served in approximately 10 mL portions in clear 50 mL plastic cups with a small white plastic spoon for each sample.

5.5.1.4 Design and procedure

The experiment was conducted at the Crossmodal Research Laboratory at the University of Oxford. Participants were seated at a table in front of a computer monitor with a keyboard, mouse, and headphones in an experimental booth. On the side table were four salsa samples (each labelled with a random 3 digit code) with spoons, a cup of water, and saltine crackers.

For each trial, the participants were instructed to taste the sample whose number was shown on the computer screen, accompanied by a soundtrack (the spicy soundtrack

was presented at approximately 80 dB(A)). The participants were instructed to taste the salsa and listen to the soundtrack at the same time. Only after the soundtrack had finished could they evaluate the salsa in terms of its spiciness and flavour intensity (all on scales of 0-10). Each trial lasted for 60 seconds, and the participants were given a two-minute break between trials to cleanse their palates with water and crackers. The order of the soundtrack/salsa sample combination was determined using a Latin square design; participants were divided into four groups, where each group heard the soundtrack in a specific order. Overall, each participant heard each soundtrack (spicy and silent) twice and tasted each salsa sample (mild and hot) twice. Finally, participants specified their liking for the spicy soundtrack and for spicy foods in general (on 1-7 scales), as well as detailing their years of musical training.

The study lasted for approximately 20 minutes. Participants were paid £4 or else awarded course credit.

5.5.2 Results and discussion

The average ratings for the salsa with different heat levels and under different sound conditions are shown in Figure 5.7. Repeated-measures analysis-of-covariant tests were conducted on participants' spiciness and flavour intensity ratings, with sound condition and salsa heat level as within-participant factors. The covariates⁵ used were liking for the spicy soundtrack, liking for spicy foods, and years of musical training.

Analyses revealed a significant interaction effect between sound condition and heat level ($F(2,35)=6.14$, $p=.005$, Wilks' Lambda.74). The interaction had a significant effect on salsa spiciness ratings ($F(1,36)=17.57$, $p=.001$, $\eta^2=.26$), but not on flavour

⁵ The Delaney-Maxwell method of mean-centering the covariate values was applied to prevent the covariates from altering the main effects of the repeated measures factors (Delaney & Maxwell, 1981).

intensity ($F(1,36) < 1$, n.s.). More specifically, the participant rated the hot salsa as significantly more spicy when listening to the spicy soundtrack as compared to silence ($p = .019$); in contrast, for the mild salsa, there was a borderline-significant effect where the salsa was rated as being less spicy when listening to the spicy soundtrack as compared to silence ($p = .052$).

There was also a significant main effect of salsa heat level ($F(2,35) = 142.75$, $p < .0005$, Wilks' Lambda = .11) Further univariate ANOVA tests showed that salsa heat level had a significant main effect on both salsa spiciness ($F(1,36) = 281.15$, $p < .0005$, $\eta^2 = .89$) and flavour intensity ($F(1,36) = 6.19$, $p = .018$, $\eta^2 = .15$), where the Doritos Hot Salsa was rated as both more spicy ($p < .0005$) and more intense in flavour ($p = .018$) than the Doritos Mild Salsa (all p-values in post hoc comparison tests have been Bonferroni corrected).

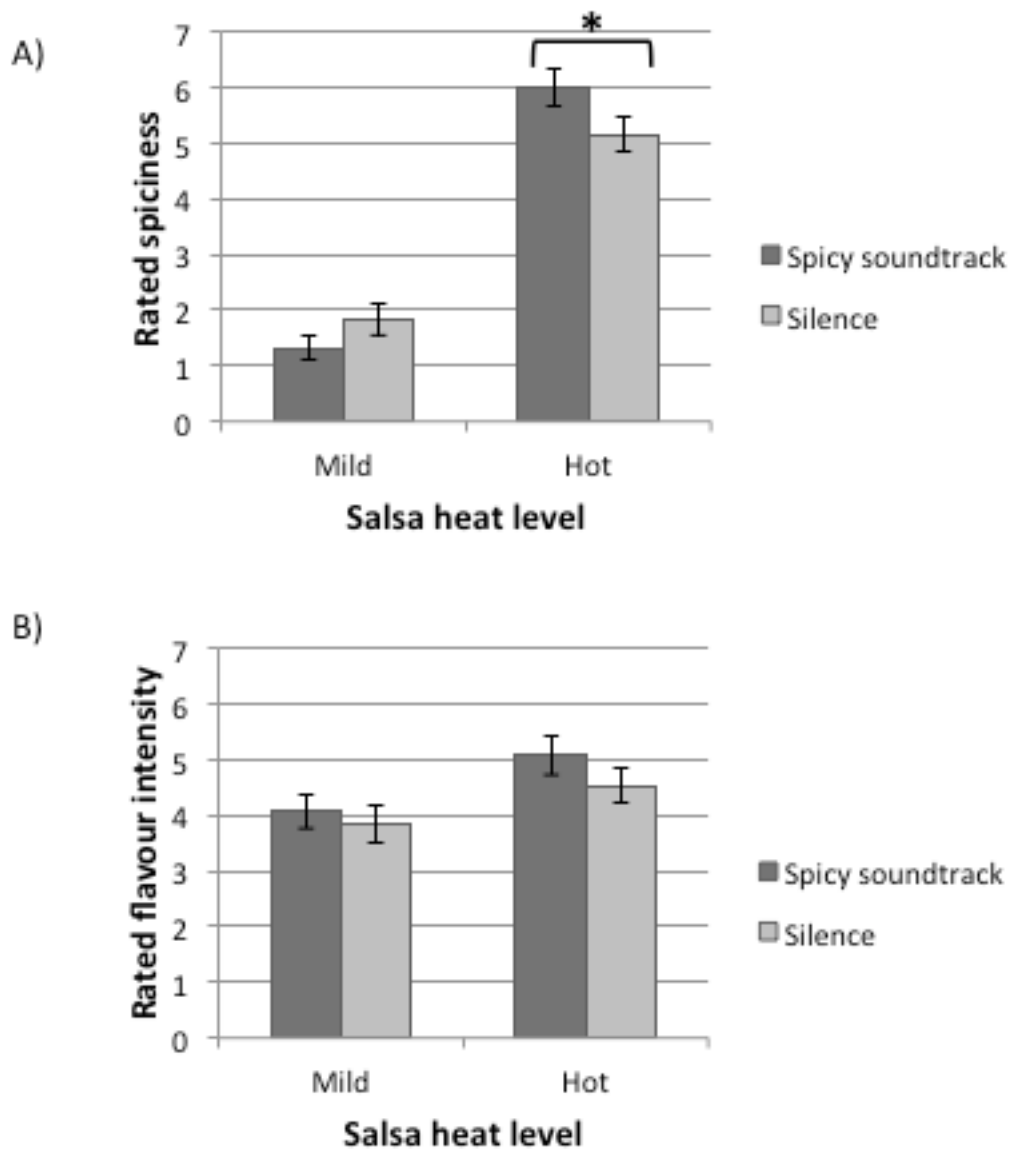


Figure 5.7. Participants' ratings of the spiciness (A) and flavour intensity (B) of salsa in Experiment 7D, for two salsa heat levels and two sound conditions (spicy soundtrack and silence). Error bars represent the standard error of the means. Asterisks (*) denote statistical significance ($p < .05$).

The results of Experiment 7D revealed that ratings of the spiciness of the salsa samples were significantly higher while listening to the spicy soundtrack as compared to silence, but only for the hot salsa sample. In fact, an opposite trend was seen with

the mild salsa sample, where the participants rated the salsa as being less spicy while listening to the spicy soundtrack as compared to silence.

The effect of sound on spiciness ratings can thus be explained in terms of an assimilation-contrast model (Cardello 2007; Deliza & MacFie, 1996). The sound cue given by the spicy soundtrack leads to an expectation about piquancy, and assimilation occurs if the actual experienced piquancy seems to be more or less in-line with what was expected (e.g., Tuorila et al., 1994). In this case, assimilation could account for the fact that the hot salsa rated to be spicier during the spicy soundtrack condition compared to the silent condition. When the expectation set by the sound stimuli is incongruent with the experienced piquancy however, contrast may occur with the resulting percept being biased in the direction opposite to expectations (Cardello & Sawyer, 1992). Contrast could also explain why the mild salsa was rated as being less spicy during the spicy soundtrack condition as compared to the silent condition.

Under this assimilation-contrast theory, one can imagine that those individuals who are more sensitive to spiciness might be more prone to the effect of the soundtrack even for mildly spicy foods, if they perceive that even a low level of spice is quite high. No spiciness sensitivity information was collected in Experiment 7D, although it is perhaps plausible to assume that anyone who would volunteer for such a study involving tasting potentially spicy food would be self-selective towards those who can tolerate and enjoy eating spicy foods. In fact, 29 out of 40 participants reported liking spicy foods (giving a rating of 5 or higher on a 7-point scale). An interesting future study could be to recruit a group of participants with high spiciness sensitivity and assess the effect of the spicy soundtrack on mildly spicy foods; one would expect to

observe the soundtrack-induced spiciness enhancement effect for these participants, but not for average spiciness sensitivity participants.

An alternative interpretation for the effect observed for the mild salsa is that it seemed spicier in silence compared to the spicy soundtrack because it was easier to pick up the subtle spiciness of the salsa in silence. The spicy soundtrack, played at 80 dB(A) (around the same loudness as a vacuum cleaner), might have distracted the participant from paying attention to the mild spiciness of the salsa.

Next, Experiment 8 follows-up with the findings of Experiment 7 by directly examining the role of soundtrack timing on auditory taste modulation, especially in the case where the soundtrack is heard only before food is consumed.

5.6 Experiment 8: Assessing the role of expectations on auditory taste modulation of chocolate

5.6.1 Methods

5.6.1.1 *Participants*

A total of 56 participants (44 women, 12 men) aged 22-57 years ($M=36.48$, $SD=9.67$) took part in the study. The participants were recruited at the Food Matters Live conference at the London ExCeL center, November 22, 2016. All of the participants gave their informed consent to take part in the study. None of the participants reported a cold nor any other known impairment of their sense of smell, taste, or hearing at the time of the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

5.6.1.2 *Auditory stimuli*

I used the same sweet and bitter soundtracks as in Experiment 6.

5.6.1.3 Gustatory stimuli

70% Lindt chocolate was used for the study for its fairly complex taste, ambiguous sweet-bitter balance, and commercial availability. Each sample consisted of approximately 3g of chocolate. All samples were served on a white Styrofoam plate with 3-digit labels.

5.6.1.4 Design and procedure

Experiment 8 was designed with timing condition (sound before or during tasting) as a between-participant factor and soundtrack type (bitter or sweet) as a within-participants factor. Each participant was randomly assigned to one timing condition, completed four trials, and heard each soundtrack twice.

Each participant was seated in front of a computer screen with headphones, a plate with chocolate samples (labelled with 3-digit identifiers 719, 420, 536, and 148), and a cup of water to cleanse their palates. The experiment was programmed on the Qualtrics online survey platform and participants responded by using the mouse to click or drag the indicator on the continuous rating scales. Participants were randomly assigned to either the before-tasting (N=29) or during-tasting condition (N=27).

Participants in the before-tasting condition first looked at the target chocolate sample for 30 seconds while listening to either the sweet or bitter soundtrack through headphones. Next, they tasted the chocolate sample in silence for 30 seconds. Participants in the during-tasting condition were first asked to look at the chocolate samples for 30 seconds in silence. Next, they tasted the chocolate sample for 30 seconds while listening to either the sweet or bitter soundtrack. For both groups, participants rinsed their mouths out with water before moving onto the evaluation page, where they rated the chocolate on an 11-point bitter-sweet scale (0=much more

bitter than sweet, 10=much more sweet than bitter) and a 11-point pleasantness scale (0=very unpleasant, 10=very pleasant). Participants cleansed their palates right after tasting to ensure that the soundtracks during the evaluation stage did not act on any residual tastes in the mouth.

The order of the soundtracks and the order of the rating scales were randomised.

Finally, the participants listened to each soundtrack again. For each soundtrack, they rated it on 7-point bipolar scales of valence (unpleasant to pleasant) and arousal (calm to exciting). Participants also rated how much each soundtrack matched each of four basic tastes (sweet, bitter, sour, and salty) on a scale of 0-10 (0=not at all matching, 10=very much matching).

The experiment lasted for around 10 minutes and participants were debriefed afterwards.

5.6.2 Results and discussion

Bittersweet and liking ratings were first averaged over the two identical soundtrack trials for each participant. A mixed RM-MANOVA was conducted with ‘soundtrack type’ as within-participants factor and ‘timing’ as the between-participants factor. The model included bittersweet and liking ratings as measures.

The mean values of the participants’ taste and liking ratings for the chocolates in both timing conditions are shown in Figure 5.8. For bitter-sweet ratings, the RM-MANOVA revealed a significant main effect of soundtrack ($F(2,52)=8.43$, $p=.001$, Wilk’s Lambda=.76), but no significant main effect of timing ($F(2,52) < 1$, n.s.), and no significant interaction between timing and soundtrack type ($F(2,52)=2.70$, $p=.08$). More specifically, participants rated the chocolates higher towards the sweet end of the bitter-sweet scale when listening to the sweet soundtrack – either before or during

tasting – as compared to the bitter soundtrack ($M_{\text{sweet_soundtrack}}(\text{SE})=5.18(0.23)$, $M_{\text{bitter_soundtrack}}(\text{SE})=4.02(0.21)$, $p<.0005$). Similarly, the chocolates tasted with the sweet soundtrack either before or during tasting was also liked more than with the bitter soundtrack ($M_{\text{sweet_soundtrack}}(\text{SE})=6.74(0.19)$, $M_{\text{bitter_soundtrack}}(\text{SE})=5.99(0.19)$, $p=.001$).

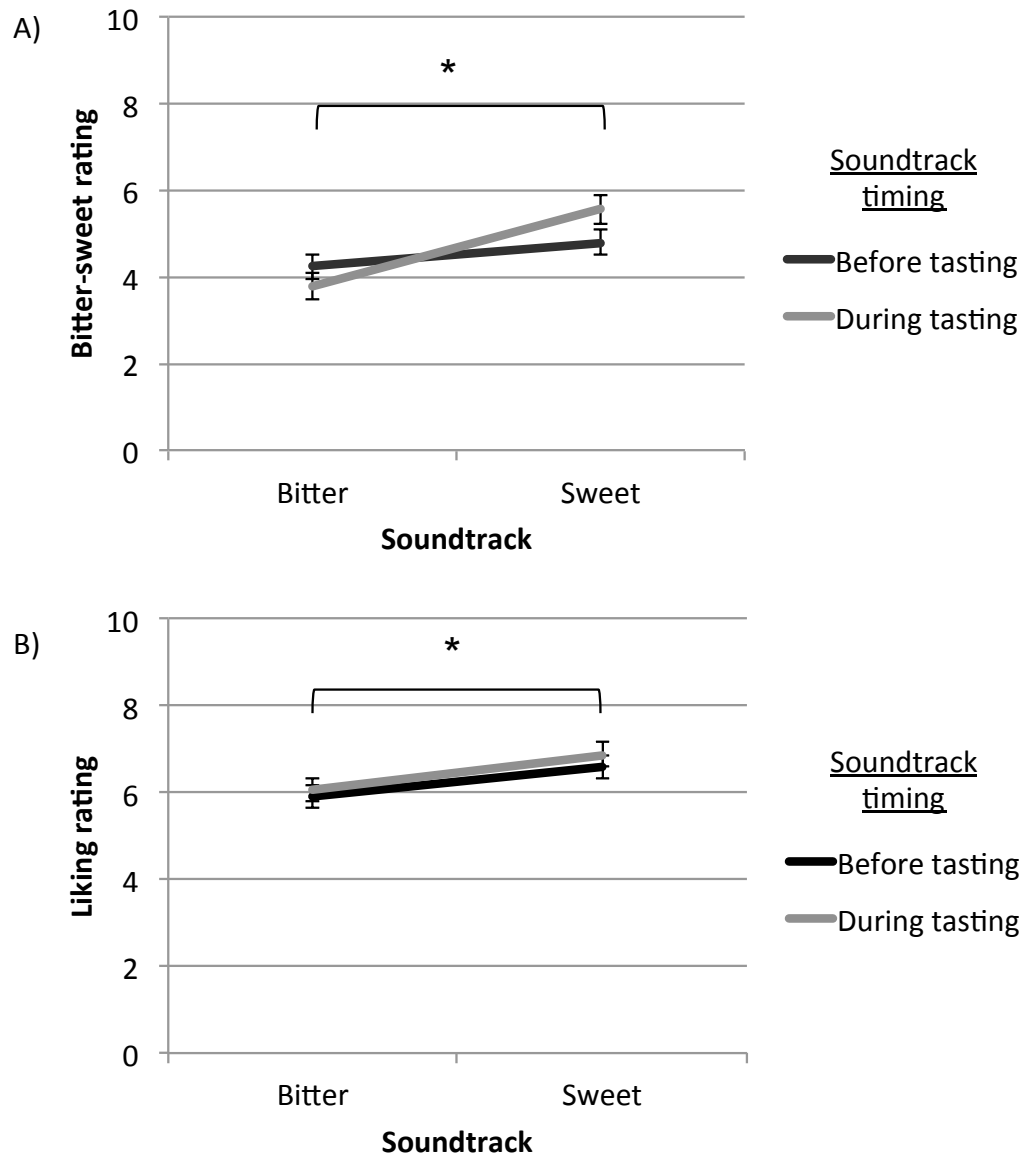


Figure 5.8. Mean values of bitter-sweet (A) and liking (B) ratings for chocolates in Experiment 8, for both auditory stimuli timing conditions. Error bars indicate standard error. Asterisk “*”, indicates statistical significance at $p<.05$.

Assessing the emotion ratings of the soundtracks, participants rated the sweet soundtrack as significantly more pleasant than the bitter soundtrack ($M_{\text{bitter}}(\text{SE})=3.45(0.22)$, $M_{\text{sweet}}(\text{SE})=4.43(0.22)$, $p=.01$). The sweet soundtrack was also rated as more arousing than the bitter soundtrack ($M_{\text{bitter}}(\text{SE})=3.29(0.21)$, $M_{\text{sweet}}(\text{SE})=4.25(0.22)$, $p=.008$). In terms of participants' sound-taste association ratings for each soundtrack (see Figure 5.9), the bitter soundtrack was rated as matching bitterness significantly better than the other tastes ($p<.0005$ for all comparisons), and the sweet soundtrack was rated as matching with sweetness significantly better than with the other tastes ($p<.0005$ for all comparisons).

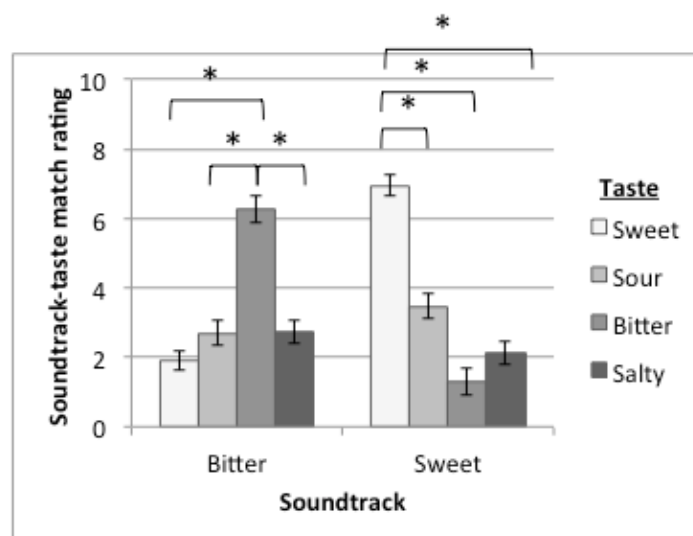


Figure 5.9. Mean values of soundtrack-taste match ratings for the sweet and bitter soundtrack used in Experiment 8. Participants rated how well each soundtrack matched four basic tastes (sweet, sour, bitter, salty) on a scale of 0-10. Error bars indicate standard error. The asterisks “*”, indicates statistical significance at $p<.05$.

In Experiment 8, I did not find any differences in the timing of the auditory stimuli – whether before or during chocolate tasting – on the modulatory effects of crossmodally congruent soundtracks on taste evaluation. There was an overall effect

of soundtrack type, where the sweet soundtrack enhanced the sweetness of chocolates as compared to the bitter soundtrack. This result is in line with results of Experiment 7, where soundtrack was shown to alter participants' expectations about the spiciness levels of foods. In this case, the timing condition where the sweet/bitter soundtracks were heard before the tasting phase presumably influenced the participants' expectations about the chocolate they were about to consume.

Alternatively, the fact that participants' liked the chocolate more when the sweet soundtrack was played – regardless of timing – suggests that the listeners might have merely transferred their feelings about the soundtracks to the taste stimuli. Post-experiment questions showed that the sweet soundtrack was rated to be significantly more pleasant than the bitter soundtrack. Therefore, it is possible that instead of having different expectations about the taste of the chocolate samples, participants shifted their positive or negative feelings about the soundtrack onto the samples. What's more, these feelings might have been translated into differences in taste ratings in a halo dumping effect (Kappes et al., 2006), so that, for instance, the more liked chocolate might have been rated as sweeter whereas the less liked chocolate might have been rated as more bitter. It should be noted, however, that in the typical cases of halo dumping, the effect goes away when participants are asked to rate both taste and liking so that there is no need for participants to “dump” the values of liking onto the bittersweet scale (Clark & Lawless, 1994). More discussions on the role of emotions will be presented in Chapter 7.

5.7 General Discussion and conclusions

Overall, the research reported here demonstrates how crossmodally congruent soundtracks can influence participants' expectations about the food they are about to consume. Experiment 7B demonstrated that participants' expectations were influenced by the spicy soundtrack, which shows that the influence of music begins before participants even taste the food. This agrees with the results of Experiment 8, which showed that sweet/bitter soundtracks could influence taste evaluation regardless of whether the soundtracks are heard before tasting or during tasting. In addition, the results of Experiments 7D are consistent with those findings showing that expectations about the taste of foods can bias participants' perceived ratings towards what is expected, as long as the difference between expectation and the actual experience is small (Shermer & Levitan, 2014; Tu et al., 2016; Woods et al., 2011).

A potential issue with Experiments 7 and 8 is that Experiment 7B was conducted using a between-participants design while Experiments 7C, 7D, and 8 were conducted using a within-participants design. As a result, the latter experiments may have been more prone to demand effects, where participants might have wanted to please the experimenters by giving higher spiciness ratings when the spicy music was playing (although note that the results of Experiment 7D, where only the spicier salsa showed enhanced spiciness effects, contradict such a theory). It is possible that the crossmodal correspondence between sound and spiciness is a relative effect, like many other correspondences have been shown to be, and hence requires explicit comparisons between distinctively different auditory stimuli (Gallace & Spence, 2006).

The studies reported in this chapter have shown that crossmodally congruent soundtracks could give rise to sensory expectations, which might then modify our perception of food and beverages. One way that expectations can shape perception is

drawing people's attention towards specific tastes experienced in the eating process (Deliza & MacFie, 1996; Piqueras-Fiszman & Spence, 2015). In the next chapter, I will examine in more detail the role of attention in crossmodal perception.

6 Attention as a possible mechanism underlying the auditory modulation of taste/flavour

6.1 Introduction

Attention is intrinsic to how we perceive sensory inputs (Chen et al., 2013), and is considered by some psychologists and neuroscientists to be closely linked with consciousness itself (Crick & Koch, 1990; Dehaene et al., 2006; though see Koch & Tsuchiya, 2006, for an argument against conflating the two notions). We are constantly bombarded with vast amounts of sensory information, and an attentional mechanism prioritises limited neural resources towards more relevant information for further processing and storage (Spence, 2014c). In addition, attention can be oriented in at least two ways: either in a bottom-up fashion, where sensory stimuli can automatically shift attention/processing resources independently of higher levels goals; or in a top-down manner, where stimuli that match the observers' goals are consciously selected (Talsma et al., 2010). Attention plays a crucial role in determining what we perceive in food. In his review, Stevenson (2012) illustrates the role of attention in flavour perception, where the unitary experience of flavour is attributed to attentional capture by somatosensation over olfaction. Moreover, extending the spotlight metaphor, attended flavour elements become relatively more salient than relatively less attended elements (Marks & Wheeler, 1998; Ashkenazi & Marks, 2004).

The question here, then, is the role of attention on the way sound-flavour correspondences may influence the eating/drinking experience. In Experiment 9, I focus on bottom-up attention processing to assess whether the effect of auditory

stimuli is automatic, or whether one needs to have their attention drawn to the sound (and its possible taste/flavour associations) in order for “sonic seasoning” effects (Chapter 3) to occur. In Experiment 10, I use time-based sensory evaluation methods to examine whether different sound stimuli can draw people’s attention to different taste/flavour attributes (Experiment 10A) or modify the perception of a particular taste attribute (Experiment 10B).

Multisensory integration is often characterised as an automatic process (Talsma et al., 2010), but it is unclear whether crossmodal correspondences automatically affect people’s performance (or eating/drinking experience, in this case). Automaticity can be defined in a variety of ways (Moors & De Houwer, 2006); one feature of automaticity that I will focus on for the purpose of this thesis is insensitivity to task load (see Spence & Deroy, 2013a, for a review). The measurement of task efficiency is usually measured in dual-task studies where participants perform a primary task at the same time as a second, unrelated task that putatively consumes neural resources. The degree of efficiency is then based on the degree to which the primary task is affected by the secondary task. Over the years, it has been shown that cognitive task load can attenuate the intensity of people’s affective experiences (Kron et al., 2010; van Dillen & Koole, 2007). Moreover, a cognitive task load that involves memorising seven-digit or one-digit numbers has been shown to reduce perceived taste intensity (van der Wal & van Dillen, 2013). It follows that cognitive task load might also reduce the ability of participants to pay attention to the background soundtrack and/or discern small differences in taste. Therefore, in Experiment 9, I investigate whether the “sonic seasoning” effect is sustainable under high cognitive load (e.g., memorising a seven digit number), with the assumption that the memory task would divert attention away from the chocolate tasting task at hand. The experimental

hypothesis is that under high cognitive load, participants would not experience any sonic seasoning effects because they would pay more attention to the cognitive task instead of the background soundtrack. In comparison, under low cognitive load, one would expect to observe the same sonic seasoning effect as observed in Experiment 6.

In Experiment 10, I explore another aspect of attention, namely, whether sound operates by shifting our attention to different flavour aspects in a complex mixture. Crossmodal correspondences between pitch and spatial location have been shown to modulate attentional orienting (Chiou & Rich, 2012), so it is conceivable that auditory stimuli (specifically taste-congruent soundtracks) may be able to shift our attention towards specific taste/flavours. Moreover, since music and food/drink are both time-varying in nature, it seems only appropriate to take temporality into account when studying the impact of music on the eating/drinking experience.

Several recent studies have examined the impact of music on food/drink evaluations, but none have relied on time-based methods. Two common methods of time-based sensory evaluation of food products are time-intensity (TI), where a specific sensory attribute is measured over time, and temporal dominance of sensations (TDS), a relatively new technique used to record several sensory attributes simultaneously over time.

TDS was first introduced at the 5th Pangborn symposium (Pineau et al., 2003) and quickly caught on in the food science community. The technique works by introducing multiple flavour attributes to the participant, where the latter is asked to assess which attribute is perceived as dominant at any given time. It has been used to characterise beverages like blackcurrant squash (Ng et al., 2012) and wine (Meillon et al., 2009; Sokolowsky & Fischer, 2012). To my knowledge, no published data have studied the effect of soundtracks/music on food/beverage perception using TDS.

Anecdotal evidence (e.g., Gray, 2007) reports that the same wine can taste different under the influence of different music. In Experiment 10A, TDS is used to assess whether music can draw participants' attention to different aspects of a wine. If music does direct one's attention to different tastes/flavours, then it should be possible to observe different patterns of attended flavours with different auditory conditions.

In comparison, time intensity (TI) is a much more established method which works by measuring the sensory perception of a specific attribute's intensity and enables the monitoring of temporal changes during product evaluation (Lee & Pangborn, 1986). (see Cadena et al., 2011). It has been used on a large variety of food products such as ice cream (Cadena & Bolini, 2011), beer (King & Duineveld, 1999), and wine (Goodstein et al., 2014; Sokolowsky & Fischer, 2012), to name but a few examples. In Experiment 10B, TI is used to show whether changes in intensity can be observed in a specific taste in a wine if the auditory stimuli changes over time. If crossmodally congruent soundtracks work by shifting people's attentional focus, then we should be able to see changes in taste intensity (as shown by Marks & Wheeler, 1998) if the soundtracks change during the tasting procedure. The experiment focused on two distinct basic tastes – sweet and sour – since they would be easy for participants to attend to. Accordingly, I chose soundtracks which have been previously shown to correspond to sweet/sour tastes, as well as a wine with pronounced sweet and sour tastes.

6.2 Experiment 9: The role of task load on perceptual effects of sound-taste correspondences

6.2.1 Methods

6.2.1.1 Participants

A total of 49 participants (29 women, 20 men) aged 20-34 years ($M=24.92$, $SD=3.64$) took part in the study. The data collection was spread out over two sessions. Participants were recruited via the BI Norwegian Business School's participant recruitment platform. All of the participants gave their informed consent to take part in the study. None of the participants reported a cold, nor any other known impairment of their sense of smell, taste, or hearing at the time of the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

6.2.1.2 Auditory stimuli

The sweet and bitter soundtracks were the same as in Experiment 6.

6.2.1.3 Gustatory stimuli

70% Lindt chocolate was used for the study for its fairly complex taste, ambiguous sweet-bitter balance, and commercial availability. Each sample consisted of approximately 3g of chocolate served in a small clear plastic cup.

6.2.1.4 Design and procedure

The study was designed with load condition (low or high load) and soundtrack type (bitter or sweet) as within-participants factors. Each participant completed four trials, and experienced each load condition-soundtrack combination once.

Experimental sessions comprised of up to ten participants at a time. Each participant was seated in front of a computer screen in an experimental cubicle, isolated from other cubicles by opaque plastic separators. No two participants sat immediately adjacent to one another during the experiment sessions. The experiment was programmed on the Qualtrics online survey platform and participants responded by using the mouse to click or drag the indicator on the continuous rating scales. Each participant was given four chocolates on labelled plates (A, B, C, and D) as well as tap water and crackers to cleanse their palates. For each trial, participants were first asked to memorise a string of numbers, either short (“9”, “2”) or long (“7453624”, “7396054”). Next, they were instructed to start tasting a specific chocolate sample (A, B, C, or D) once they heard a 60 second soundtrack (the soundtracks were cued to play automatically). Once they finished tasting the sample, the participants arrived at a new page where they were asked to recall the number they had memorised. Finally, they were asked to evaluate the balance of bitter and sweet tastes of the chocolate sample on a 0-100 point scale (0=much more bitter than sweet, 100=much more sweet than bitter). They rinsed their mouths out with water and ate a piece of cracker to cleanse their palates between each trial. The order of the soundtracks and load conditions were randomised.

Finally, the participants listened to each soundtrack again. Each soundtrack was rated on 7-point bipolar scales of valence (unpleasant to pleasant) and arousal (calm to exciting). The participants also rated how much each soundtrack matched each of four basic tastes (sweet, bitter, sour, and salty) on a scale of 0-10 (0=not at all matching, 10=very much matching).

The experiment lasted for around 20 minutes and the participants were paid 50 NOK for their time.

6.2.2 Results and discussion

Participants (14 out of 49) who did not perform the memory task correctly were eliminated from the analysis.

The mean values of the participants' ratings for the chocolates are shown in Figure 6.1, for both low and high cognitive load conditions. A RM-ANOVA was conducted on participants' chocolate bitter-sweet ratings with 'load condition' (high or low) and 'soundtrack type' (sweet or bitter) as within-participants factors. There was a significant main effect of soundtrack ($F(1,34)=5.17$, $p=.029$, partial $\eta^2=0.13$), where the chocolates experienced with the sweet soundtrack (either with low or high cognitive load) was rated as sweeter and less bitter than the chocolates experienced with the bitter soundtrack ($M_{\text{bitter}}(\text{SD})=42.41(2.43)$, $M_{\text{sweet}}(\text{SD})=47.56(2.17)$, $p=.029$). However, no significant effect of load condition was observed ($F(1,34)<1$, n.s.), nor any interaction between load condition and soundtrack ($F(1,34)<1$, n.s.).

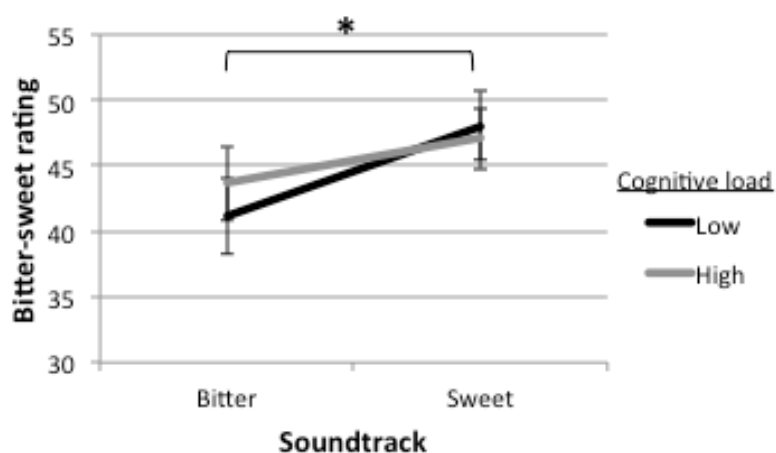


Figure 6.1. Mean values of bitter-sweet ratings for chocolates, for both high and low cognitive load conditions, in Experiment 9. Error bars indicate standard error. Asterisk “*”, indicates statistical significance at $p<.05$, in the comparison between bitter and sweet soundtrack conditions (taken into account both cognitive load conditions).

Assessing the emotion ratings of the soundtracks, participants rated the sweet soundtrack as significantly more pleasant than the bitter soundtrack ($M_{\text{bitter}}(\text{SD})=2.80(0.24)$, $M_{\text{sweet}}(\text{SD})=5.51(0.18)$, $p<.0005$). There was no significant difference in terms of arousal ratings between the two soundtracks ($M_{\text{bitter}}(\text{SD})=3.06(0.22)$, $M_{\text{sweet}}(\text{SD})=3.57(0.33)$, $p=.23$). In terms of participants' sound-taste association ratings for each soundtrack (see Figure 6.2), the bitter soundtrack was rated as matching bitterness significantly better than the other tastes ($p<.0005$ for all comparisons), and the sweet soundtrack was rated as matching with sweetness significantly better than with the other tastes ($p<.0005$ for all comparisons).

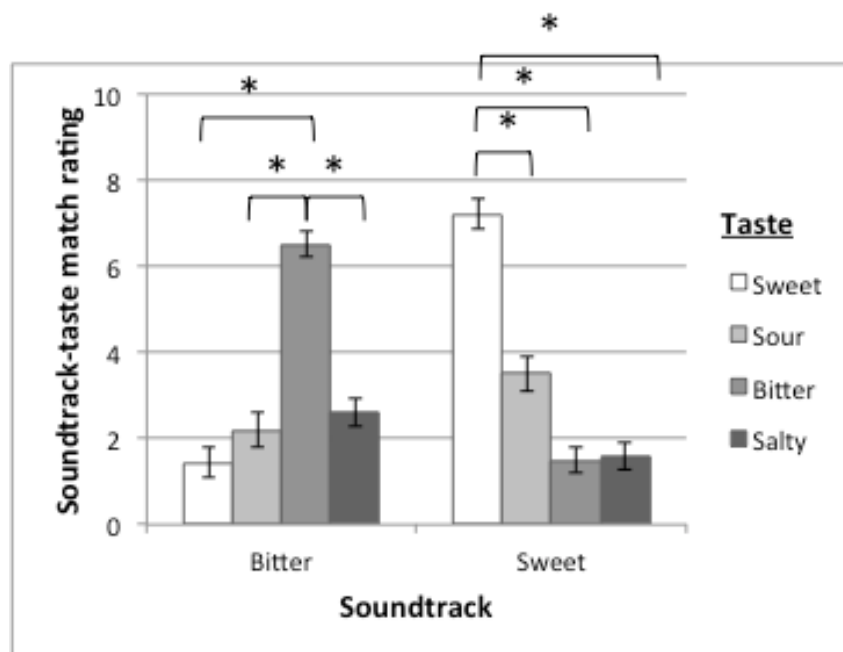


Figure 6.2. Mean values of soundtrack-taste match ratings for the sweet and bitter soundtrack used in Experiment 9. Participants rated how well each soundtrack matched four basic tastes (sweet, sour, bitter, and salty) on a scale from 1-10. Error bars indicate standard error. Asterisk ‘*’, indicates statistical significance at $p < .05$.

The aim of Experiment 9 was to evaluate the effect of cognitive load on the perceptual effects of sound-taste correspondences. While soundtrack type (bitter,

sweet) did have the predicted effect on participants' taste ratings for chocolate, no interaction effect between load condition and soundtrack type was observed. Therefore, there was no support for the thesis that high cognitive load would reduce the influence of taste-congruent soundtracks.

Even though no effect of cognitive load was observed in this experiment, it is nevertheless unclear how cognitive load might have influenced the tasting experience. Might it have affected people's ability to perceive the soundtrack? Or might it have influenced their ability to taste the chocolate? A future study that compares the influence of cognitive load on people's taste experience (conducted in silence), might serve as a useful baseline for comparing the effect of cognitive load.

Another important limitation of Experiment 9 is that cognitive load was manipulated, in terms of working memory, as opposed to perceptual load. There has been much research showing that perceptual load can interfere with unisensory processing (Lavie, 1995; Lavie & Cox, 1997; Macdonald & Lavie, 2011) and multisensory, specifically audiovisual, integration (Alsius et al., 2005, 2007; Mattingley et al., 2006). However, at least when it comes to spatial attention, multisensory cues seemingly retain their attention-capturing ability under perceptual load, when compared to unisensory stimuli (see Spence & Santangelo, 2009; and Spence, 2010; for reviews). Take the realistic scenario of diners engaged in a conversation at a restaurant; could background music influence their dining experience? The task of listening and speaking has been shown to hinder the evaluation of alcohol levels in a drink (Stafford et al., 2012, 2013) as well as reduce people's susceptibility to the McGurk illusion (Alsius et al., 2005). However, if attention to specific tastes/flavours in the food is comparable to spatial attention, then one might predict that presenting

multisensory cues like matching background music¹ would be immune to any increase in perceptual load, such as conversing with one's dining companions. Future studies examining the role of perceptual load will be crucial too determining the feasibility of applying such "sonic seasoning" effects in restaurants.

6.3 Experiment 10A: Analysing the impact of music on wine perception via time-based methodologies: Temporal Dominance of Sensations (TDS)

6.3.1 Methods

6.3.1.1 Participants

A total of 21 participants (11 women, 10 men) aged 21-69 years ($M=37.6$, $SD=12.8$) took part in the study. The participants were recruited at the Universidad de Tres de Febrero in Buenos Aires, Argentina. All of the participants gave their informed consent to take part in the study. None of the participants reported a cold nor any other known impairment of their sense of smell, taste, or hearing at the time of the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

6.3.1.2 Auditory stimuli

Two pieces of music were chosen for the study that varied in tempo, mode, and instrumentation. Soundtrack 1 was Brian Eno's "Discreet Music", and soundtrack 2 was Mussorgsky's "A Night on Bald Mountain". Both pieces have been used in

¹ Especially if the dish is also presented with congruent visual (colours, shapes) and tactile (texture and weight of servingware) cues.

various wine-music demos and talks related to sound-taste interactions in previous events. 45 second excerpts were taken from the beginning of both pieces, and a pre-test (N=19) showed that they significantly changed the perceived fruitiness and tannin levels of a light red wine (Georges du Boeuf's Beaujolais-Village 2014).

6.3.1.3 Wine

Experiment 10A involved a Pinot Noir produced in Argentina – Manos Negras Red Soil Select Pinot Noir, 2014. The wine has medium body and alcohol, soft tannins, and crisp acidity, with flavours of violets, cherries and earth. The wine was served in 15 mL samples, inside clear plastic 50 mL cups.

6.3.1.4 Design and procedure

Each participant was seated in front of a computer screen with headphones and a cup of water to cleanse their palates. The experiment was programmed using the Sensomaker tool for sensorial characterisation of food products (Pinheiro et al., 2013).

For each trial, the participants were given a sample of wine by the experimenter. They were instructed to start the trial as soon as they put the wine in their mouth, and to hold the wine there for the 45 second duration of the trial. During the trial, the TDS computerised system showed the participant the entire list of 8 adjectives in two columns (red fruit, tannins, alcohol, woody, sweet, acidic, spicy, and bitter). The attributes were selected based on a similar TDS study on Pinot Noir wines (Visalli, 2016). The participant was instructed to click on the start button as soon as the wine sample has been put in their mouth, then to consider which attribute is perceived as the most dominant. Each time they feel like the perception has changed, they click on a new attribute which they perceive to be most dominant. The participant is free to

select an attribute several times during the course of the trial, or not at all. Participants first practiced with a weak yerba mate tea solution 1-3 times, until they could operate the TDS software with ease.

The order of adjectives was randomised for each participant to avoid any order effect of the list of attributes. However, for each participant, the order was always the same so learning the terms and scoring was facilitated. The participants always tasted the wine in the silent condition first, but the order of the two music soundtracks was randomised. In those trials with a soundtrack, the experimenter started the music at the same time as the participant clicked on the start button.

The experiment lasted for around 10 minutes and the participants were debriefed afterwards.

6.3.2 Results and discussion

Figure 6.3 shows TDS curves for each sound condition. TDS graphs were produced by SensoMaker software. Two lines are drawn on the graph, “chance level” is the dominance rate that an attribute would be chosen by chance, in this case equal to $1/8$, since there are 8 attributes. The second line shows “significance level”, which is the minimum value for the dominance level to be considered significantly higher than chance. It is calculated using the confidence interval of a binomial proportion based on a normal approximation (Pineau et al., 2009). Concentrating on dominance ratings at above significance level, three major differences are shown. First, onset time for acid occurs at around 9 seconds in the silence condition, whereas acid peaks at 23 seconds for soundtrack 1 and 27 seconds for soundtrack 2. Second, bitterness is prominent from 25-38 seconds in soundtrack 2, and from 8-14 seconds in soundtrack

1 (and is barely registered at 29 seconds for the silent condition. Finally, astringency is significant throughout wine evaluation in the silent condition, but is not significant during the two soundtracks. For soundtrack one, acid is registered before bitterness, whereas in soundtrack 2, bitterness is registered before acidity.

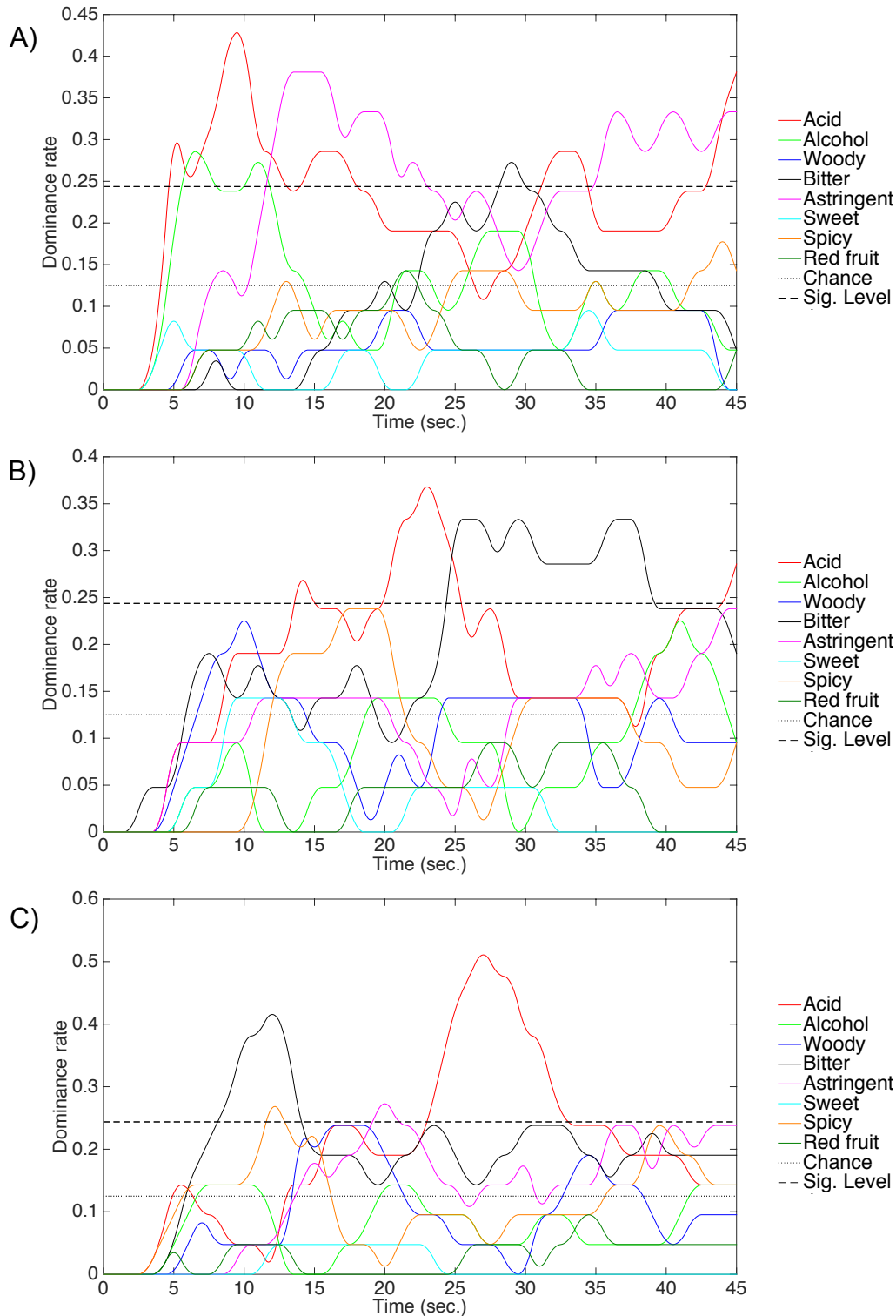


Figure 6.3. Experiment 10A’s TDS graphical representation for Manos Negras Pinot Noir wine, under different background sound conditions: A) silence, B) soundtrack 1 – Brian Eno’s Discreet Music, C) soundtrack 2 – Mussorgsky’s Night on Bald Mountain.

For each sound condition, the total number of times each adjective was rated as well as the duration of dominance for each of the 8 adjectives was calculated. Since none of the measures were normally distributed according to the Shapiro-Wilk test, non-parametric rank-sum tests were used. A Friedman test revealed there were no significant differences in the total number of adjectives participants' used for each sound condition ($\chi^2(2)=4.00$, $p=.14$). There were also no significant differences in dominance durations of acidity ($\chi^2(2)=2.00$, $p=.37$), alcohol ($\chi^2(2)=2.92$, $p=.23$), woodiness ($\chi^2(2)=3.86$, $p=.14$), sweetness ($\chi^2(2)=1.81$, $p=.41$), spiciness ($\chi^2(2)=2.68$, $p=.26$), or red fruit ($\chi^2(2)=1.27$, $p=.53$).

There was, however, a difference in dominance durations for bitterness ($\chi^2(2)=7.75$, $p=.021$) and astringency ($\chi^2(2)=7.55$, $p=.023$). Post-hoc analysis with Wilcoxon signed-rank tests was conducted. Compared to the silence condition, there were significant increase in bitterness dominance durations for both soundtrack 1 ($Z=-2.28$, $p=.023$) and soundtrack 2 ($Z=-2.30$, $p=.021$). There were, however, no differences when comparing the two soundtrack conditions with each other ($Z=-0.024$, $p=.98$).

There were also significant reductions in astringency dominance durations for both soundtrack 1 ($Z=-2.04$, $p=.042$) and soundtrack 2 ($Z=-2.19$, $p=.029$), compared to the silence condition. There were no differences when comparing the two soundtrack conditions with each other ($Z=-0.75$, $p=.45$).

To see more clearly the effect of each soundtrack on the wine, TDS difference curves were plotted (see Figure 6.4), where the difference between each soundtrack and the silent condition are plotted at points when the difference is significantly different from zero using a test to compare to binomial proportions (Pineau et al., 2009). The effect of soundtrack 1 (Brian Eno's Discreet Music) on the wine, compared to the

silent condition, is an enhancement of bitterness and a reduction of alcohol in the 0-15 second timeframe, and then a reduction in alcohol at around the 30 second mark. The effect of soundtrack 2 (Mussorgsky's Night on Bald Mountain) on the wine, compared to the silent condition, is a longer and more prominent enhancement of bitterness during the 0-15 second timeframe along with a reduction in acidity and astringency. There follows an enhancement in acidity around the 25-30 second timeframe.

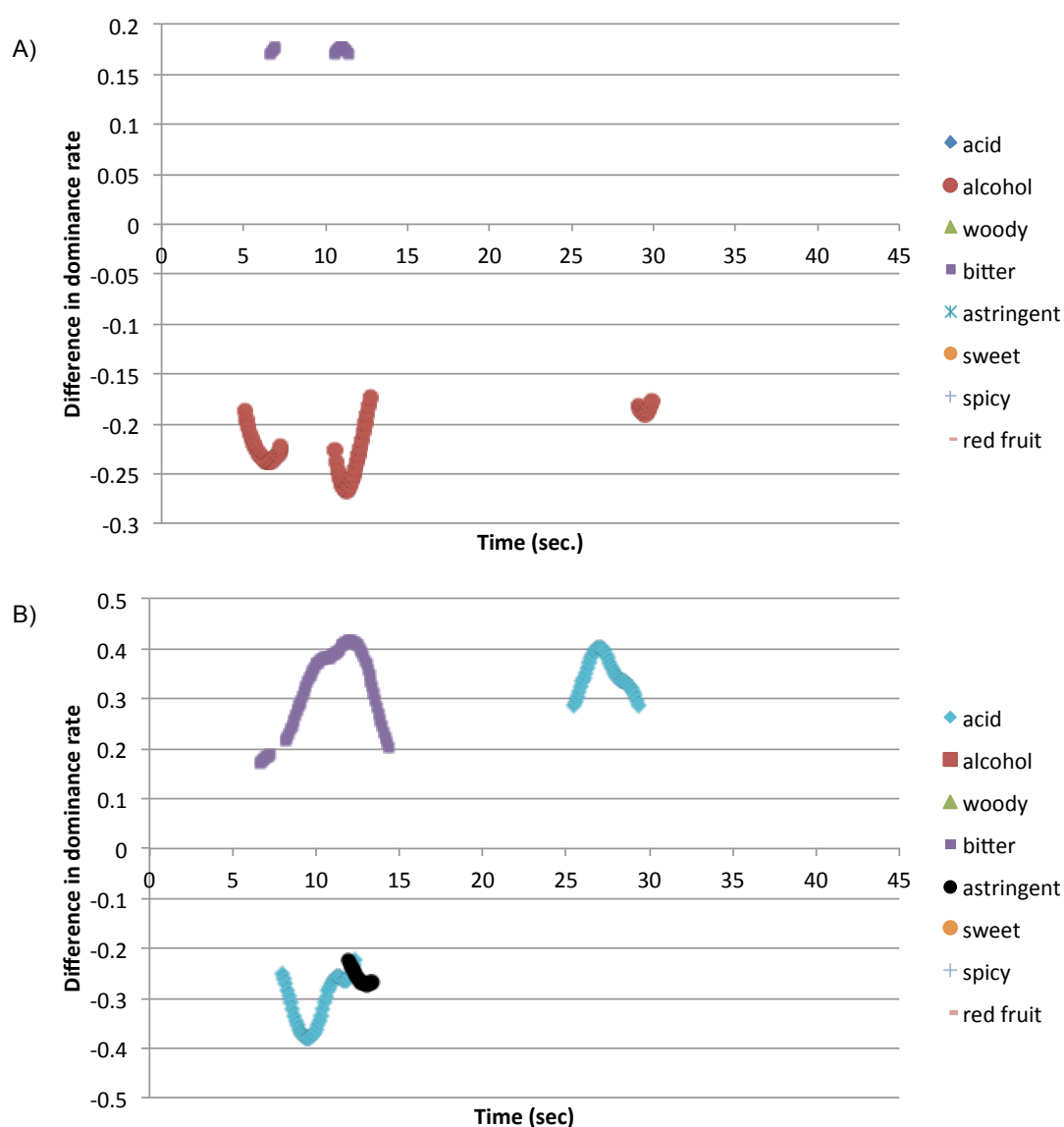


Figure 6.4. Experiment 10A's TDS difference curves between A) soundtrack 1 and silence condition, and B) soundtrack 2 and silence condition. Only differences significantly greater than

zero (i.e., significantly differences between attribute dominance rates evaluated while listening to a soundtrack compared to listening in silence) are plotted.

The results of Experiment 10A demonstrate that, as predicted, different soundtracks revealed different patterns of dominant flavours. Overall, the attack of acidity occurs earlier in the silent condition than in either of the soundtrack conditions, and astringency is less noticeable when there is music playing. Bitterness is more prominent in the attack of the wine for the Mussorgsky piece, whereas for the Brian Eno piece bitterness comes after the initial registration of acidity. Analysing dominance durations supported results from TDS graphs, where bitterness was dominant for longer – but astringency shorter – during the two soundtrack conditions when compared to the silence condition.

There were no significant differences in the number of adjectives participants selected – on average the number was around 4, or half of the 8 available adjectives. For this group of participants, basic tastes such as acid, bitter was used more often ($Dur_{acid}=9.1$ seconds, $Dur_{bitter}=7.4$ seconds) than more descriptive terms such as red fruit and woody ($Dur_{red_fruit}=1.6$ seconds, $Dur_{woody}=3.7$ seconds). This might either be attributable to the fact that the participants simply did not taste the more descriptive attributes in the wine, or that basic tastes were simply more dominant (or more easily came to mind) in the wine, especially under experimental conditions.

Future work lies in associating specific characteristics of music with peaks in the TDS graphs, to tease apart exactly what in the music triggered specific attention to different attributes in the wine.

6.4 Experiment 10B: Analysing the impact of music on wine perception via time-based methodologies: Time Intensity (TI)

6.4.1 Methods

6.4.1.1 Participants

The same participants who participated in Study 10A also took part in Study 10B. One participant misheard the instructions and rated the wrong tastes, so their results were excluded.

6.4.1.2 Auditory stimuli

Experiment 10B involved soundtracks designed to match with sweet and sour tastes. The sweet soundtrack came from Jialin Deng's composition (see Experiment 2). The sour soundtrack, *Superscriptio* by Brian Ferneyhough, has been shown to be associated with sour tastes (Kontukoski et al., 2015). Two mixed soundtracks were then created based on these sweet and sour soundtracks. One was a sour-sweet soundtrack with 15 seconds of the sour soundtrack immediately followed by 15 seconds of a sweet soundtrack. The other was a sweet-sour soundtrack produced in reverse order.

6.4.1.3 Wine

Study 10B involved an off-dry chenin blanc produced in Argentina – Santa Julia Chenin Dulce Natural 2015. This wine has low alcohol (7.5%), crisp acidity, and a medium level of residual sugar (73.3 g/l). It has aromas of white peach, apricot, and citrus and citrus flavours with a good balance between acidity and sweetness.

6.4.1.4 Design and procedure

Each participant was seated in front of a computer screen with headphones and a cup of water to cleanse their palates. The experiment was programmed using the Sensomaker tool for sensorial characterisation of food products (Pinheiro et al., 2013).

For each trial, participants were given a sample of wine by the experimenter and told to focus on either the sweetness or sourness of the wine. They were instructed to start the trial as soon as they put the wine in their mouth, and to hold it there for the 30 second duration of the trial. The participant was instructed to click on the start button as soon as the wine sample had been put in their mouth, then to move the cursor along a horizontal scale to rate the sensation of sweetness or sourness as it involves in the mouth, until the end of the trial. The experimenter started the music at the same time as the participant clicks on the start button.

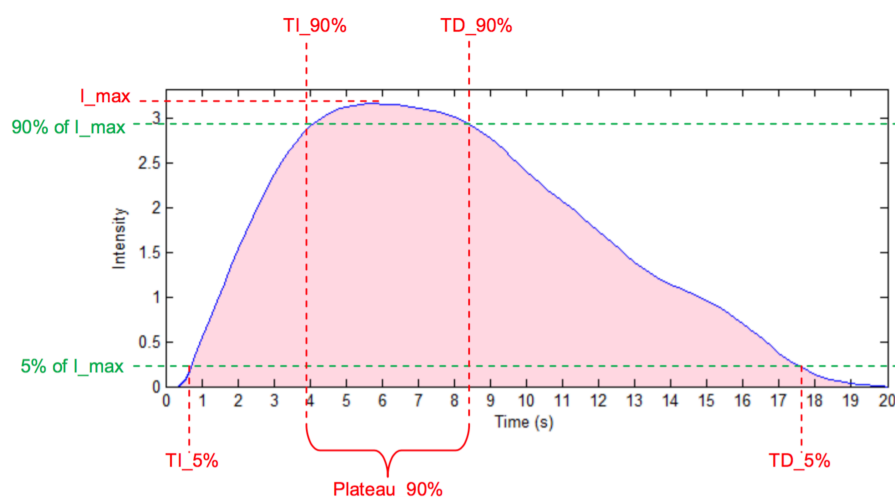
There were a total of four trials, two trials where the participants had to rate sweetness intensity (once listening to sour-sweet soundtrack, once listening to sweet-sour soundtrack), and two trials where the participants had to rate sour intensity (once while listening to the sour-sweet soundtrack, once while listening to the sweet-sour soundtrack). The order of the trials was randomised using a Williams Design Latin Square.

The experiment lasted for around 10 minutes and participants were debriefed afterwards.

6.4.2 Results and discussion

TI graphs and parameters of I_{\max} (maximum intensity), TI_{90} (time when intensity is at 90% of I_{\max} at increasing part of curve), TD_{90} (time when intensity is at 90% of I_{\max}

at decreasing part of curve), and the areas under the curve (AUCs) were produced by Sensomaker software (see Figure 6.5 for an illustration). An I_{30-15} measure, which is the difference in participants' intensity ratings at 15 and 30 seconds, was calculated to take into account participants' taste ratings with changes in the soundtrack (namely at the beginning and end of the second part of the two-part soundtracks). RM-MANOVA with within-participant factors of soundtrack type (sour-sweet, sweet-sour) and taste rating (sour, sweet) were conducted on the time intensity graph parameters.



I_max: maximum intensity

TI_5%: time when intensity is 5% of I_max at increasing part of the curve

TD_5%: time when intensity is 5% of I_max at decreasing part of the curve

TI_90%: time when intensity is 90% of I_max at increasing part of the curve

TD_90%: time when intensity is 90% of I_max at decreasing part of the curve

Plateau_90%: time interval which the intensity is $\geq 90\%$ of I_max

Area: area under the curve

Figure 6.5. Time-intensity curve parameters used in Experiment 10B. Reproduced from Sensomaker user guide (Nunes & Pinheiro, 2015).

The average TI graphs for each of the 4 conditions are shown in Figure 6.6. The 15-second point is demarcated on the graphs, showing when the soundtracks changed

from sour to sweet (see Figures 6.6A and B) or from sweet to sour (see Figures 6.6C and D). Visually inspecting the graphs, there is a notable difference in the slope of the average TI curve in the 15-30 second region between 6A and C), where 6A is positive and 6C is negative (in other words, sweetness rating during the sweet soundtrack section increases, but decreases during the sour section). A similar difference can be seen between 6B and D.

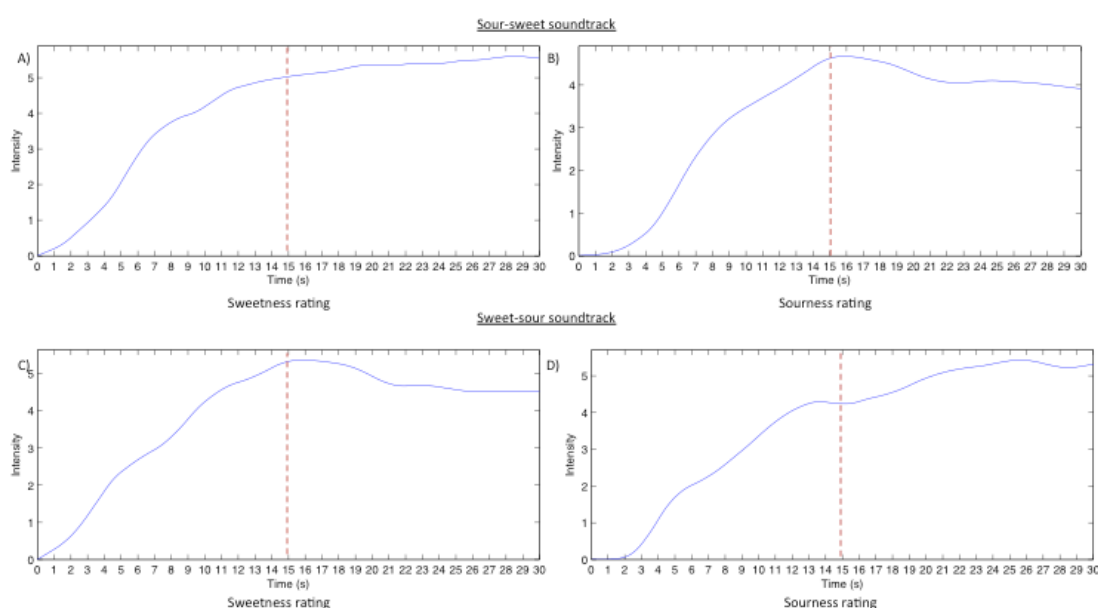


Figure 6.6. Mean values of time intensity (TI) values for each of the four soundtrack-rating conditions in Experiment 10B: A) sour-sweet soundtrack, sweetness rating; B) sour-sweet soundtrack, sourness rating; C) sweet-sour soundtrack, sweetness rating; D) sweet-sour soundtrack, sourness rating. The dotted line is shown at the 15-second mark, where the soundtrack changes from either sweet-to-sour or sour-to-sweet.

Average values of I_{\max} , TI_{90} , TD_{90} , and AUC in the four experimental conditions are shown in Table 6.1. A RM-MANOVA showed that there was a significant interaction effect between soundtrack type and taste rating, $F(4,16)=5.32$, $p=.006$, Wilk's $\Lambda=0.43$. Further univariate ANOVAs revealed that the significant interaction

effect applied to the TI_{90} ($F(1,19)=9.66$, $p=.006$, $\eta^2=0.34$) and TD_{90} ($F(1,19)=13.00$, $p=.002$, $\eta^2=0.41$) ratings. Post hoc pairwise comparison tests with Bonferroni corrections revealed that, for the sweet-sour soundtrack, TI_{90} was before the 15 second mark for sweet intensity evaluation but after the 15 second mark for the sour intensity evaluation ($TI_{90,sweet}=12.56$, $TI_{90,sour}=17.60$, $p=.035$). Similarly, TD_{90} occurred significantly earlier for the sweet intensity evaluation than for the evaluation of sour intensity ($TD_{90,sweet}=23.88$, $TD_{90,sour}=28.05$, $p=.023$). In other words, when participants listened to the sweet-sour soundtrack, the period of 90% taste intensity registered occurred significantly earlier for sweetness evaluation than sourness evaluation. A similar, but not significant, trend was observed during the sour-sweet soundtrack, where the period of 90% sour taste intensity occurred earlier than the period of 90% sweet intensity ($TI_{90,sweet}=16.22$, $TI_{90,sour}=13.91$, $p=.19$; $TI_{90,sweet}=27.26$, $TI_{90,sour}=23.66$, $p=.051$).

There were no significant main effects of soundtrack ($F(4,16)<1$, n.s.) or taste rating ($F(4,16)<1$, n.s.).

Soundtrack	Taste rating	I_{max}	AUC	TI_{90}	TD_{90}
Sweet-sour	Sweet	5.91 (0.46)	117.75 (11.81)	12.56 (1.47)	23.88 (1.29)
Sweet-sour	Sour	5.32 (0.50)	110.14 (12.63)	17.60 (1.77)	28.05 (0.87)
Sour-sweet	Sweet	6.31 (0.47)	125.44 (13.66)	16.22 (1.89)	27.26 (1.16)
Sour-sweet	Sour	5.32 (0.50)	97.29 (10.10)	13.91 (1.61)	23.66 (1.53)

Table 6.1. Time-intensity graph parameters from Experiment 10B, grouped by soundtrack condition and which taste the participant was asked to focus on during the trials. Standard errors of the means are shown in parentheses.

For a more precise look at how changes in the soundtrack might have influenced participants' taste ratings, the differences in participants' intensity ratings between 15 and 30 seconds (I_{30-15}) were measured (see Figure 6.7). When the participants were asked to rate the sweetness of the drink, I_{30-15} was negative for the sweet-sour soundtrack but positive for the sour-sweet soundtrack. In other words, sweetness decreased when the soundtrack went from sweet to sour, but increased when the soundtrack went from sour to sweet. The exact opposite pattern was observed for the sour ratings, where I_{30-15} was positive for the sweet-sour soundtrack but negative for the sour-sweet soundtrack. (This provides evidence that a changing soundtrack can modify people's taste ratings over time, even when the beverage is the same.)

A RM-ANOVA analysis of I_{30-15} revealed a significant interaction effect between soundtrack and taste rating ($F(1,19)=17.96$, $p<.0005$, $\eta^2=0.49$). Post hoc tests with Bonferroni corrections revealed that, for sweet ratings, I_{30-15} is lower during the sweet-sour soundtrack as compared to the sour-sweet soundtrack ($p=.044$). On the other hand, for sour ratings, I_{30-15} is higher during the sweet-sour soundtrack as compared to the sour-sweet soundtrack ($p=.022$). There were no significant main effects of soundtrack ($F(1,19)<1$, n.s.) or taste rating ($F(1,19)<1$, n.s.).

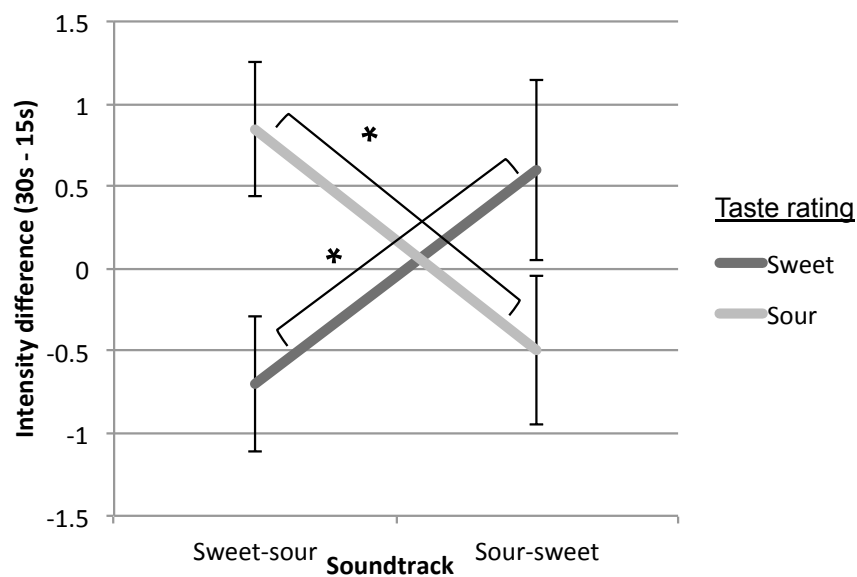


Figure 6.7. Mean values of time intensity (TI) differences between participants' ratings at 30 seconds and 15 seconds in Experiment 10B. For the sweet-sour soundtrack, 15 seconds denotes the switch in the music from the sweet soundtrack to the sour soundtrack. The same holds, reversed, for the sour-sweet soundtrack. Error bars denote the standard error of the means. Asterisk '**', indicates statistical significance at $p < .05$.

The results of Experiment 10B demonstrate that, as predicted, a change of soundtrack results in a change in taste intensity. Furthermore, the change in taste intensity is in the same direction as the soundtrack in question – a switch from sour to sweet soundtrack enhanced sweetness intensity, whereas the opposite switch from sweet to sour soundtrack enhanced sour intensity. These results are especially striking considering that participants evaluated the same mouthful of wine at different time points, with the knowledge that the wine itself cannot have changed (see Woods et al., 2010, for a demonstration of the unity assumption).

Figure 6.6 shows an inflection point at around 15 seconds (especially evident in panels 6.6B and C). Since 15 seconds is where the soundtrack changed from sweet to sour, or from sour to sweet, this supports the anecdotal report (see Crisinel et al.,

2012) that changes in auditory stimuli can influence taste evaluation quickly. Moreover, this could be interpreted as a counterexample to the theory that music might operate by changing the taster's emotion state, since it would take several seconds (eight, according to Bachorik et al., 2009) for the emotional impact of music to build up (see Spence & Wang, 2015c, for a more in-depth discussion).

6.5 General discussion and conclusions

In this chapter, the role of attention on crossmodal interactions between sound and taste/flavour was assessed in several different ways. First, Experiment 9 evaluated the effect of cognitive load, via a working memory task, on the perceptual effect of sound-taste correspondences. No interaction effect between task load (high, low) and soundtrack type (bitter, sweet) was observed on the participants' taste ratings, so the automaticity of the sonic seasoning effect remains inconclusive. Future studies involving perceptual load, especially vocal shadowing tasks (which, in some sense, mimics restaurant dining conditions), are needed in order to understand how sound-taste correspondences influence information processing and evaluate the feasibility of playing flavour-congruent soundscapes in real-world applications.

Second, Experiments 10A demonstrated that different pieces of music revealed different patterns of attended flavours in a wine over a period of 45 seconds. This finding is a first step towards supporting the theory that what we hear can (perhaps automatically) shift our attention to specific taste/flavour attributes. A possible follow-up study might involve a TDS test using soundtracks with known taste correspondences (as tested in Experiment 2, for example), in order to verify that

listening to such a soundtrack does indeed lead to the sensation dominance of the corresponding taste.

Finally, Experiment 10B showed that the influence of sound on taste can act fairly quickly (on the order of seconds), so that when the soundtrack changes from a sweet-congruent one to a sour-congruent one (and vice versa), the change is reflected in participant ratings at around the same time as the soundtrack change. This finding agrees with anecdotal evidence from musicians and wine experts (reported by Crawshaw, 2012, and reviewed in Spence & Wang, 2015c), where “the gustatory and olfactory changes [in the wine] perceptually appear to occur immediately, coinciding with changes in mode, tempo, and more generally, the style/genre [of the music]”.

Taken together, the results of Experiments 10A and 10B document the first time that time-based methodologies are used to investigate the influence of auditory stimuli on the tasting experience. The successful deployment of TDS and TI here opens many future avenues of research, with a focus on the temporal aspects of the music listening experience as well as the tasting experience. As discussed in Spence and Wang (2015b, most off-the-shelf music has not been ideal for research purposes since stylistic changes often occur and, unless one is careful, there is a real danger in a piece of music corresponding to different tastes/flavours. Learning more about the temporal characteristics of such “sonic seasoning” effects could free researchers from such constraints, as well as enable experience designers to create more fluid and sophisticated experiences which take advantage of the evolving nature of both the listening and the eating/drinking experience.

7 Emotion as a possible mechanism underlying the auditory modulation of taste/flavour

7.1 Introduction

The results of the experiments reported back in Chapter 2 illustrated the role of emotion mediation as a way of explaining the crossmodal correspondences between sound and taste. In this chapter, I extend that idea to assess whether emotion – specifically valence – is a plausible mechanism underlying the influence of soundscapes on the tasting experience. The basis for this chapter comes from Experiment 3, where harmonic consonance was shown to be associated with sweetness while harmonic dissonance was associated with sourness. Furthermore, participants rated the same juice mixture as sweeter or more sour (and also more or less liked) depending on whether the melody was harmonized with consonant or dissonant intervals.

One possible explanation for the taste modulation effect observed in Experiment 3 was that the soundtracks were priming certain emotions in the mind of the participant. Neurologically, consonant and dissonant intervals are deeply ingrained in low-level sensory processing; consonant intervals evoke more pronounced brainstem responses than dissonant intervals (Bidelman & Krishnan, 2009), and various degrees of consonance and dissonance have highlighted increased activity in the paralimbic and neocortical regions responsible for processing pleasant/unpleasant emotional states (Blood et al., 1999). It is well-known that people's emotional perception of stimuli in one sensory modality can be altered by the emotional information from another modality; for instance, vocally expressed affective information has been shown to

bias the interpretation of facial emotions (de Gelder & Vroomen, 2000; Ethofer et al., 2006; Massaro & Egan, 1996), postures (Van den Stock et al., 2007), and body movements (Van den Stock et al., 2008). In terms of musical stimuli, music-elicited emotion has been shown to influence people's performance in those tasks involving the processing of visual emotional stimuli (Logeswaran & Bhattacharya, 2009; see Gerdes et al., 2014, for a review). In addition, musical emotion has also been shown to alter brightness judgment (Bhattacharya & Lindsen, 2016). Returning to Experiment 3, it is plausible that any positive feelings induced by consonant music (and equally, any negative emotions induced by the dissonant music) may simply have been transferred to the juice-rating task, thus resulting in higher reported pleasantness (and sweetness) ratings for the juice.

In order to test this theory, I decided to re-run Experiment 3 with the introduction of additional visual stimuli. If visual images – with similar valence ratings – induced a similar amount of taste modulation as the consonant/dissonant soundtracks, this would provide evidence that the mechanism behind the results reported in Experiment 3 was some kind of sensation transfer. I used visual images from the International Affective Picture System (IAPS) with similar valence ratings as the consonant/dissonant soundtracks in Experiment 3. IAPS is an on-going effort at the NIHM Center for the Study of Emotion and Attention to create and distribute stimuli with standardised emotion ratings (valence, arousal, and dominance) for use in studies of emotion and attention (Lang et al., 2008). The juice mixture and music stimuli were the same as in Experiment 3.

In Experiment 11, the experimental hypothesis is that, if participants are influenced by the emotional content of what they see or hear, then the positive stimuli – regardless of modality – should be liked more and induce a sweeter taste in the juice

mixture. Likewise, it is expected that exposure to the negative stimuli should induce a sourer taste rating and diminished liking for the juice mixtures.

7.2 Experiment 11: Assessing the role of valence on the influence of visual and auditory stimuli on taste perception

7.2.1 Methods

7.2.1.1 *Participants*

A total of 49 participants (36 women, 13 men) aged 18-60 years ($M=30.86$, $SD=10.98$) took part in the study. The participants were recruited at the Food Matters Live conference at the London ExCeL center, November 22, 2016. All of the participants gave their informed consent to take part in the study. None of the participants reported a cold nor any other known impairment of their sense of smell, taste, or hearing at the time of the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

7.2.1.2 *Auditory stimuli*

Consonant and dissonant versions of a short melody were created. The consonant version was harmonised with major and minor thirds (three and four half-steps, respectively), while the dissonant version was harmonised with minor second intervals (one half-step). The music scores were created using the MuseScore software and exported to midi form, and instrumented using GarageBand's Steinway Grand Piano plugin. The same soundtracks have been used in Wang and Spence (2016) and have been shown to be associated with different tastes (on a 10 point sour-

sweet scale, $M_{\text{consonant}}=6.78$, $M_{\text{dissonant}}=3.27$) and different pleasantness ratings (on a 10 point unpleasant-pleasant scale, $M_{\text{consonant}}=5.23$, $M_{\text{dissonant}}=2.51$). The sounds can be downloaded from <https://soundcloud.com/janicewang09/sets/harmony>.

7.2.1.3 Visual stimuli

Two images, of a crying and laughing child, respectively, were chosen from the International Affective Picture System (IAPS) (Lang et al., 2008), slide numbers 2900.1 and 2900.2. The two images were chosen because they were identical except for the facial expressions of the child, and because they had very different valence ratings (on a scale of 1-9, $M_{\text{valence}_{2900.1}}(\text{SD})=2.56(1.41)$, $M_{\text{valence}_{2900.2}}(\text{SD})=6.62(1.97)$), but similar arousal ratings (on a scale of 1-9, $M_{\text{arousal}_{2900.1}}(\text{SD})=4.61(2.07)$, $M_{\text{arousal}_{2900.2}}(\text{SD})=4.52(1.92)$).

7.2.1.4 Gustatory stimuli

A juice mixture was prepared with a blend of 2:2:1 ratio Tropicana white grapefruit juice, Sainsbury's smooth orange juice, and Sainsbury's apple juice. This was the same ratio of juices as used in Experiment 3A. A juice mixture was used instead of a single juice in order to minimize the possibility that the participants would guess the nature of the juice sample and hence perhaps feel overconfident giving their responses, especially given the within-participants nature of the experimental design (cf. Experiment 3). Each sample consisted of 15 mL of juice and was served in a transparent plastic shot glass. Each participant received four samples, served on a white Styrofoam plate with 3-digit labels.

7.2.1.5 Design and procedure

The study was designed with valence (pleasant or unpleasant sound or visual image) and modality (sound or image) as the within-participants factors. Each participant completed four trials and experienced all four stimuli.

Each participant was seated in front of a computer screen with headphones, a plate with four juice samples (labelled with 3-digit identifiers 719, 420, 536, and 148), and a cup of water to cleanse their palates. The experiment was programmed on the Qualtrics online survey platform and participants responded by using the mouse to click or drag the indicator on the continuous rating scales. Participants were randomly assigned to experience either the two visual stimuli first or the two auditory stimuli first. For the visual stimuli, the participants had to taste the juice sample while looking at the image. For the auditory stimuli, participants were asked to play the sound clip and taste the juice sample once they could hear the music. For each trial, participants were asked to keep the entire juice sample in their mouths for the duration of the 15-second tasting period. Next, the participants navigated to a new page and evaluated the juice sample on an 11-point sour-sweet scale (0=much more sour than sweet, 10=much more sweet than sour) and an 11-point pleasantness scale (0=very unpleasant, 10=very pleasant). The participants were instructed to rinse their mouths out with water between each trial. The order of presentation of the pleasant/unpleasant stimuli and the order of each set of questions were randomised for each participant.

Finally, the participants viewed each of the stimuli again. For each soundtrack or image, they rated it on 7-point bipolar scales of valence (unpleasant to pleasant) and arousal (calm to exciting).

The experiment lasted for around 5 minutes and participants were debriefed afterwards.

7.2.2 Results and discussion

The mean values of the participants' taste and liking ratings for the juice samples in all conditions are shown in Figure 7.1. RM-MANOVA was conducted with 'stimuli valence' and 'stimuli modality' as the within-participants factors and sour-sweet and liking ratings as response measures.

Overall, there was a significant main effect of valence ($F(2,47)=10.02$, $p<.0005$, Wilks' Lambda=0.70), but not of modality ($F(2,47)<1$, n.s.). There was also no significant interaction between valence and modality ($F(2,47)<1$, n.s.). More specifically, valence had a significant effect on both sour-sweet taste ratings ($F(1,48)=6.73$, $p=.013$, $\eta^2=0.12$) and liking ratings ($F(1,48)=19.88$, $p<.0005$, $\eta^2=0.29$). Participants rated the juice as being sweeter and less sour ($M_{\text{positive}}(\text{SE})=4.72(0.20)$, $M_{\text{negative}}(\text{SE})=3.99(0.24)$, $p=.013$) while exposed to more positive stimuli (the laughing image and the consonant music) compared to the more negative stimuli (the crying image and dissonant music). In addition, they liked the juice more while exposed to the more positive stimuli compared to the more negative stimuli ($M_{\text{positive}}(\text{SE})=5.74(0.23)$, $M_{\text{negative}}(\text{SE})=4.77(0.22)$, $p<.0005$).

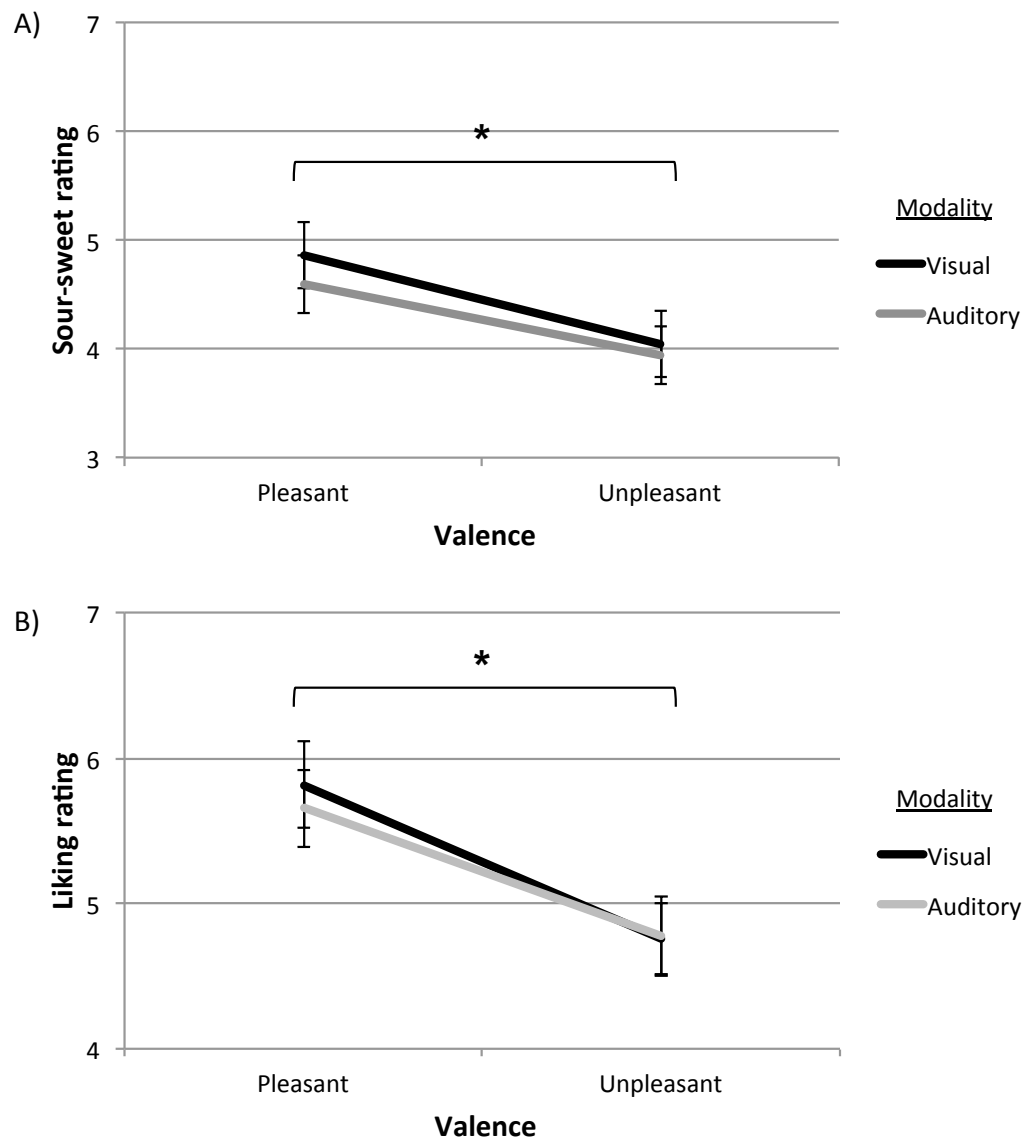


Figure 7.1. Mean values of sour-sweet (A) and liking (B) ratings for juice samples under both stimuli modality (visual, auditory) and valence (pleasant, unpleasant) types in Experiment 11. Error bars indicate standard error. Asterisk '**', indicates statistical significance at $p < .05$.

Assessing the emotion ratings of the soundtracks, the mean values of the participants' valence and arousal ratings for the visual and auditory stimuli are shown in Figure 7.2. To ensure that the auditory and visual stimuli were comparable in terms of valence ratings, I conducted an RM-MANOVA with valence and modality as within-participants factors, and participants' valence and arousal ratings as measures.

Overall, there was a significant main effect of valence on participants' valence ratings ($F(1,48)=367.17$, $p<.0005$, $\eta^2=0.88$) and arousal ratings ($F(1,48)=8.30$, $p=.006$, $\eta^2=0.15$). The more pleasant stimuli (consonant music, laughing face) were rated as being significantly more pleasant than the more unpleasant stimuli (dissonant music, crying image), $M_{\text{pleasant}}(\text{SE})=5.70(0.12)$, $M_{\text{unpleasant}}(\text{SE})=2.12(0.11)$, $p<.0005$, as expected. The more pleasant stimuli were also rated as more calming and less arousing than the more unpleasant stimuli, $M_{\text{pleasant}}(\text{SE})=3.79(0.20)$, $M_{\text{unpleasant}}(\text{SE})=4.55(0.15)$, $p=.006$.

There were no significant interaction effects between stimuli valence and modality for valence ratings ($F(1,48)=3.19$, $p=.07$), which validated the choice of stimuli in Experiment 11 – in the sense that the stimuli only differed by valence and not by modality – so the visual and auditory stimuli had similar levels of valence. For arousal ratings, however, there was a significant interaction effect ($F(1,48)=10.33$, $p=.01$, $\eta^2=0.13$). For the pleasant stimuli, the consonant soundtrack was rated as less arousing than the laughing image ($M_{\text{auditory}}(\text{SE})=3.55(0.23)$, $M_{\text{visual}}(\text{SE})=4.02(0.23)$, $p=.048$). For the unpleasant stimuli, however, the dissonant soundtrack was rated as more arousing than the crying image ($M_{\text{auditory}}(\text{SE})=4.78(0.18)$, $M_{\text{visual}}(\text{SE})=4.33(0.19)$, $p=.035$). In other words, the auditory stimuli seemed to evoke greater differences in arousal compared to the visual stimuli.

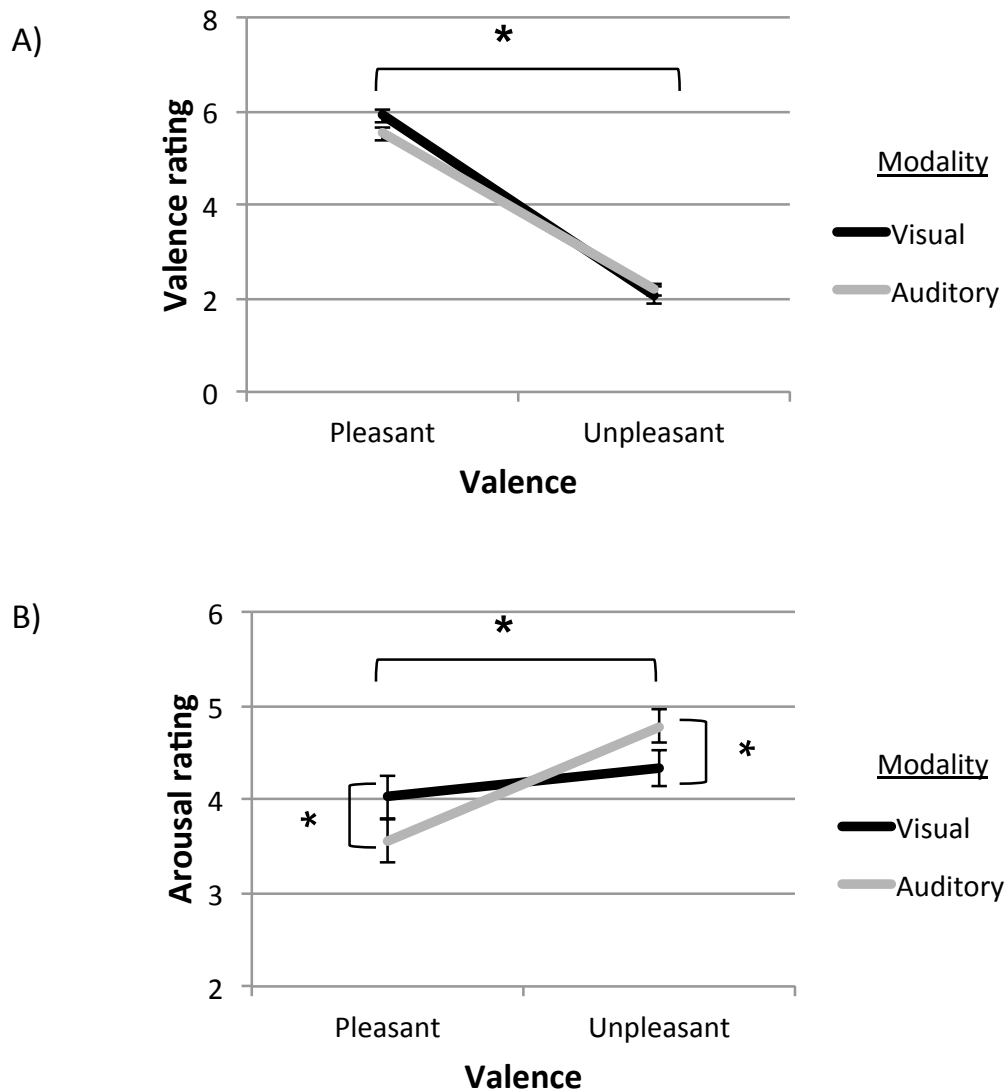


Figure 7.2. Mean values of valence (A) and arousal (B) ratings of the visual and auditory stimuli used in Experiment 11, categorised by modality (visual, auditory) and valence (unpleasant vs pleasant). Error bars indicate standard error. Asterisk ‘*’, indicates statistical significance at $p < .05$.

The aim of Experiment 11 was to assess whether emotionally positive or negative stimuli might influence the eating experience regardless of the sensory modality of the external stimuli. The outcome of the experiment indeed revealed that those participants who tasted the juice mixture while presented with positively-valenced

stimuli rated the juice as sweeter compared to negatively-valenced stimuli, regardless of whether the stimuli were visual or auditory.

The visual stimuli for Experiment 11 were chosen carefully from the IAPS database in order to make sure that the valence levels of the visual stimuli would match those of the consonant/dissonant musical stimuli. The results revealed that this was indeed the case (see Figure 7.2), lending support to the hypothesis that participants were more influenced by the emotion content of the stimuli than by any sensory-modality specific traits.

The fact that emotional valence can influence taste evaluation once again brings into focus emotion transfer as a possible mechanism underlying some part of the auditory modulation of flavour. Experiment 1 showed that taste-congruent soundtracks have very different valence and arousal ratings; for instance, sweet soundtracks generally have much higher valence ratings than bitter soundtracks. In addition, those soundtracks with higher valence ratings tends to be associated (but not always) with increased food liking, and higher sweetness ratings. See Table 7.1 for a summary of studies in this thesis where valence ratings of the sound stimuli were measured.

Experiment #	Increased liking?	Increased sweetness?	Auditory stimuli	Food/beverage
3	Yes	Yes	Melody with consonant and dissonant harmonies	Juice mixture
5	No	Yes	Creamy/rough soundtracks	Chocolate
6	Yes	Yes	Sweet/bitter soundtracks	Chocolate
8	Yes	Yes	Sweet/bitter soundtracks	Chocolate

11	Yes	Yes	Melody with consonant and dissonant harmonies; images with happy/sad child	Juice mixture
13B	No	N/A	Staccato/legato soundtracks	White wines (Canadian Chardonnays)
14	Yes	No	Sweet/bitter soundtracks	Chocolate

Table 7.1. Summary of those experiments in the present thesis where the valence of the auditory stimuli was measured. Columns show whether listening to the more pleasant soundtrack was associated with higher ratings for gustatory stimuli liking and sweetness ratings. This table shows potential evidence of sensation transfer in the experiments.

It is worth noting a few patterns that emerge from visual inspection of Table 7.1. First, the consonant/dissonant melodies used in Experiments 3 and 11 seem to influence both liking and sweetness, which could support the sensation transfer theory. On the other hand, the creamy/rough soundtracks in Experiment 5 influences sweetness ratings, but not liking; and the more liked legato soundtrack Experiment 13B resulted in the wine being enjoyed *less* as compared to the staccato soundtrack. The sweet/bitter soundtracks used in Experiments 6, 8, and 14 enhance sweetness but does not consistently increase liking. What's more, the results of Experiment 6 revealed that the soundtracks enhanced sweetness even when controlling for differences in soundtrack liking. In conclusion, it would appear that soundtracks with clear emotional associations and no other differences (e.g., consonant/dissonant harmonies) operate via influencing participants' feelings about the food/drink they are

about to consume. With soundtracks where emotion attributes are not obviously clear to the participants (especially when the soundtracks are quite different, as with the sweet/bitter ones), the influences of mechanisms like expectations and attention are at least more apparent than sensation transfer *per se*.

Experiment 11 showed that the valence of the stimuli can influence food/drink evaluation regardless of sensory modality, which is especially important for innovators in the restaurant/food industries. It follows that all aspects of the environment – whether packaging or background music or the choice of decorations on the wall – could influence the experience of eating/drinking. Furthermore, the idea the valence of the stimuli can potentially extend to the tactile modality as well. For instance, might the same juice mixture taste sweeter if consumed from a cup with a pleasant texture (e.g., smooth, velvety, or fuzzy) compared to one with an unpleasant texture (e.g., rough, slimy, or sticky)? Finally, the question remains as to whether combining congruent stimuli from multiple sensory modalities could have a more powerful impact on the taste experience. For instance, Spence et al. (2014a) has shown that congruent combinations of coloured lights and music (for instance, red lighting and sweet music) influences the wine drinking experience more than just coloured lights alone. In this case, might there be an additive effect of sweetness enhancement if participants were presented with both consonant music and a happy image?

8 Physiological response as a possible mechanism underlying the auditory modulation of taste/flavour

8.1 Introduction

Chapters 4-8 addressed various cognitive mechanisms that have been suggested to underlie crossmodal modulations of flavour by sound – such as self-report bias, expectations, attention, and emotion. Another possible hypothesis that will be investigated in the present chapter, is a particular version of an embodiment account, whereby people might associate certain soundtracks with certain tastes because the soundtracks induces a similar physiological response in the listener as ingesting foods with that particular taste.

More specifically, this chapter focuses on the hypothesis that people might associate a specific soundtrack with sourness because the soundtrack, much like sour foods, can increase the rate of the listener's resting-state (non-eating) salivary flow. Salivation is a nonconscious physiological process controlled by the autonomic nervous system. Salivation aids in the ingestion process and can influence the perception of tastants in the mouth (see Spence, 2011b, for a review). Salivation can also be induced by conditioned reflexes, such as seeing or smelling appetising foods (Krishna et al., 2014; Wooley & Wooley, 1973), or even by a goal-driven material reward (Gal, 2012). Previous research has shown that while looking at a lemon does not increase salivation (Kerr 1961, Shannon 1974), sniffing or slicing lemons does (e.g., Pangborn, 1968; Pangborn et al., 1979). Moreover, looking at a video of someone else eating a lemon has been shown to induce salivation too (Hagenmuller et al., 2014). Therefore, to check the validity of the experimental methodology, a video of a man eating a

lemon was added as an experimental condition, which was expected to increase salivation (thus demonstrating the sensitivity of the measurement technique).

Furthermore, participants from various previous studies have occasionally commented on the “mouth-watering” effect of high-pitched and dissonant soundtracks which were composed to correspond to sourness. Previously, it has been shown that music can influence the composition of salivation (Suda et al., 2008), with major mode music reducing salivary cortisol levels as compared to music in the minor mode. While there are no studies relating salivary cortisol levels with specific effects on taste perception, those who exhibit higher cortisol level increases due to stress also tend to consume more foods, including more sweet foods, as compared to those who experience lesser cortisol level changes (Epel et al., 2001). Spoken food words have also been shown to increase salivation as compared to non-food words (Staats & Hammond, 1972). Based on these results, it’s possible that a putatively sour soundtrack might enhance the level of salivation in the listener, especially if the latter were to explicitly associate the soundtrack with sourness.

To test this hypothesis, Experiment 12 was designed to measure levels of salivation under different audiovisual conditions. Several methods of saliva collection have been used over the years, including the absorption of saliva by dental cotton rolls, measuring the frequency of swallows, or the electrophysiological measurement of parotid gland activity (Nederkoorn et al., 2001). The method of cotton-roll collection is used in Experiment 12, as it has been shown to provide a reliable, sensitive, and straightforward means of measuring salivary flow (White, 1977).

8.2 Experiment 12: Music to make your mouth water? Measuring the influence of sour music on salivation

8.2.1 Methods

8.2.1.1 *Participants*

36 participants (22 women, 14 men) aged between 18-49 years ($M=23.1$, $SD=6.6$) took part in the study. The participants reported no hearing impairments. The participants were recruited from the Oxford Psychology Research Participant Database and the Experimental Psychology Research Participation Scheme. The study was carried out in accordance with the recommendations of the Central University Research Ethics Committee of Oxford University with written informed consent of all participants. All participants gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Central University Research Ethics Committee of Oxford University (R47262_RE001).

8.2.1.2 *Audiovisual stimuli*

Three audiovisual stimuli were used. The sour soundtrack consisted of a high-pitched and dissonant soundtrack composed by Bruno Mesz that has been shown to reliably correspond to sourness based on previous studies (Kontukoski et al., 2015). In fact, Wang et al. (2015) compared 7 soundtracks composed to correspond to sourness, and the soundtrack by Mesz was labelled as sour, as opposed to any other basic taste, by the largest number of participants (58 out of 100).¹ A silent video of a man facing the camera eating a lemon was used as an additional experiment condition to verify the validity of the saliva measurement methodology, since it has previously been shown

¹ Note that, if participants chose tastes at random, the soundtrack would be labeled as sour by 25 out of 100 people.

to elicit salivation (Hagenmuller et al., 2014). The specific 60-second segment of the video can be viewed at <https://www.youtube.com/v/5FfHSUVBIdw/?start=63&end=123&showinfo=0&rel=0>. Finally, a silent baseline condition (via a soundtrack with the commands “start” and “stop” separated over a 60 second interval) was included as a baseline saliva measure. All three conditions were 60 seconds long. The two audio conditions were accompanied by a visual target (+) for participants to focus on while listening to the soundtracks.

8.2.1.3 Design and procedure

The experiment was conducted at the Crossmodal Research Laboratory at the University of Oxford. Participants were seated at a table in front of a computer monitor with a keyboard, mouse, and headphones in an experimental booth. On the side table were six small plates each with three 8mm dental cotton rolls, a cup of water, and a napkin.

On each trial, the participants were instructed to place three cotton dental rolls in their mouth, two buccally and one under the tongue, then immediately start playing the soundtrack or video. Once the soundtrack or video had finished, the participants were asked to remove the cotton rolls immediately, place them back on the plate, and hand them to the experimenter. Each trial lasted for 60 seconds, and the participants were given a 5 minute recovery period between trials. Each audiovisual condition was played twice (not necessarily successively), thus giving rise to a total of six trials. The order in which the trials were presented was determined using a Williams Design Latin Square in order to minimise any first order carryover effects between trials. The cotton rolls were disposed of immediately after weighing.

After the saliva collection trials, the participants rated which of the four basic tastes (sweet, sour, bitter, or salty; umami was not included as it's not a familiar description to many UK participants) the soundtrack best matched with, and more specifically, how well the sour soundtrack matched with sweet, sour, and bitter tastes on three 1-7 scales (1=does not match at all, 7=matches very well). They also reported their age and gender.

The study lasted for approximately 35-40 minutes. The participants were paid £6 or awarded with course credit for taking part.

8.2.2 Results and discussion

To determine the level of induced salivation, the cotton rolls were weighed before and immediately after each trial, on a balance with 0.0001g precision. The difference between the two weights was used as the amount of induced saliva. The mean weight of induced salivation was then calculated for each audiovisual condition and each participant. The average salivation level for each audiovisual condition is shown in Figure 8.1.

A RM-ANOVA was conducted with the factor 'audiovisual condition' (sour soundtrack, lemon video, silence). In addition, the model included participants' rating of whether they matched the soundtrack with sourness as a between-participants variable, and the interaction term of the audiovisual condition and the between-participants variable. There was a significant main effect of audiovisual condition on salivation ($F(1.82, 61.77)=6.85$, $p=.003$, $\eta^2=.17$), but no significant main effect of soundtrack-sourness match ($F(1,34)<.000$, n.s.) and no interaction effect between the two ($F(2,68)=.17$, n.s.). More specifically, pairwise comparisons with Bonferroni

corrections revealed that more salivation was measured during the lemon video condition ($M_{\text{video}}=0.93\text{g}$, $SD=0.68$) than the silent baseline condition ($M_{\text{silence}}=0.74$, $SD=0.53$, $p=.006$) or the sour soundtrack condition ($M_{\text{soundtrack}}=0.74$, $SD=0.58$, $p=.029$). Contrary to the hypothesis, the sour soundtrack condition did not significantly differ from the baseline condition ($p=1.00$).

Furthermore, the % increase in salivation was computed for each participant while listening to the sour soundtrack compared to the silent condition. Next, Pearson's correlation coefficient between this % increase and how much the participant matched the soundtrack to sourness was calculated, in order to determine whether sensitivity to the soundtrack's intended taste representation influenced their level of salivation. There was no significant correlation between % increase of salivation while listening to the sour soundtrack as compared to silence, and the rating of how much the sour soundtrack was matched to sourness ($r_{36}=-0.11$, $p=.51$). In other words, the extent to which the participants matched the sour soundtrack to sourness is not related to any increase in the amount of salivation while listening to the sour soundtrack.

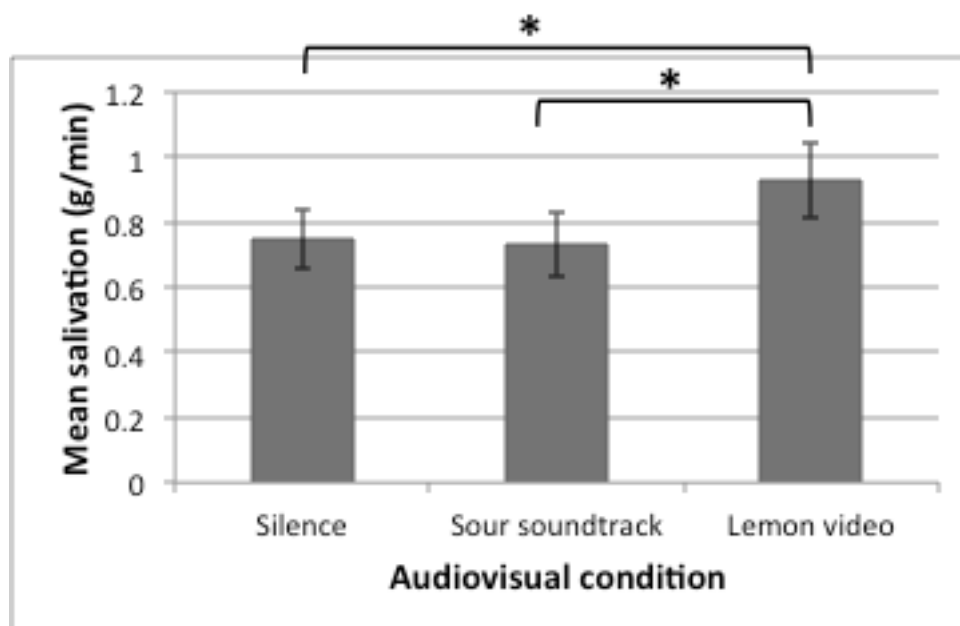


Figure 8.1. Participants' mean salivation (g/min) in all three 60-second audiovisual conditions in Experiment 12. Error bars indicate standard errors. Asterisks denote statistical significance (* $p < .05$).

The results of Experiment 12 revealed that, as reported previously, watching a video of someone else eating a lemon induces salivation compared to the baseline condition, i.e., silent while looking at a fixation cross. This replication of previous results (Hagemuller et al., 2014; see Spence, 2011b, for a review) validates the methodology of using dental rolls to measure salivation. On average, the lemon video condition increased salivation by 0.18g as compared to the baseline condition. This is similar to Hagemuller et al.'s (2014) findings, with 49 participants, where a different lemon-eating video increased salivation by approximately 0.25g over a 60-second period.

However, there was no evidence that listening to the sour soundtrack increased salivation in the participants when compared to the silent baseline condition. This result is especially telling given that Experiment 12 used the most effective – in terms of evoking sourness – soundtrack that has been tested to date (Experiment 2). Perhaps auditory stimulation is simply not sufficient to evoke a physiological response; this may be in-line with previous research which showed that while looking at lemons does not induce increased salivation (Kerr 1961; Shannon 1974), smelling or slicing lemons does (Pangborn 1968 Pangborn et al., 1979). Therefore, like a visual representation of lemons, music alone might not evoke a strong-enough representation of sourness to stimulate a physiological response. In line with this suggestion, it has been theorised that while sound is a part of multisensory flavour perception, it has a relatively weak contribution (Mroczko-Wąsowicz, 2016). The resulting percept is consciously decomposable into its component unisensory parts (e.g., hearing the

sound of crunching makes potato crisps more crunchy, but it is easy to separate the sound of mastication from the flavour of the potato crisp). In contrast, olfaction has a strong (what Mroczko-Wąsowicz termed “constitutive”) contribution which binds with information from the tongue to form a unified flavour perception (Rozin, 1982; Spence et al., 2015c).

The fact that no increase in salivation was reported between the soundtrack condition and baseline condition implies that, contrary to the initial hypothesis, there was no enhancement of salivation due to music, even when participants associated the soundtrack with sourness. While the results of Experiment 12 do not offer any support for this particular version of the embodied hypothesis involving putatively sour soundtracks and salivation, it is still possible that music might act in conjunction with other sensory stimuli to enhance physiological responses to food/drinks. An interesting future study, for instance, could involve comparing the lemon video condition with a combined lemon video plus sour soundtrack condition, in order to assess whether music might act to further enhance salivation.

9 Individual differences in sound-taste/flavour interactions

9.1 Introduction

At this point, the experiments reported in this thesis have shown that a variety of mechanisms could be at work behind how what we hear can influence what we taste (see the experiments reported in Chapters 4-8). What remains unclear, though, on the basis of the research presented thus far, is whether individual differences such as in expertise, taster status, not to mention taste preferences, influence the crossmodal interaction between taste and sound. Experiment 13 addresses the question of wine taster expertise, while Experiment 14 addresses the role of taster status and sweetness liking.

9.1.1 Tasting expertise

In terms of wine expertise, it might at first seem that wine experts would be “immune” to perceptual biases presented by external senses because they are more experienced in the objective sensory evaluation of wine. However, there is anecdotal evidence of winemakers and writers endorsing the effect of music on wine (Gray 2007; Smith 2010; see Spence & Wang, 2015c, for a review). There is also evidence that wine experts’ aroma judgments being biased by the colour of the wine (Parr et al., 2003). On the other hand, if wine experts are more attuned to subtle differences in smell or taste (e.g., through training), then it is conceivable that they would be better able to pick out any subtle changes in the tasting experience, resulting from the way in which music biases the focus of their attention. Wine expertise could therefore act as a moderator on the way people perceive sound-taste correspondences.

Training improves people's ability to discriminate flavours when tasting wine (Owen & Machamer, 1979). However, that is possibly because trained panellists and experts can adapt an analytical tasting strategy that helps them to distinguish different components of wine flavours, when compared to untrained panellists (Arvisenet et al., 2016). Moreover, wine experts use a different vocabulary when verbalising wines, using analytical and source-based terms (like vanilla or cherry) whereas non-experts tend to use holistic terms (Challoet & Valentin, 2000; Croijmans & Majid, 2016). These observations might well lead to the suggestion that wine experts would be better at separating the influence of the music from their sensory evaluation of the wine itself. Furthermore, several neuroimaging studies involving wine tasters have been conducted with the goal of pinpointing the influence of expertise on brain function. Sommeliers activate brain regions involved in high-level cognitive processes like working memory and behavioural strategies when tasting wine – unlike novices who activate the primary gustatory cortex and emotional processing areas more (Castriota-Scanderbeg et al., 2005). In a follow-up study focused on the effect of expertise during the different phases of tasting (i.e., during vs. after tasting), Pazart et al. (2014) observed that wine experts activated those brain regions responsible for multisensory integration immediately during the wine tasting phase, whereas for control participants they were only activated during the after tasting phase. This implies that experts are able to analyse the sensory properties of wine more efficiently than untrained participants by showing a more immediate and targeted sensory reaction to the wine, and by calling upon episodic memory at the same time as they are processing the sensory qualities of whatever is in their mouth.

In terms of odours, expertise has been shown to increase sensitivity and discrimination (see Royet et al., 2013, for a review), possibly giving rise to structural

reorganisation in olfactory brain regions (Delon-Martin et al., 2013). However, there seems to be no evidence that wine experts have increased sensitivity when it comes to wine tasting (that is, there seems to be no differences in sensitivity to odours – either those typically found in wine or otherwise, see Brand & Brisson, 2012; Parr et al., 2002).

Taken together, then, the evidence suggests that while wine experts might have a different way of thinking about, naming, and describing wine, they are no more sensitive to flavours (possibly because wine experts are trained to categorise and look for specific flavours, or combination of flavours, in a wine, and not to distinguish isolated flavours *per se*). In addition, the advantages of wine experts in consistently naming smell and flavours in a wine is domain-specific, and does not seem to extend to beverages outside of wine (Croijmans & Majid, 2016).

Experiments 13A and 13B were conducted on a population of highly-trained wine specialists. The aims were two-fold; first, Experiment 13A replicated the influence of specifically designed taste soundtracks on wine perception (see Reinoso Carvalho et al., 2016, for a similar study using beer). Second, Experiment 13B assessed the effect of sound on more wine-specific characteristics, such as length, balance, and body.

9.1.2 Taster status

Another important consideration when it comes to studying the auditory modulation of taste is the fact that people all live in their own taste worlds (e.g., Bartoshuk et al., 1994), where the same gustatory stimuli may appear to have a different intensity to different people (Bartoshuk et al., 2005). The most prominent example is the fact that people all vary in their ability to taste phenylthiocarbamide (PTC) and its chemical relative, 6-*n*-propylthiouracil (PROP). Individuals are grouped with labels

“nontaster”, “taster”, and “supertaster”, according to their sensitivity to PROP (Bartoshuk, 1991). Supertasters report PTC/PROP solutions as tasting extremely bitter, and tend to be sensitive to strong tastes, especially bitterness (Catanzaro et al., 2013). Taster status can be accounted for, at least in part, by a genetic variation in the *TAS2R38* gene; for a long time, it was believed that PTC/PROP tasting ability was also correlated with the density of fungiform papillae (taste buds) on the tongue (Bartoshuk et al., 1994), although that idea has been lately questioned by the results of at least one large-scale community-based study (Garneau et al., 2014).

With respect to crossmodal perception research, it seems likely that PTC/PROP tasting ability would be an important individual difference to consider. Presumably, those with better taste sensitivity (i.e., those with better ability to taste PTC/PROP) might be better able to distinguish subtle changes in taste of the food/drink due to auditory stimuli. For instance, such supertasters might be extra sensitive to changes in bitterness compared to non-tasters.

With these considerations in mind, Experiment 14 was designed such that participants tasted chocolates at different bitterness levels (70% and 85% cacao). Their PTC tasting ability, as well as their liking for sweet foods, were measured. Individual differences in sweetness liking may be a key to testing the sensation transfer theory discussed in Chapter 7. The theory assumes that the more liked soundtrack (e.g., a sweet soundtrack) might enhance sweetness as compared to the less liked soundtrack (e.g., a bitter soundtrack) because people tend to associate pleasantness with sweetness. What happens, then, for those participants who don't find sweetness pleasant¹? If sweetness enhancement is still observed with those sweet-disliking

¹ As an example showing differences in sweet liking, Crisinel and Spence (2012) showed that those who found dark chocolate pleasant tended to match dark chocolate with piano, while those who found dark chocolate unpleasant tended to match it with string and woodwind instruments.

participants, then it is likely that the mechanism behind sweetness enhancement goes beyond sensation transfer.

9.2 Experiment 13A: Assessing the influence of music on wine perception amongst wine professionals

9.2.1 Methods

9.2.1.1 Participants

154 participants (71 women, 79 men, 4 unspecified) between 20-75 years of age ($M=46.4$, $SD=12.5$) took part in the study. 138 of the participants were professionals working in wine-related occupations, of which 42 worked in education/research, 17 in journalism/wine writing, 13 in retail, 3 in restaurant service, and 63 in viniculture/viticulture. The participants were experienced wine tasters (see Figure 9.1 for histogram), with an average of 18.1 years ($SD=11.5$) of wine tasting experience (137/154 of the participants has over 5 years of wine tasting experience).

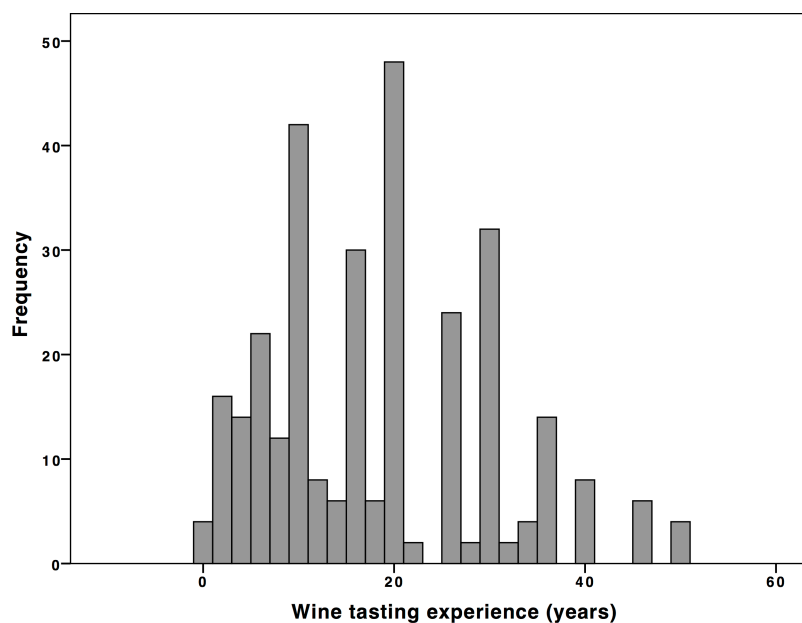


Figure 9.1. Histogram showing the distribution of participants' wine-tasting experience in terms of years, in Experiments 13A and 13B. Tasting experience is binned in intervals of 2 years (so each bar represents a span of 2 years of wine-tasting experience).

The participants gave their informed consent, and reported no impairments of hearing, smell, or taste. The participants were recruited from the International Cool Climate Wine Symposium (ICCWS) 2016, specifically from those who attended the Sensory Evaluation Seminar. The experiment was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

9.2.1.2 Auditory stimuli

Experiment 13A involved soundtracks designed to match with sweet and sour tastes. The two soundtracks have been shown to be reliably associated with sweet and sour tastes, respectively (Experiment 2), and were used recently in a study involving beer (Reinoso Carvalho et al., 2016). All of the soundtracks can be heard at <https://soundcloud.com/janicewang09/sets/iccws-2016>.

9.2.1.3 Wines

Experiment 13A involved two white blends from England - Bolney Lychgate White 2014 and 2015. The 2014 vintage is off-dry with ripe aromas of lychee and passionfruit, with the sweetness balanced by crisp acidity. The wine consists of a blend of Muller-Thurgau and Reichensteiner. The 2015 vintage is a fruity, zesty blend of Reichensteiner, Schonburger, and Wurzer grapes and has rich, ripe scents and flavours of lychee and passion fruit which are balanced by a crisp, spicy acidity and refreshing elderflower notes.

9.2.1.4 Design and procedure

The experiment took place during the Sensory Evaluation Seminar at ICCWS. The participants sat at a table in front of a placemat with four glasses of wine on top, a cup of water, spittoons, and a paper questionnaire. The wines were identified on the placemats by random 3 digit codes. Before the actual study began, the participants specified their gender, age, and whether they worked in a wine-related profession, and years of wine tasting experience.

The sweet soundtrack was played during the first trial, and the sour soundtrack during the second trial. For each trial, the participants were instructed to taste the wine once the soundtrack started. First, they were asked to list the top three flavours they perceived in the wine. Next, they rated the sweetness and acidity level of the wine on 0-10 scales (0=no sweetness/acidity, 10=very high sweetness/acidity) and how much they liked the wine (0=not at all, 10=very much). The participants were instructed to rinse their mouths out with water between wines. Half the participants tasted the Lychgate White 2014 with the sweet soundtrack and the Lychgate White 2015 with the sour soundtrack, the other half tasted the wines in the reverse order.

The study lasted for approximately 10 minutes.

9.2.2 Results and discussion

Figure 9.2 shows the average values of sweetness, sourness, and liking ratings, grouped by music condition and wine type. A multivariate analysis of variance (MANOVA) test with wine (x2) and music (x2) as factors showed a main effect of music ($F(3,282)=10.99$, $p<.0005$, Wilk's Lambda=.90), of wine ($F(3,282)=5.96$, $p=.001$, Wilk's Lambda=.94), and an interaction effect ($F(3,282)=2.69$, $p=.047$,

Wilk's $\Lambda=.97$). Further univariate ANOVAs revealed that music had a significant effect on wine liking (people liked the wine much better while listening to sweet music than to sour music, $F(1,284)=32.97$, $p<.0005$), but not on wine sweetness ($p=.41$) or sourness ($p=.79$). In terms of wine differences, the Lychgate 2014 was rated as significantly sweeter ($p<.0005$) and less sour ($p=0.004$) than the Lychgate 2015, but there were no differences in how much participants liked the two wines ($p=.41$). Importantly, there was a significant interaction effect between music and wine type which was shown in ratings of sweetness and liking. For sweetness, the Lychgate 2014 was rated as sweeter during the sweet soundtrack than while listening to the sour soundtrack ($p=.042$), but not for the Lychgate 2015 ($p=.38$). For liking, the interaction effect comes from the fact that while both wines were liked equally during the sweet soundtrack, when listening to the sour soundtrack, participants preferred the more acidic Lychgate 2015 to the sweeter Lychgate 2014 ($p=.043$).

Moderation analysis was computed using the PROCESS macro (Hayes 2013), using years of wine tasting as a moderator variable. No effect of moderation was found interacting between sound condition and ratings of sweetness ($p=.18$), sourness ($p=.30$), and wine liking ($p=.80$).

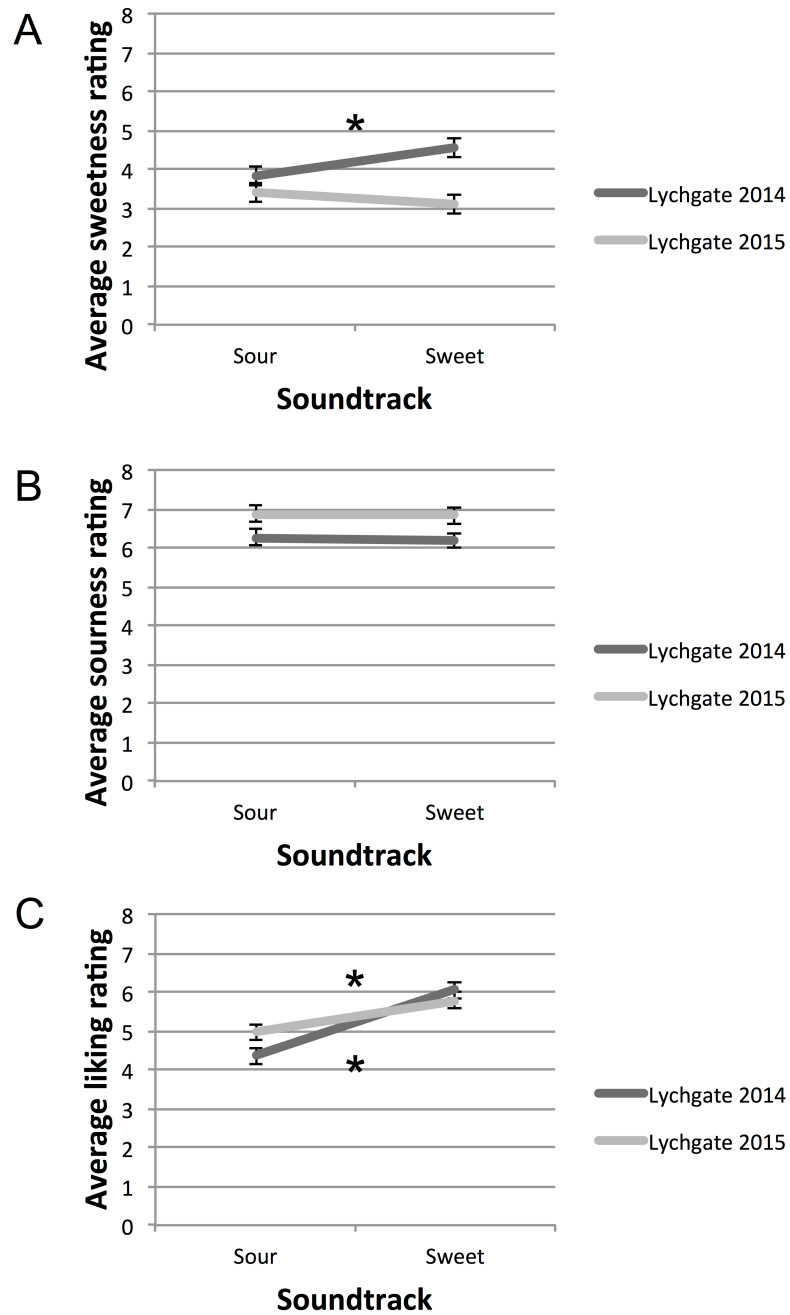


Figure 9.2. Mean values of sweetness (A), sourness (B), and liking (C) ratings in Experiment 13A, with the sound conditions on the x axis and wine type shown as separate lines. Error bars indicate standard error. Asterisk ‘**’, indicates statistical significance at $p < .05$.

Experiment 13A revealed that participants rated the wines as sweeter while listening to the sweet soundtrack compared to the sour soundtrack, as predicted, but only for

the sweeter of the two wines. This finding is in line with observations showing that a secondary sensory stimulus that is congruent with some taste/flavour property can influence the evaluation of a food item mostly when the food item itself already has that taste/flavour property. For instance, Shermer and Levitan (2014) reported that, as redness is associated with piquancy, increasing the intensity of red colour only enhanced the piquancy of medium spicy – and not of mildly spicy – salsa. Similarly, a soundtrack composed to correspond with spiciness enhanced reported spiciness only for medium spicy and not mildly spicy salsas (Experiment 7). In contrast, when it comes to the evaluation of wine liking, both wines were liked more when tasted while listening to the sweet soundtrack, regardless of the sweetness level of the wines. This seems to imply different mechanisms operating for taste evaluation (e.g., the soundtrack may have enhanced expectations, see Tu et al., 2016 for an example of expectations mediating the influence of colour on perceived food spiciness) than for hedonic evaluation (e.g., the participants' liking for the music transferring to their liking for wine).

Interestingly, the more acidic wine (Lychgate 2015) was liked more than the sweeter wine (Lychgate 2014) during the sour soundtrack, which suggests that music-taste congruency could also have played a role in participants' liking ratings. According to the theory of processing fluency (Labroo et al., 2008; Winkielman et al., 2003), the better the match between the music and the wine, the more easily participants could assess the sensory properties of the wine, and consequently, might find the wine more pleasant.

An important limitation of Experiment 13A is that order effects cannot be ruled out as an explanation for the results that were observed. Due to the public nature of the event, the music conditions could not be counterbalanced across participants (even

though the order of wines tasted were counterbalanced). In Experiment 13A, the sweet soundtrack was played first, followed by the sour soundtrack. Therefore, it is possible that the perception of pleasantness decreases with time (hence the wines tasted during the first, sweet, soundtrack were rated to be more pleasant than during the second, sour, soundtrack). Further studies with a counterbalanced order of music presentation will be needed to examine this possible explanation.

Beyond the influence of soundtracks on basic taste attributes, the question remains as to whether sound can affect other more complex features of wine. Given the high level of wine expertise of the participants, it was decided to examine the effect of sound on more wine-specific characteristics, such as length (the duration of aftertaste), balance (how much the different components of wine are in harmony), and body (viscosity).

9.3 Experiment 13B: Assessing the influence of music on wine perception amongst wine professionals

9.3.1 Methods

9.3.1.1 Participants

The same participants who participated in Experiment 13A also took part in Experiment 13B.

9.3.1.2 Auditory stimuli

Experiment 13B involved two abstract soundscapes composed by Ben Houge, a researcher and sound designer specialising in aleatoric music composition (i.e., compositions involving random elements). The first soundtrack is sparsely textured

and staccato, the second soundtrack is less sparse, with overlapping legato woodwind lines. The soundtracks can be heard at <https://soundcloud.com/janicewang09/sets/iccws-2016>.

9.3.1.3 Wines

Both Experiment 13A and 13B involved a pair of white wines. The wines were chosen to be different but similar, so that the participants would not necessarily assume them to be the same (and possibly respond accordingly). They were chosen to have the same colour to control for the effect of colour on the evaluation of the wine.

Experiment 13B involved two Chardonnays from Ontario, Canada. Tawse Quarry Road Organic Chardonnay 2012 and 2013 Speck Family Reserve Chardonnay. Both wines presented fresh acidity, medium alcohol (13%), and had been aged in French oak barrels.

9.3.1.4 Design and procedure

Experiment 13B took place several minutes after Experiment 13A. The participants cleansed their palate with water between studies.

The staccato soundtrack was played during the first trial, and the legato soundtrack was played during the second trial. For each trial, the participants were instructed to taste the wine once the soundtrack started. The wine attributes that the participants evaluated on 0-10 scales were: body (0=very light, 10=very full), balance (0=not balanced, 10=very balanced), length (0=very short, 10=very long, 10+ seconds), wine liking (0=not at all, 10=very much), how well the music matches the wine (0=not at all, 10=very much), and music liking (0=not at all, 10=very much). The participants were instructed to rinse their mouths out with water in between wines. Half of the participants tasted the Quarry Road Chardonnay with the staccato soundtrack and the

Speck Family Chardonnay with the legato soundtrack, and the other half tasted the wines in the reverse order.

The study lasted for approximately 10 minutes.

9.3.2 Results and discussion

A MANOVA with wine (x2) and music (x2) as factors again revealed a significant main effect of music ($F(6,287)=6.49$, $p<.0005$, Wilks' Lambda=.88; see Figure 9.3), wine ($F(6,287)=4.36$, $p<0.0005$, Wilks' Lambda=.92), and an interaction effect ($F(6,287)=3.05$, $p=0.007$, Wilks' Lambda=.94). Further univariate ANOVAs revealed that wines tasted while listening to the staccato soundtrack were rated as significantly fuller in body ($F(1,292)=15.45$, $p<.0005$), more balanced ($F(1,292)=5.82$, $p=.016$), having a longer finish ($F(1,292)=11.66$, $p=0.001$), and were liked more ($F(1,292)=5.30$, $p=.022$). That said, participants liked the staccato soundtrack significantly less than the legato soundtrack ($F(1,292)=5.68$, $p=.018$). In addition, the Speck winery Chardonnay was rated as having a fuller body ($F(1,292)=6.07$, $p=.014$) and a longer finish ($F(1,292)=8.53$, $p=.004$) than the Tawse Chardonnay. Finally, the interaction effect between music and wine is shown in wine liking – the Speck Chardonnay was liked more while participants were listening to the staccato soundtrack as compared to the legato soundtrack ($p=.001$), but no such differences were observed for the Tawse Chardonnay ($p=.96$). All p-values in post-hoc pairwise comparisons have been Bonferroni corrected.

Moderation analysis was computed using the PROCESS macro (Hayes, 2013), using years of wine tasting as a moderator. No effect of moderation was found interacting

between sound condition and ratings of body ($p=.78$), balance ($p=.86$), length ($p=.66$), wine liking ($p=.80$), music-wine match ($p=.35$), and music liking ($p=.80$).

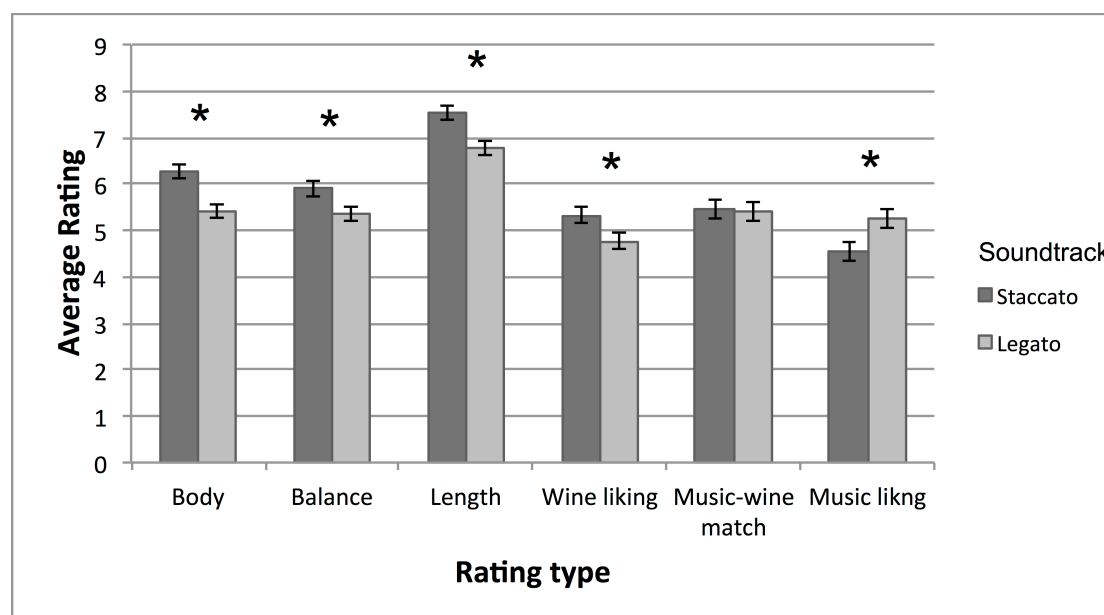


Figure 9.3. Mean values of wine body, balance, length, wine liking, music-wine match, and music liking in Experiment 13B, with the sound conditions shown in different colours. Results are averaged over both wine types. Error bars indicate standard errors. Asterisk ‘*’, indicates statistical significance at $p<.05$. All p -values in post hoc comparison tests have been Bonferroni corrected.

The results of Experiment 13B revealed that the soundtracks influenced wine evaluation significantly in terms of body, balance, length, and wine liking. Somewhat surprisingly, listening to the relatively more disliked music (staccato soundtrack) was associated with fuller body, better balance, and longer length, features that are usually associated with greater wine enjoyment. While this seems counterintuitive at first, it might be possible that the sparseness of the staccato soundtrack provided a better contrast for the oaked Chardonnays and made them seem ‘fuller’ by comparison. In general, having a knowledgeable sample of participants in Experiment 13B meant more sophisticated wine-related attributes could be used. The fact that differences in

the evaluation of these attributes were observed suggests that future studies could focus on the effect of music on more complex and nuanced food/drink characteristics i.e., going beyond the basic tastes.

9.4 Experiment 14: The influence of chocolate cacao content and individual differences on perceptual effects of sound-taste correspondence

9.4.1 Methods

9.4.1.1 Participants

A total of 27 participants (6 women, 21 men) aged 26-73 years ($M=39.82$, $SD=11.40$) took part in the study. The participants were recruited from the London Branch of the Institute of Acoustics. All of the participants gave their informed consent to take part in the study. None of the participants reported a cold nor any other known impairment of their sense of smell, taste, or hearing at the time of the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

9.4.1.2 Auditory stimuli

I used the same sweet and bitter soundtracks as Experiment 6.

9.4.1.3 Gustatory stimuli

70% and 85% Lindt chocolates was used for the study. The two cacao content chocolates were chosen to be distinctly different in taste but with each still being both

sweet and bitter. Each sample consisted of approximately 3g of chocolate. All samples were served on a white Styrofoam plate with 3-digit labels.

9.4.1.4 Design and procedure

The study was designed with cacao content (70% or 85%) and soundtrack type (bitter or sweet) as within-participant factors. Each participant completed four trials, tasted each chocolate twice and heard each soundtrack twice.

At the Institute of Acoustics, participants rated the chocolates on paper forms instead of using a computer. People were divided into two groups, one group experienced 70% cacao chocolate + bitter soundtrack, 70% cacao chocolate + sweet soundtrack, 85% cacao chocolate + sweet soundtrack, 85% cacao chocolate + bitter soundtrack, and the other group experienced 85% cacao chocolate + bitter soundtrack, 85% cacao chocolate + sweet soundtrack, 70% cacao chocolate + sweet soundtrack, and 70% cacao chocolate + bitter soundtrack. In other words, for each group, participants experienced both order of soundtracks.

Finally, the participants listened to each soundtrack again. For each soundtrack, they rated which basic taste (sweet, bitter, sour, or salty) best matched the soundtrack. Finally, the participants rated how much they enjoyed eating sweet foods on a scale from 0-10 (0=hate it, 10=love it), and took a PTC taste strip test to determine their taster status. They were asked to report which taste was most detectable on the strip (sweet, bitter, salty, sour) and how intense was the taste on a scale from 0-10 (0=no taste, 10=extremely intense).

The experiment lasted for around 10 minutes and participants were debriefed afterwards.

9.4.2 Results and discussion

The mean values of the participants' taste and liking ratings for the chocolates for both cacao contents and both soundtrack conditions are shown in Figure 9.4. A RM-MANOVA was conducted with 'soundtrack type' and 'cacao content' as the within-participant factors. The model included bitterness and liking ratings as measures; strip taste intensity, sweetness liking, and liking for each soundtrack were analysed as covariates. Overall, the RM-MANOVA revealed that, controlling for strip taste intensity, sweetness liking, and liking for each soundtrack, there were main effects of soundtrack type ($F(2,21)=3.60$, $p=.045$, Wilks' Lambda=.74) and cacao content ($F(2,21)=24.13$, $p<.0005$, Wilks' Lambda=0.30), but no significant interaction effect between the two ($F(2,21)=0.83$, n.s.). Follow-up univariate tests revealed that for soundtrack type, there were significant differences between taste ratings while listening to the bitter soundtrack compared to the sweet soundtrack ($F(1,22)=7.11$, $p=.014$, $\eta^2=0.24$). Comparing the estimated marginal means showed that the chocolates were rated as tasting sweeter while listening the sweet soundtrack compared to the bitter soundtrack ($M_{\text{sweet_soundtrack}}(\text{SE})=5.00(0.22)$, $M_{\text{bitter_soundtrack}}(\text{SE})=4.33(0.21)$, $p=.014$).

For cacao type, as expected, 70% cacao chocolate was rated to be more sweet than 85% cacao chocolate ($F(1,22)=44.96$, $p<.0005$, $\eta^2=0.67$; $M_{70}(\text{SE})=5.87(0.28)$, $M_{85}(\text{SE})=3.46(0.22)$) and liked more than 85% cacao chocolates ($F(1,22)=14.29$, $p=.001$, $\eta^2=0.39$; $M_{70}(\text{SE})=6.54(0.24)$, $M_{85}(\text{SE})=5.39(0.39)$).

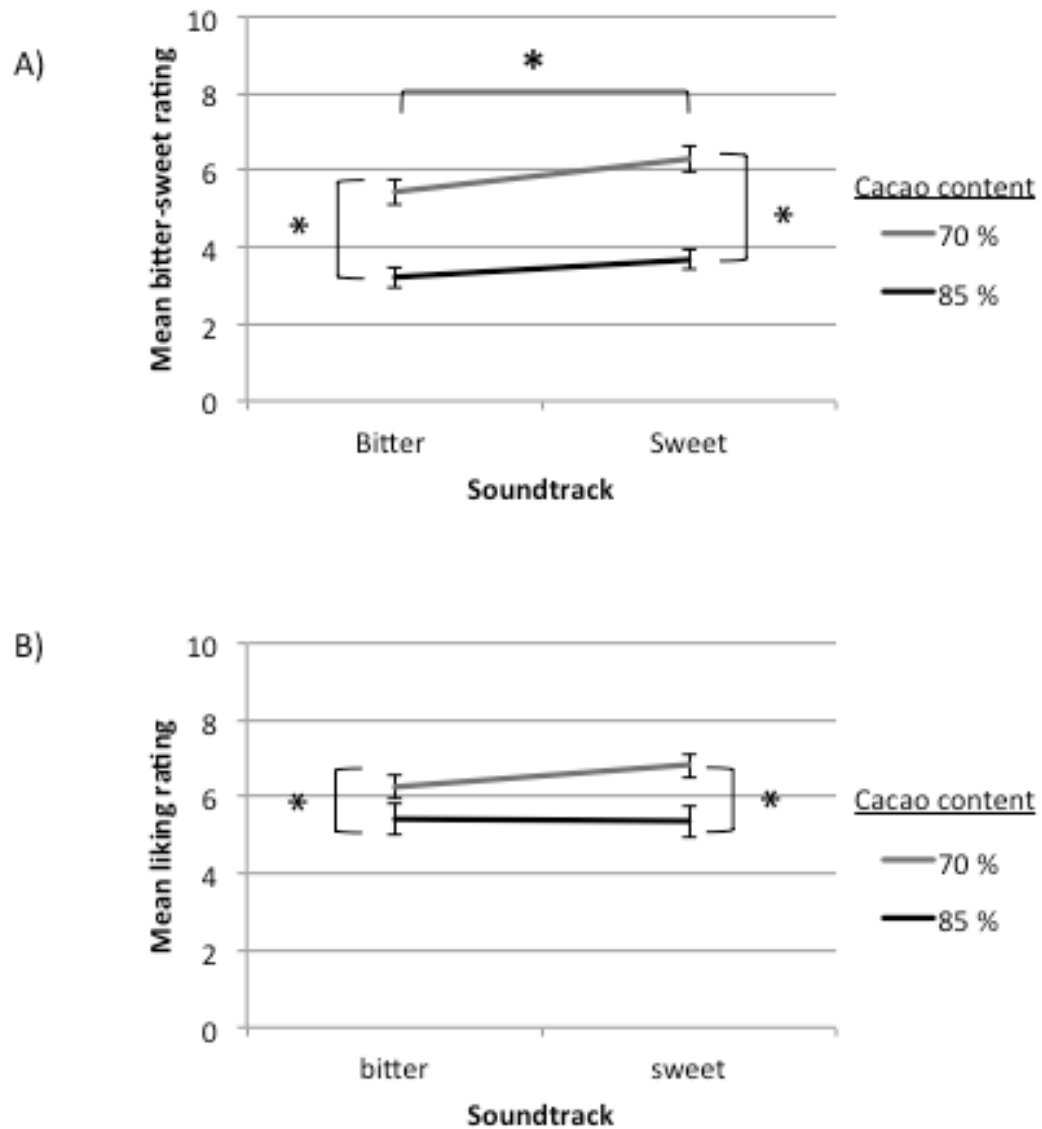


Figure 9.4. Mean values of bitter-sweet (A) and liking (B) ratings for chocolates, for both cacao content (70%, 85%) and soundtrack conditions (sweet, bitter) in Experiment 14. Error bars indicate the standard error of the means. Asterisk ‘*’, indicates statistical significance at $p < .05$.

The covariate, tasting strip intensity, was significantly related to participants’ taste ratings ($F(1,22)=5.22$, $p=.032$). Additional analysis revealed a borderline significant correlation between strip intensity and average taste ratings, $r(27)=-.36$, $p=.06$. This suggests that the more intense the participants found the tasting strip, the more bitter

they rated the chocolates. This is in line with previous research showing that supertasters are more sensitive to bitterness than non-tasters (Bartoshuk et al., 1994).

For a more in-depth analysis of PTC tasting ability, see Figure 9.5. It presents a histogram of participants' taste strip intensity ratings, showing a clear bimodal distribution. Based on this distribution, participants were divided into two groups, namely low and high taste sensitivity (with the dividing value being 4 out of 7). The difference in taste ratings between the bitter soundtrack and sweet soundtrack were calculated for each type of chocolate. A Bonferroni-corrected comparison of these taste rating differences between the two groups (see Figure 9.6) showed that those in the high-sensitivity group made bigger differences in taste ratings between the two soundtracks while tasting 85% chocolates, compared to those in the low-sensitivity group ($M_{\text{low_sensitivity}}(\text{SE})=-0.42(0.47)$, $M_{\text{high_sensitivity}}(\text{SE})=1.20(0.42)$, $p=.016$). There was no significant difference in soundtrack-based taste differences between the two taste sensitivity groups for the 70% chocolate ($M_{\text{low_sensitivity}}(\text{SE})=0.83(0.57)$, $M_{\text{high_sensitivity}}(\text{SE})=0.87(0.51)$, $p=.97$). In other words, while the two sensitivity groups behaved similarly while tasting the 70% chocolate, when it came to the more bitter 85% chocolate, the high sensitivity group appeared to be more influenced by the different soundtracks than the low sensitivity group.

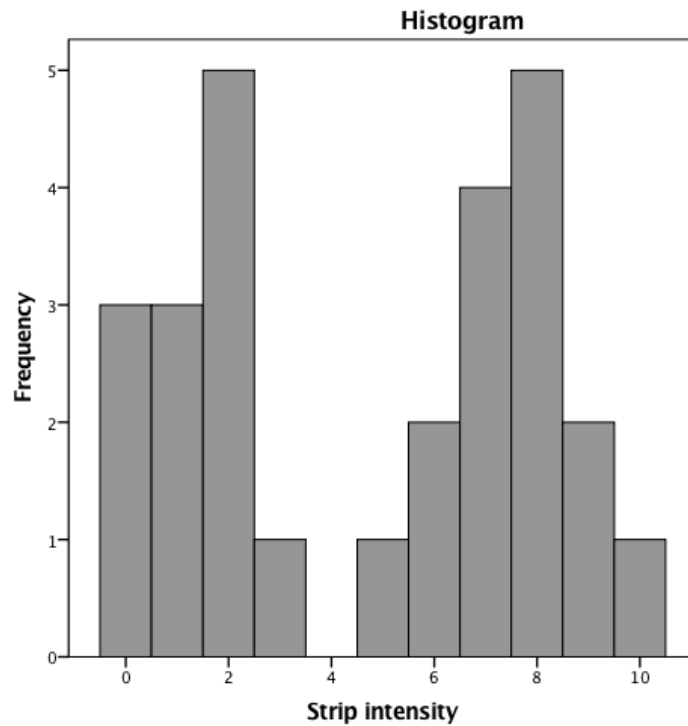


Figure 9.5. Histogram showing the distribution of participants' PTC taste strip intensity ratings in Experiment 14, on a scale of 0-10 (0=no sensation, 10=extremely intense).

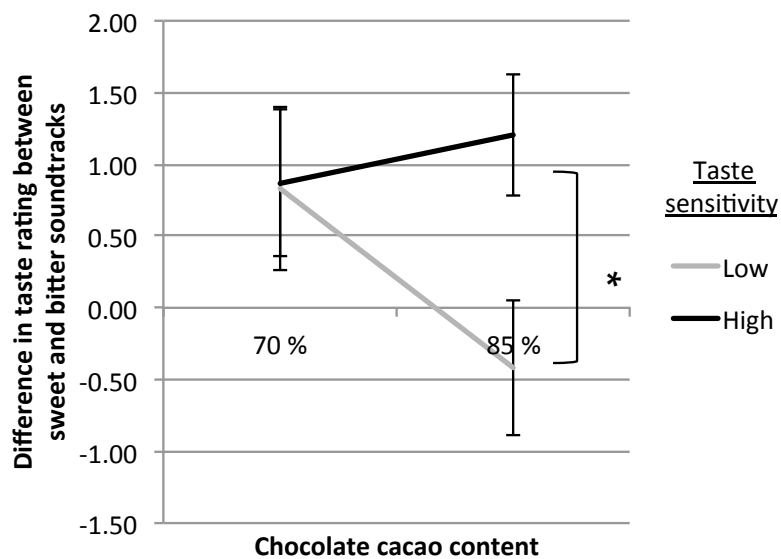


Figure 9.6. Mean values of differences in taste ratings between the two soundtrack conditions ($Taste_{sweet_soundtrack} - Taste_{bitter_soundtrack}$), for each taste sensitivity group (low, high) in Experiment 14. Error bars indicate the standard error of the means. Asterisk "*", indicates statistical significance at $p < .05$.

Finally, in terms of participants' sound-taste association matches, 10/27 participants matched the bitter soundtrack with bitterness, 8/27 with salty and 9/27 with sourness. For the sweet soundtrack, 23/27 participants matched it with sweetness, 2/27 with bitter, 1/27 with salty, and 1/27 with sour. Chi square tests of goodness of fit revealed both distributions were significantly different from chance ($X^2_{\text{bitter_soundtrack}}(3)=9.30$, $p=.026$; $X^2_{\text{sweet_soundtrack}}(3)=52.26$, $p<.0005$). In addition, participants rated the sweet soundtrack as significantly more pleasant than the bitter soundtrack ($M_{\text{bitter}}(\text{SE})=3.89(0.48)$, $M_{\text{sweet}}(\text{SE})=7.04(0.31)$, $p<.0005$).

The aim of Experiment 14 was to explore the role of individual differences on the perceptual effects of sound-taste correspondences, more specifically, focusing on taster status, sweetness liking, and soundtrack liking. The results first reaffirmed the fact that sweet and bitter soundtracks did indeed influence participants' bitter/sweet ratings of chocolates, even when controlling for these individual differences. This demonstrates that individual differences do not dominate the effect of soundtracks (although there were not enough non-sweet-likers to make any conclusive statements).

That said, sensitivity to the PTC taste strip did have a significant relationship with participants' taste ratings, specifically in the bitter soundtrack + 85% cacao chocolate condition. Presumably, the combination of bitter soundtrack and dark chocolate induced a bitterness level where taste perception clearly varied depending on taster status. With less bitter chocolates, or when the dark chocolate was paired with a sweet soundtrack, the correlation between taste sensitivity and taste rating was not significant. This bifurcation in bitter taste ratings is what putatively drove the differences in auditory taste modulation, where the high taste sensitivity group found

a bigger difference in the taste of the 85% chocolate between the bitter soundtrack and sweet soundtrack, as compared to the low taste sensitivity group. This result provides an example where, given the same foods, people might experience the influence of taste-inspired soundtracks differently, depending on their individual taste worlds.

9.5 General discussion and conclusions

Experiments 13 and 14 revealed that individual differences play a role in sound-taste interaction under select circumstances. While there was no evidence in Experiment 13 that wine-tasting expertise moderated the impact of music on wine evaluation, the expert participant pool showed that music could influence wine-specific terminology (compare with Experiment 4, where only differences in acidity and fruitiness were observed with a participant pool of novice tasters). On the other hand, Experiment 14 demonstrated that supertaster status made a difference in how soundtracks influenced chocolate ratings, but only for bitter-tasting foods (e.g., 85% cacao chocolate).

In both Experiment 13A and 13B, it can be seen that experienced wine tasters are also influenced by background music when it comes to wine evaluation. Moreover, years of wine tasting experience do not moderate the impact of music on wine evaluation. One might argue that the results observed here are in-line with previous studies showing that wine experts exhibit no differences in olfactory sensitivity from novices (Brand & Brisson, 2012; Parr et al., 2002) – in other words, wine experts are not *unique* in their ability to perceive qualities in wines *per se*, but maybe better in their ability to identify and categorise flavours found in wines. In this case, it is not surprising that wine experts can also be influenced by music, since music only seems

to enhances/detracts certain flavours of the wine without creating new ones. Furthermore, it is worth pointing out that Experiment 13 involves a large number (154) of very experienced wine tasters (with an average of 18.1 years of experience), which makes this a unique study in the area of music-wine correspondences.

In addition, it is also worth noting that both Experiment 13A and 13B involved pairs of white wines. They were chosen to have the same colour in order to control for the effect of colour on the wine evaluation, since most previous studies involving wine and music had used wines of different colours (e.g., Spence et al., 2013; Experiment 4). The issue here is that the colour of the wines, if they were different, may have entered into the crossmodal matching process (cf. Palmer et al., 2013; see Spence, 2015e, for a review of impact of food colour on taste). Similarly, the chocolates used in Experiment 14 were also of the same colour, to avoid participants making taste judgments based on visual cues (for instance, lighter coloured chocolates, e.g. milk chocolates, tend to be sweeter than dark coloured chocolates).

In terms of practical applications, these results demonstrate how subtly individual differences might influence a particular music and food experience. In order to construct such an experience, all aspects of the design – from the music through to the food/drink and the participants' particular taste worlds – should be taken into consideration. Going forward, Experiment 13 demonstrated that music-wine experiences can be designed for those at various knowledge levels, not just novice drinkers. Furthermore, a quick supertaster status test might be in order at the beginning of future multisensory dining experiences, especially when chefs venture into the realms of taste where there might be a sharp divide in the population (e.g., bitter foods).

10 General discussion: Conclusions, limitations, and directions for future research

Over eight experimental chapters and 14 experiments, the work presented in this thesis has demonstrated a range of modulatory effects of audition on taste/flavour on various different foods and beverages. Importantly, six out of the eight chapters focused on assessing the different possible mechanisms and moderating factors underlying the effects of sonic seasoning. In Section 10.1, these findings are summarised in light of existing theories and a model for how auditory crossmodal correspondences may influence multisensory flavour processing is introduced. In Section 10.2, limitations of the experiments presented in this thesis and potential directions for future study are discussed, including two case studies with real-world applications. Namely, Section 10.2.2 addresses questions around the practicalities associated with the use of sonic seasoning in real-life dining situations while Section 10.2.3 explores how taste-congruent soundtracks can potentially be used to encourage healthier eating behaviours. Finally, Section 10.3 draws together some concluding remarks concerning sound-taste correspondences that have emerged from the research reported in this thesis.

10.1 Summary of findings

After reviewing the literature on crossmodal correspondences and multisensory flavour perception in Chapter 1, Chapter 2 explored the role of emotion mediation in the formation of sound-taste correspondences. The results of Experiment 1 revealed

that participants' choices of auditory pitch to match basic taste solutions were partly mediated by the emotional valence (pleasantness) and arousal ratings associated with the music and taste stimuli. In Experiment 2, the scope of the auditory stimuli was extended from single musical notes to pieces of soundtracks; an analysis of 24 different taste-corresponding soundtracks that had been composed by various researchers and designers demonstrated that valence and arousal once again played a role in how participants matched basic taste words to soundtracks.

In Chapter 3, a set of experiments highlighted the effects of taste/flavour-congruent soundtracks on the perception of fruit juice, wine, and chocolates. Experiment 3 demonstrated that varying the consonant/dissonant harmonies of a melody can alter the rated sweetness/sourness of juice samples. The results of Experiment 4 verified that people can consistently match certain wines and pieces of music, and that listening to the latter can significantly influence the perceived acidity and fruitiness of a wine. Experiment 5 demonstrated that soundtracks designed to correspond with roughness and creaminess can influence the perceived mouthfeel of chocolates.

Chapters 4-8 then explored mechanisms underlying perceptual effects of sound-taste correspondences. The results of the study reported in Chapter 4 revealed that the timing of soundtrack presentation has a significant influence on the sonic seasoning effects that are observed. More specifically, in Experiment 6, sweet and bitter soundtracks influenced the participants' sweet/bitter ratings of chocolates, but only when the soundtracks were played *while* tasting. When they were played *after* tasting, i.e., during evaluation, no similar crossmodal modulation of taste was observed. This result demonstrates that the soundtracks were genuinely influencing people's taste experience in real time, and not just acting by altering their taste memories, say.

In Chapter 5, the ability of taste/flavour-congruent soundtracks to modify sensory expectations was assessed via two experiments. First, Experiment 7 provided evidence that a spicy soundtrack could enhance both people's expectations about the spiciness of a dish as well as the perceived spiciness of the food itself. Specifically, the results of Experiment 7 demonstrated that the influence of the spicy soundtrack on participants' spiciness ratings could be explained via an assimilation-contrast model of expectation disconfirmation. Participants' spiciness ratings for salsa samples were enhanced when listening to a spicy soundtrack (compared to silence) only when the spiciness level of the salsa was fairly close to what was expected. No spiciness enhancement by music was observed for the mildly spicy salsa, possibly because the difference between the expected and actual spiciness levels was too great. The results of Experiment 8 further supported the theory that playing the appropriate soundtracks can bias sensory expectations by demonstrating that sweet and bitter soundtracks can influence the sweet/bitter rating of chocolates, regardless of whether the soundtracks are heard only before tasting or only during tasting.

The role of attention was investigated in the two experiments reported in Chapter 6. In Experiment 9, cognitive load in the form of a memory task was introduced as a within-participants variable in a sonic seasoning study involving bitter and sweet soundtracks. The soundtracks influenced the participants' taste evaluation regardless of the level of cognitive load (high or low) and sound condition, thus not supporting the hypothesis that high cognitive load might reduce the influence of sonic seasoning. The results of Experiment 10, however, provided evidence that, as in the case of sounds orienting people's spatial attention (e.g., Chiou & Rich, 2012; Orchard-Mills et al., 2016), different soundtracks can draw people's attention to different

taste/flavour attributes in a wine (Experiment 10A), as well as modify the perception of specific tastes in real time (Experiment 10B).

The experiment reported in Chapter 7 then extended Chapter 2's focus on emotion mediation between sound-taste correspondences in order to evaluate whether the emotional valence of what we hear can influence what we taste. Namely, Experiment 11 demonstrated that the valence of external stimuli, both music and images, can influence participants' taste evaluation of beverages. A juice mixture was rated as sweeter when participants were exposed to pleasant stimuli (a smiling image of a child or consonant music) as compared to unpleasant stimuli (an image of a crying child or dissonant music).

The goal of the experiment described in Chapter 8 was to investigate whether taste-congruent soundtracks can evoke the same kind of physiological reactions in the listener as the taste that it is matched with. Experiment 12 focused specifically on the known fact that ingesting (or even priming thoughts of ingesting) sour tastes can enhance salivary flow (Pedersen et al., 2002). Contrary to the hypothesis, however, the results of Experiment 12 provided no evidence that a putatively sour soundtrack can increase people's salivation.

Finally, the experiments reported in Chapter 9 explored individual differences in sonic seasoning. Experiment 13 demonstrated that wine-tasting expertise did not moderate how music influenced wine perception, either in terms of basic taste qualities (sweet/sour) or in terms of more complex wine-tasting terminology (e.g., balance, length). In contrast, the results of Experiment 14 revealed that taster status – another source of individual variability – influenced the degree to which participants' chocolate taste ratings were influenced by bitter and sweet soundtracks. Specifically, those with higher sensitivity to PTC reported a greater auditory modulation of taste

with the bitter (85% cacao) chocolate, as compared to those with a lower sensitivity to PTC.

Taken together, then, the results of the various experiments reported in this thesis demonstrate that taste-congruent soundtracks can influence the perception of taste/flavour on multiple levels. Such soundtracks can shape people's expectations about the food that they are about to consume, focus their attention towards different tastes/flavours, and induce them to transfer their feelings about the soundtrack to the food that they happen to be tasting.

10.2 A framework for how crossmodal correspondences influence multisensory flavour perception

Based on the summary of findings reported in Section 10.1, it would appear that taste-congruent soundtracks could potentially influence the tasting experience at various points in time, including before and during tasting. A parallel can be drawn to the role of food-intrinsic sounds, which act to shape our expectations concerning the food we are about to eat (e.g., the sound of the steak sizzling on the hotplate, see Wheeler, 1938) as well as influence food evaluation during eating (e.g., mastication sounds). Similarly, taste-congruent soundtracks can modify people's sensory expectations about the food that they are about to consume. They can also direct people's attention towards specific tastes/flavours in the food, and prime people with specific emotions that can consequently affect the tasting experience. Figure 10.1 demonstrates this

framework for how such soundtracks may influence multisensory flavour perception and evaluation¹.

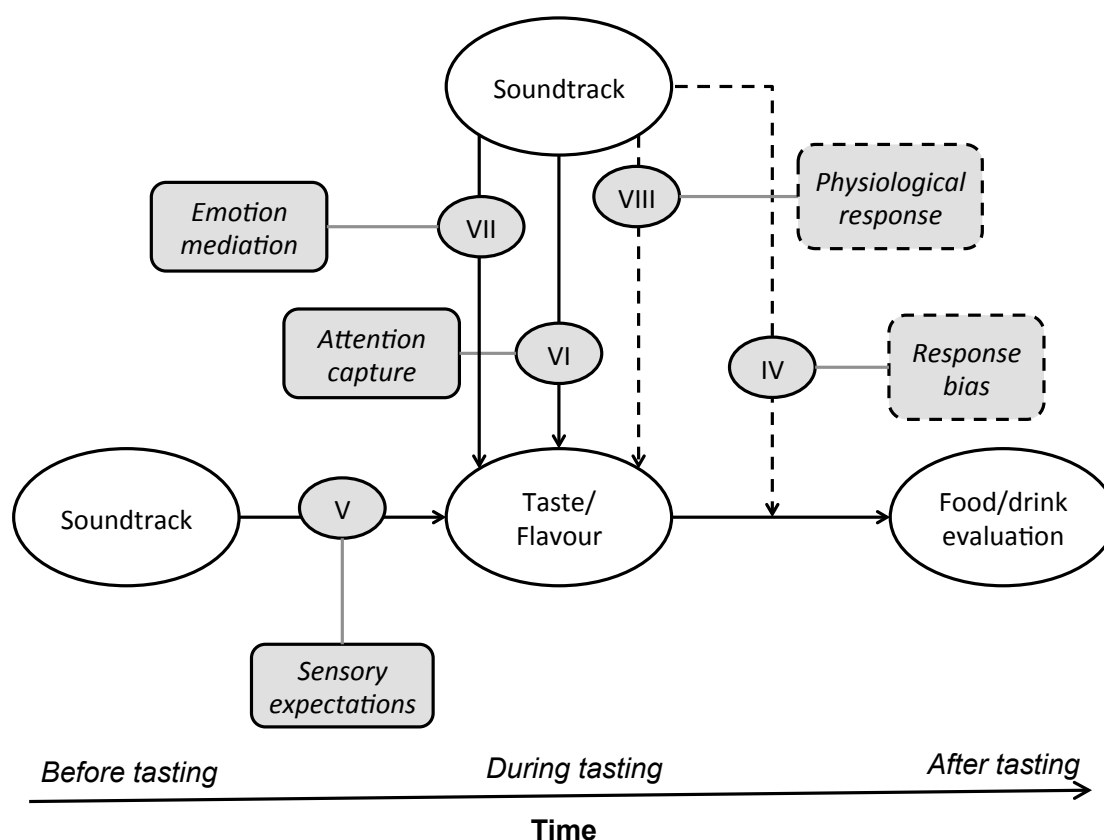


Figure 10.1. Schematic summary of the various ways in which sound might influence taste/flavour evaluation at different points in time. The relevant chapters addressing the possible link are shown in Roman numerals. Dashed lines denote mechanisms for which no evidence was found in the studies reported in this thesis, whereas undashed lines denote those mechanisms garnering empirical support from the results reported in the various experiments reported in this thesis.

It is important to note here that sound-taste correspondences are most likely a key element in both endogenous and exogenous attentional capture. According to the endogenous account, crossmodal correspondences can influence top-down attentional capture in terms of how sounds are decoded into sensory (taste) expectations, such

¹ Note here that synaesthesia was not considered as a possible mechanism, for the reasons explained in Section 1.4.3.

that hearing a sweet soundtrack might bias us to expect sweet tastes in the food we are about to eat. According to the exogenous account, correspondences can also affect bottom-up attentional capture in terms of how sound can draw attention to different components of taste/flavour experience (see Figure 10.2 for a possible model).

Unlike in the case of attentional capture, sound-taste correspondences may or may not play a role in the emotion mediation mechanism (Chapter 7). Most likely, background sounds could modulate the emotional state of the listener, who might then be expected to view the gustatory stimuli in the same manner². For instance, those participants who listened to the consonant music sample in Experiment 11 might have experienced positive feelings in response to the music. Consequently, they might have viewed the simultaneously-tasted juice sample more favourably, which, in turn, might have enhanced the evaluation of pleasant (sweet) tastes in the sample.

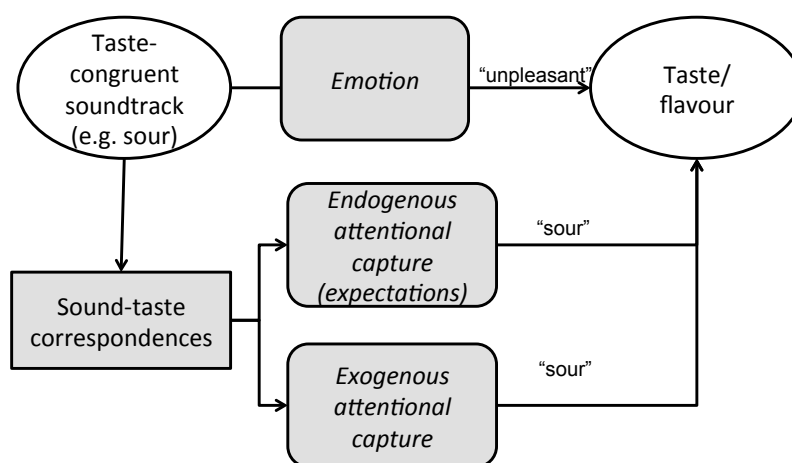


Figure 10.2. Various ways in which a sour taste-congruent soundtrack might influence taste/flavour perception. Note that sound-taste/flavour correspondences are likely to play a role in some pathways, such as the setting of expectations of sourness and focusing one’s attention on

² There is a subtle point to be made here about the influence of emotions. The results from the experiments reported in Chapter 2 demonstrated that emotions partially mediate pitch-taste correspondences. However, this doesn’t necessarily mean that crossmodal correspondences are involved in the emotion mediation pathway as described in Chapter 7.

the sour tastes in the food. The influence of emotion on the sonic seasoning effect, on the other hand, may or may not explicitly involve crossmodal correspondences (see footnote 10.2).

Of course, looking at the bigger picture, there is no reason to believe that the framework proposed in Figure 10.2 is necessarily unique to sound. Other non-chemosensory senses such as vision and touch likely operate on a similar basis in terms of influencing our evaluation of food. However, it should be noted that visual and tactile cues like colour, shape, and texture are often attributes of the food itself (i.e., they can often be considered to constitute ‘product intrinsic’ cues, see Szybilo & Jacoby, 1974), whereas auditory cues, specifically background sounds, are not intrinsic properties of the food itself (that is, they are ‘product extrinsic’). In a sense, background sound may influence taste/flavour in a similar way as other product-extrinsic cues such as the colour, shape, and texture of packing and servingware. For example, it is well known that the colour and shape of the food as well as the container in/on which the food is presented can alter people’s expectations of the food they are about to eat (e.g., Clydesdale et al., 1992; Fairhurst et al., 2015; Johnson & Clydesdale, 1982; Morrot et al., 2001; Piqueras-Fiszman et al., 2012; Zampini et al., 2008; see Spence et al., 2010, and Spence & Deroy, 2013c, for reviews). An interesting future study would be to conduct a visual analogue of Experiment 10B, where the colour of a food item is changed in real time as it is consumed (possibly using a similar projection set-up to that reported by Nishizawa et al., 2016). This hypothetical study would then assess whether the change in colour might trigger, in real time, a change in perceived flavour. Granted, such a setup might work against people’s unity assumption that the same food does not change taste midway through (see Woods et al., 2010), but note here that some food/drinks (especially wine) do, in

fact, evolve and change in the mouth and the glass³ (Fielden, 2009). In terms of touch, studies involving the impact of tactile stimuli on the perception of food are beginning to emerge. Some have focused on the texture of servingware (Biggs et al., 2016), while others have examined the interaction between the perceived texture of the food and of the servingware (Piqueras-Fiszman & Spence, 2012). However, it remains to be seen whether touch/texture-taste correspondences can modify sensory expectations as well as actual taste evaluations.

According to the flavour perceptual system proposed in Section 1.3.3, background (product-extrinsic) sounds, colours, shapes, and textures can all potentially influence the perception of taste/flavour. This is possibly because the brain is designed to gather as much information as possible about the food that we are about to eat, therefore all the senses can influence the tasting experience (see Figure 10.3). Extrapolating from the results of this thesis, it is possible that each sensory input can heighten our awareness in a variety of other sensory attributes in a spreading activation model. Spreading activation models have traditionally used to in cognitive psychology as a way to represent how mental processes retrieve information, particularly in semantic memory networks (Anderson, 1983; Collins & Loftus, 1975; Foster et al., 2013; see also Foster et al., 2016, for evidence of spreading activation in non-verbal memory networks). Concepts are modelled as nodes in a network, and two related nodes/concepts are connected by an edge whose length reflects the strength of their shared associations (Collins & Loftus, 1975). Applied in the context of sensory processing, it is plausible that the brain maintains a network of sensory features that are interconnected via crossmodal correspondences.

³ The fact that wines often evolve in the mouth might explain why Experiment 10B revealed significant differences in sweetness/acidity ratings in the same mouthful of wine over time.

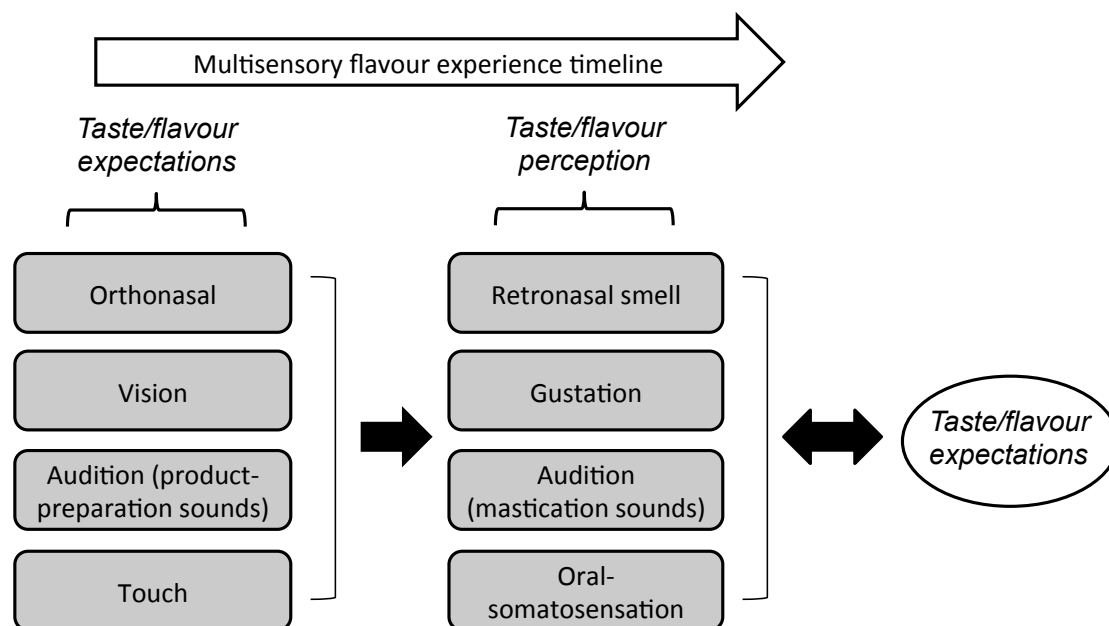


Figure 10.3. A timeline illustrating how the different senses affect the eating/drinking experience at different points in time. The straight black arrow indicates that taste/flavour expectations can influence taste/flavour perception (Piqueras-Fizman & Spence, 2015; Yanagisawa, 2017). The double-headed arrow signals the complex interaction between expectations and perception, since taste/flavour expectations are continually being updated even during the process of eating/drinking.

10.3 Limitations and directions for future research

10.3.1 Limitations

10.3.1.1 Experimental design

An overarching limitation of the majority of the experiments reported in this thesis (with the exception of Experiment 12) is that they require the participants to eat or drink something. Trials with a tasting component tend to be time consuming and fatiguing (cf. Amerine et al., 1965), in comparison to more traditional audiovisual psychophysics studies in which the participants can run through hundreds, if not thousands, of trials at a time. Due to this limitation, the experiments reported in this thesis usually involved a small number of trials (under 12), which can obviously give

rise to statistical power issues and may yield less robust results as compared to other cross-sensory studies. Furthermore, taste/flavour ratings can be very subjective (e.g., Bartoshuk et al., 1986, 2005; O'Mahony, 1973; O'Mahony et al., 1979), and most casual participants (like the University recruits in many of the experiments reported here) who are not specifically trained on tasting panels might have trouble making consistent ratings across trials (e.g., Clapperton & Piggott, 1979; Meiselman & Dzendolet, 1967; Robinson, 1970). Notably, most of the experiments (with the exception of Experiment 12, where salivation was measured via cotton rolls), involved participants giving self-reported taste ratings. In the future, studies that either measure neurophysiological response (e.g., Woods et al., 2011) or some downstream behavioural consequence (e.g., Kontukoski et al., 2013) would presumably reduce inconsistencies in participant self-report⁴.

Furthermore, the majority of the experiments reported in this thesis (with the exception of Experiment 7B) required the participants to make explicit comparisons between distinctly different music. Beyond this thesis, all of the sound-taste experiments reviewed in Section 1.2.2.1 (with the exception of North's, 2012, study) also used within-participants designs. Therefore, it remains unclear whether similar effects would be observed in a between-participants scenario where each participant is only exposed to one sound condition. Of the experiments that used a between-participants experimental design, Experiment 7B reported a difference in the expected spiciness of a dish between those who listened to a spicy soundtrack, as compared to those who listened to a sweet soundtrack, white noise, or silence. However, there were no significant differences between the groups when it came to the actual taste of the dish. Elsewhere, North (2012) revealed a significant effect of music condition on

⁴ On the other hand, physiological measurements would induce additional noise in the data.

wine tasting in a between-participants setting. Those participants who listened to a piece of music that could be associated with a specific characteristic (powerful and heavy, subtle and refined, zingy and refreshing, or mellow and soft) while tasting a wine tended to rate the wine higher in that same category, as compared to those who listened to one of the other soundtracks. For example, the participants who listened to a powerful and heavy soundtrack rated the wine higher on the powerful and heavy scale as compared to those participants who listened to other soundtracks. It should, however, be noted that North's study can be viewed as a case of the metaphorical transfer of abstract ideas, rather than necessarily a genuine perceptual effect *per se* (Spence & Deroy, 2013b). The difference being here that, while ratings of *descriptive* attributes of the wines (e.g., powerful and heavy) were altered in different musical conditions, there was no evidence that the *flavour* evaluation of the wines were actually affected⁵.

Granted, as crossmodal correspondences generally tend to involve relative mappings (see Section 1.1.4.2), it would not be surprising if within-participant experimental designs involving explicit comparisons between different sound conditions were to yield more consistent/robust auditory modulation effects (see Chiou & Rich, 2012; Gallace & Spence, 2006; Orchard-Mills et al., 2013; and Walker et al., 2015, for examples of relative comparisons in audiovisual correspondence studies). Moreover, it may be that participants need to be explicitly made aware of the pertinent correspondences for sonic seasoning to be maximally effective. There is evidence from audiovisual research that crossmodal correspondences helped participants to

⁵ According to Spence and Wang (2015c), there are four types of judgments that people can make about wine: Sensory – evaluations of physical properties of the drink (e.g., flavour, grip, length); Descriptive – metaphorical assessments; Analytic – concerning deductive qualities such as age, complexity, balance, and price assessment; and Hedonic – liking for the wine.

respond more quickly on a visual search task only when participants were explicitly told about the correspondence before the experiment (Klapetek et al., 2012).

Finally, given that individual differences can moderate sonic seasoning effects (Chapter 9), it would perhaps be good practice in future studies to collect relevant information, such as taster status and music expertise, from each participant. Experiment 14 revealed that taster status could play a role in moderating the impact of sonic seasoning, especially for bitter foods. In addition, Experiment 2 demonstrated that music expertise is likely a factor in how people make sound-taste correspondences. In the study, a dissonant high-pitched piano soundtrack was more often matched with sweetness by those without musical training, and more often matched with bitterness by those with musical training. One plausible explanation for this result was that those with musical training tend to focus on melody and harmony, and therefore matched the dissonant harmony with bitterness; in contrast, those without musical training tend to focus on the timbre of the music, and therefore matched the piano instrumentation with sweetness (Wolpert, 1990).

Besides taster status and level of expertise, another potential metric to collect from participants, which was not addressed in this thesis, is their arousal seeking tendency (AST). Lin (2010) asked her participants to evaluate (in terms of arousal and liking) videos of fictitious hotel bars and guest rooms with manipulated colour schemes and music designed to create putatively Gestalt (dynamic bar, tranquil guest room) and non-Gestalt (tranquil bar, dynamic guest room) scenarios. The use of Gestalt situations here is relevant here because, according to some, crossmodal correspondences can be viewed as a type of multisensory Gestalt grouping based on similarity (Spence, 2015a). The results revealed that those who scored higher on the assessment of AST also tended to rate the Gestalt settings to be more arousing and

pleasurable than non-Gestalt settings. This finding is potentially useful for those looking to create novel multisensory experiences in the hospitality industry (see Section 10.3.2), since those people who have more arousal seeking tendencies may find such experiences more enjoyable.

10.3.1.2 Open Questions

Perhaps the biggest open question to date is the origin of sound and taste/flavour correspondences. Section 1.1.5 discussed some possible theories underlying crossmodal correspondences in general – statistical, structural, affective, semantic, and polar. Table 10.1 below summarises crossmodal correspondences between auditory features and taste/flavour/mouthfeel attributes that have been demonstrated to date. One of the most consistent and widely demonstrated correspondences is that seen between pitch and basic taste. As revealed by the results of Experiment 1, this crossmodal correspondence is partially mediated by emotional valence and arousal. Some researchers have also proposed a statistical hypothesis, according to which the correspondence originates in innate stereotypical orofacial gestures that people make in response to ingesting different tastes (Knöferle & Spence, 2012; Spence, 2012b). Evidence consistent with this suggestion comes from observations that babies protrude their tongue out and up in response to pleasant tastes such as sweetness (Rosenstein & Oster, 1988; Steiner et al., 2001). This, in turn, produces a high vowel sound when air is exhaled (Ladefoged & Johnson, 2011). In contrast, the tongue goes out and down in response to unpleasant tastes like bitterness, which then produces a low vowel sound upon exhalation⁶.

⁶ Note that this orofacial gesture theory does not account for the fact that sourness, an aversive taste, corresponds to high pitch.

Yet another theory of taste-pitch correspondence makes use of the purported transitive property of crossmodal correspondences (Section 1.1.4.3). It proposes that the correspondence comes from mutual associations with a third sensory feature, such as colour (see Deroy et al., 2013, for an argument of transitivity for non-obvious correspondences such as those that have been documented between sound and odour). For instance, bitterness is usually matched to darker colours (see Spence et al., 2015b, for a review), and dark colours are matched to low pitch (Marks, 1989; Melara, 1989). By transitivity, bitterness is therefore matched to low pitch. Incidentally, if sound-pitch correspondences are really due to their mutual associations with colour, then it raises the intriguing question of whether those born blind might not experience sound-taste correspondences in the same way as sighted people.

Taste/Flavour	Source (Researcher/Experiment)	Pitch	Tempo	Instrumentation	Harmony	Articulation	Loudness	Distortion/Roughness	Sharpness
Sweet	Experiment 1	high							
	Experiment 2	high	average		consonant	both			
	Experiment 3				consonant				
Sour	Crisinel & Spence, 2010b	high		piano	consonant	legato	soft		
	Mesz et al., 2011	average			consonant	legato			
	Knoeferte et al., 2015	high			consonant	legato	low		low
Salty	Experiment 1	high							
	Experiment 2	high	fast	flute	dissonant	staccato			
	Experiment 3				dissonant				
Bitter	Crisinel & Spence, 2010b	high		brass	dissonant	average	average		
	Mesz et al., 2011	high			dissonant	average			
	Knoeferte et al., 2015	high			dissonant	staccato	high		high
Spicy	Experiment 1	average							
	Experiment 2	average	average		ns	staccato			
	Experiment 3	average			ns				
Creamy	Crisinel & Spence, 2010b	average		brass	average	staccato	average		
	Mesz et al., 2011	average			consonant	staccato			
	Knoeferte et al., 2015	low			consonant	staccato	high		low
Rough	Experiment 1	low							
	Experiment 2	low	slow	cello	dissonant	legato			
	Experiment 3	low		brass	average	legato	average		
Sharpness	Mesz et al., 2011	low			dissonant	legato			
	Knoeferte et al., 2015	low			dissonant	legato	high		low
	Experiment 7	high	fast		ns	ns	high		high
Distortion/Roughness	Experiment 5	high		flute	consonant	legato			
	Experiment 5				consonant	legato			
	Experiment 5			violin (pizzicato)	dissonant	staccato			

Table 10.1. A summary of sound-taste/flavour/mouthfeel correspondences demonstrated to date.

In terms of other sound-taste/flavour correspondences, it is possible that spiciness and its matching auditory properties (high pitch, fast tempo, high distortion), uncovered in Experiment 7, are linked because they are all associated with high levels of emotional arousal (for spiciness, see Watanabe et al., 1987; for auditory properties, see Experiment 2 and Zwaag et al., 2011). On the other hand, the fact that the legato creamy soundtrack enhanced ratings of chocolate creaminess in Experiment 5 can perhaps be most easily interpreted as a connection between legato music and the slow rate at which creamy chocolates melt in the mouth.

Another fundamental question that has yet to be explored concerns the neural mechanism(s) underlying the sonic seasoning effects documented behaviourally. As an example, Woods et al. (2011) conducted a functional MRI (fMRI) study to prove that expectations about the sweetness of a drink can, in fact, influence taste perception. The results demonstrated enhanced activation in the primary taste cortex when participants expected a very sweet drink versus a less sweet one. Similarly, one might want to conduct an fMRI experiment in which the participants listen to different taste-congruent soundtracks while sampling a beverage. Depending on the soundtrack heard, one would then expect to observe increased activation in different taste areas in the primary taste cortex (see Schoenfeld et al., 2004, and Chen et al., 2011, for evidence of gustotopic mapping in the mammalian brain). For instance, listening to a sweet soundtrack might increase activation in the region in the primary taste cortex corresponding to sweetness, compared to when listening to a bitter soundtrack. If this were the case, then there would be a convincing case of sound genuinely affecting the participants' perceptual experience.

Chapter 9 addressed some individual factors such as expertise and taster status, but many more factors remain to be examined. The most important of which is probably

the role of cultural background. All of the participants in the present thesis – even though they were recruited from the U.K., Belgium, Norway, U.S.A., and Argentina – still fit the WEIRD (Western, Educated, Industrialised, Rich, and Democratic) model (Henrich et al., 2010). The quest to determine whether non-WEIRD populations – especially those from non-western cultures – also share these sound-taste correspondences and experience the sonic seasoning effect in the same way is an on-going effort. Such studies would provide valuable information on which sound-taste correspondences are products of culture/environment, and which might originate from mental processes innate to all humans. Colour-flavour correspondences, for instance, exhibit clear cultural differences which hint at their statistical origin (Shankar et al., 2010; Wan et al., 2014, 2015, 2016). For instance, Wan et al. (2016) demonstrated that blue drinks were associated with blueberry flavour by participants from China, India, Korea, the UK, and the USA; while the French, Japanese, and Norwegian participants made significantly different choices (many of them matching blue with a mint flavour instead). Shape-taste correspondences also exhibit cultural differences: Members of the Himba tribe in Northern Namibia, for instance, tend to match less bitter chocolate samples to angular shapes, rather than the rounded shapes commonly chosen by Westerners (Bremner et al., 2013). In terms of sound-taste correspondences, some recent research has demonstrated that Indian participants – who are used to a different musical system than a traditional Western music (Agarwal et al., 2013) – perform at above-chance-level⁷ at decoding basic taste information embedded in complex musical compositions (Knoeferle et al., 2015). In general, cross-cultural variations in sound-taste correspondences remain an area to be investigated in future studies. The caveat here, of course, being that re-testing each

⁷ Albeit at a slightly lower level than U.S. participants.

experiment for cross-cultural effects and comparing across a sufficient number of cultural backgrounds would rapidly multiply the number of experiments that need to be conducted.

On a separate point, most of the experiments reported in this thesis involved a contrast comparison with a pair of soundtracks (e.g., sweet and bitter), where a difference in taste ratings was observed as a function of soundtracks the participants listened to. What remains unknown, however, is whether say the sweet soundtrack enhances sweetness, or whether a bitter soundtrack decreases sweetness (or both). Only Experiment 7B-D compared the spicy soundtrack with a silent control condition, which demonstrated that spicy-congruent soundtrack can indeed enhance spiciness. In the future, more experiments are needed with a neutral⁸ (silent) condition included, to study the relative contributions of congruent versus incongruent soundtracks on the taste experience.

Another open question involves the bidirectional property of crossmodal correspondences, as discussed in Section 1.1.4.1. Mesz and his colleagues (2011) demonstrated that some correspondences between sound parameters (pitch, articulation, loudness, duration, and harmonic dissonance) and basic tastes seem to be bidirectional. In their first study, musicians were given the taste words sweet, sour, bitter, and salty and were asked to improvise short musical pieces based on those cue words. In their second study, the task was reversed; the participants who listened to these improvised musical pieces were able to perform at above-chance-level at decoding the taste words which inspired the improvisations. What remains to be determined, in terms of bi-directionality, is whether sound-taste correspondences can

⁸ That said, it is a non-trivial task to determine the most appropriate neutral condition (Jonides & Mack, 1984), especially for studies which aim to be ecologically valid.

influence auditory as well as gustatory perception. For instance, if high pitched, dissonant music can make juice mixtures taste sourer (see Experiments 3 and 11), might tasting something sour make a simultaneously-presented auditory stimulus sound more high pitched and dissonant (or at least focus one's attention on the higher pitched/dissonant element)? A future study could involve an ambiguous auditory signal with high and low tones (see van Ee et al., 2009, Experiment 6), where the participants were instructed to hold one of two keys depending on whether the high or low tone series is dominant at any point in time. One could then compare the differences in durations of dominance for the two tones, depending on whether the participants were simultaneously experiencing a sour or bitter taste.

In summary, there are many open questions to be answered in the area of sound-taste/flavour correspondences. In addition to examining the area from a theoretical viewpoint, some practical applications of sonic seasoning are reviewed in light of the results from the present thesis. Assessing the practicalities of such applications will help to elucidate a number of avenues for future research.

10.3.2 Case study 1: Taste-congruent music in restaurants

So-called “multisensory” or “experiential” dining has become increasingly popular in recent years (Mitchell, 2012; Spence, 2017; Spence et al., 2015c; Spence & Piqueras-Fiszman, 2014; Wang, 2013), but it remains unclear just how effective taste-congruent music might be in a restaurant setting if one wants to design a sonic seasoning experience. Previously, in my Master's thesis at the MIT Media Lab (see Wang, 2013), I discussed some of the challenges and pitfalls around designing music-food experiences, such as allowing for social interaction and accounting for individual differences. Extending the work I started there, factors involved in designing a sonic-

seasoning enhanced dining experience are explored below, including both applications from results in this thesis and open questions for future experiments.

In terms of the food/drink being served, experiments have typically used food/drink samples with complex and nuanced flavours, such as wine, beer, or bittersweet chocolate. As presented in Chapter 7, taste-congruent soundtracks can shift people's attention towards different elements in the food/drink. Since it has been argued that the effect of attention on awareness tends to be more pronounced with more complex input (Lavie, 1995, 2005), it follows that more complex and nuanced food/drink might result in more pronounced sonic seasoning effects. Wine is an especially good target for sonic seasoning since it can express a myriad of aromas and flavours. In a restaurant setting, for example, music can be selected/designed to round out green tannins in a young wine (similar to the creamy music from Experiment 4), or to invigorate a tired or old wine by bringing out acidity and fruitiness (similar to how Rachmaninoff's Vocalise enhanced fruitiness in Experiment 3). Music can even be customised to reflect personal preferences, so that diners can enjoy the same bottle of wine even if, say, one person prefers crisp and elegant wines while the other prefers powerful and fruity wines. A caveat that needs to be addressed here though is that background music may also have a negative impact on the wine if one is not careful. Fine wines are often characterised by a delicate balance between acid, alcohol, oak, and tannins (Fielden, 2009). Therefore, there could be a real danger of playing music that might adversely affect the balance of the wine and spoil the wine tasting experience.

With regard to the music itself, one important question must be addressed. Namely, will it be enough for the music to be merely playing in the background, or do the diners need to actively pay attention to it? Experiment 10 demonstrated that music can

indeed focus participants' attention on different taste/flavour attributes, but did not address the issue of having participants pay attention to the music in the first place. In all the experiments reported in this thesis and in those reported in Section 1.2.2.1, the taste-congruent soundtracks were presented in isolation. Furthermore, the participants were instructed either explicitly or implicitly (i.e., by being asked to wear headphones) to pay attention to the sounds they hear. In a restaurant setting, on the other hand, the background music would have to compete with the sound of service and of any others conversations happening around the diner (which can be very loud; Spence, 2014b). It is unclear therefore how loud the music would need to be in order for the diner to experience the sonic seasoning effect; in an experimental setting, the music is usually presented at a comfortable listening level of around 70 dB(A), but variations in listening volume have not been tested empirically (although Zellner et al., 2017, who conducted a study in a student cafeteria, pointed out that the null result they observed might have been due to the fact that the noise of background conversations in their study drowned out the sound of any experimental music conditions).

Another issue here regarding using sonic seasoning in restaurants is that even the level of background noise may affect participant's taste ratings (Stafford et al., 2012; Woods et al., 2011; Yan & Dando, 2015; see Spence, 2014b, for a review). The Fat Duck, an adventurous Michelin-starred restaurant, mitigated this issue by presenting their Sound of the Sea dish (see Section 1.2.2.2) over ear buds, which implicitly focuses the diners' attention on the sound and minimises background noise and distractions from fellow diners. However, eating an entire meal while wearing ear buds or headphones would probably not appeal to most people. A more sociable solution might be to introduce music at the moment when the dish arrives at the table,

when diners naturally break away from conversations to focus on the food. As demonstrated in Chapter 6, introducing the music at this point could also enhance the diners' expectations about the dish they are about to eat, and subsequently influence the taste/flavour experience⁹.

In terms of the target audience for such experiential dining events, Experiment 13 revealed that wine experts also experience sonic seasoning effects, so any wine-related experiences need not be limited to amateur wine drinkers. However, according to the results of Experiment 14, those who detect more intense bitterness on PTC taste strips might have a different taste experience than others, especially when it comes to bitter dishes (Experiment 14). To account for individual differences, one can, for instance, imagine a restaurant where a PTC taste strip and an AST assessment test is sent out at the time the reservation is made, so that the diners can receive individualised experiences (see Spence, 2017).

Finally, what kind of music should restaurants play? The comparison of 24 taste soundtracks in Experiment 2 revealed that participants gave most consistent choices when asked to match sweet soundtracks to basic tastes (56.9% averaged over all sweet soundtracks, where the chance level is 25%) and worst for bitter soundtracks (31.4% averaged over all bitter soundtracks). In addition, there tends to be some confusion between sour and bitter soundtracks (e.g., Ishii & O'Mahony, 1987; O'Mahony et al., 1979; Spence et al., 2015a; Wise & Bresin, 2011). This suggests that sonic seasoning experiences are probably most effective if designed using sweet soundtracks, since participants will likely easily associate the soundtracks with sweetness. In further support of using sweet soundtracks, there is evidence that when

⁹ With the caveat that the induced expectations should be reasonably close to the actual taste of the dish!

participants are exposed to both a sweet soundtrack and a sour or bitter soundtrack, the degree of modulation in sweetness ratings is greater than that of bitter or sour ratings (Reinoso Carvalho et al., 2016). Beyond basic tastes, Experiments 5 and 10 demonstrated that music could also be used to modify oral-somatosensory perception like creaminess and spiciness (see Table 10.1 for a summary of auditory features crossmodally corresponding to various taste/flavour attributes). In terms of music presentation, as mentioned in Section 10.3.1.1, most experiments have been performed using explicit comparisons between distinctly different pieces of music. Similarly, it might be more effective for designers to follow suit and present multiple pieces of music with the same dish in order to maximise the sonic seasoning effect.

As a final note, Reinoso Carvalho and his colleagues (2015a) demonstrated that people's liking for a novel flavour of chocolate is enhanced when they are told they are listening to a soundtrack specifically designed to match the chocolate. This certainly bodes well for the future of restaurants (or bars) offering sonic seasoning!

10.3.3 Case study 2: Promoting healthier eating

In the age of widespread obesity (e.g., Finucane et al., 2011; Lee et al., 2017), it is becoming increasingly important to invent new ways to promote healthier eating behaviours. Using taste-congruent soundtracks to enhance taster satisfaction while at the same time reducing sugar, salt, or fat could be a viable solution in reducing obesity in some small way, but there are a few issues to consider in order to implement soundtrack-based health interventions.

First, there is the question of whether sensory cues from different modalities can be combined in order to deliver maximum impact in terms of modifying flavour perception. There is already some evidence that combining music and coloured lights

is more effective at enhancing fruitiness in a red wine than with colour alone (Spence et al., 2014a). Experiment 11 demonstrated that both a happy face image and a consonant soundtrack enhanced the participants' sweetness ratings of a juice sample, compared to a sad face or a dissonant soundtrack. In terms of enhancing sweetness perception, it would therefore be worth exploring the combination of congruent cues from different modalities (such as happy images, appropriately coloured servingware and lighting, and sweet soundtracks) in a future study (see Woods et al., 2013, for an example of the additive effect of emotional information on a sound-image matching task).

Second, it is crucial to remember that using taste-congruent soundtracks to promote healthful outcomes involves more than a one-time effort. To date, no study has examined whether repeated exposure to the same soundtrack (e.g., daily exposure during meal times) might reduce the effectiveness of sonic seasoning. Even at a single session, exposure to taste soundtracks has been limited to a short period of time (Experiment 7B took five minutes per trial, and every other study took under 60 seconds per trial). It remains to be seen whether such soundtracks might be effective over an entire meal, or whether people simply become habituated to the soundtracks. If habituation were a concern, then perhaps alternating between different taste soundtracks during the course of the meal would emphasise contrast effects (as observed in most experiments in this thesis, where participants heard two different taste soundtracks). Clearly, future studies would be necessary in order to determine the feasibility of taste-congruent soundtracks as a long-term solution as far as promoting healthy eating is concerned.

The results of the experiments presented in this thesis demonstrate that taste-congruent sounds can modify taste/flavour ratings by an average of 14% (SD=3%)¹⁰. This has led to speculation that with the simultaneous presentation of a sweet soundtrack, participants/consumers would be equally satisfied with a food/drink with 10% less sugar. To date, no study has been conducted to compare hedonic reports from participants who taste a food item alongside a reduced-sugar (but sonically enhanced) substitute. Such a study would be critical for the advent of technologically enhanced servingware, whose makers are ready to make use of the claim that sonic seasoning can reduce the intake of sugar (and possibly other unhealthy substances). In fact, a prototype of “Sonic Sweetener” beverage cup, by Xin Café, has already been created (<http://xincafe.cn/sonicsweetener/en/>); the plastic cup plays putatively sweet soundtracks and glows pink whenever someone drinks from it, and is designed to enhance the sweetness of the beverage inside without adding sugar.

Another way in which music might encourage healthy eating takes advantage of the possibility that (healthy) foods with unpleasant tastes may be subject to the distraction effects of soundscapes/music (see Silvestrini et al., 2011, for evidence that auditory distraction can reduce pain). This is a potential strategy for persuading people to eat healthier foods; for instance, playing a pleasant soundtrack might make bitter vegetables such as kale more palatable as one’s attention is diverted away from any unpleasant bitter tastes.

Finally, it remains to be seen whether, beyond merely modifying taste/flavour ratings, taste-congruent soundtracks could also be used strategically to change people’s eating behaviour. For example, it would be useful to assess if a sweet soundtrack might induce people to add less sugar to their coffee or consume a smaller portion of

¹⁰ More specifically, for sweet tastes, the average amount of taste modulation is 15% (SD=3%).

dessert. After all, music has been demonstrated to change people's purchase behaviour (e.g., Areni & Kim, 1993; North et al., 1997, 2003) and speed of eating/drinking (e.g., Caldwell & Hilbert, 1999; MacElrea & Standing, 1992; Milliman, 1986; Roballey et al., 1985; see North & Hargreaves, 2008, and Spence & Piqueras-Fiszman, 2014, for reviews).

10.4 General conclusions

The fact that eating experiences involve extra senses beyond just taste and smell has become increasingly clear to researchers in recent years (e.g., Auvray & Spence, 2008; Spence & Piqueras-Fiszman, 2014). People's perception of food and beverages can be influenced by means of aromas, shapes, colours, and even sounds. Over-and-above any food-intrinsic sounds (e.g. mastication sounds), various studies (reviewed in Section 1.2) have begun to demonstrate the relationship between auditory attributes and taste/flavour. Results from this thesis have demonstrated that people make non-random matches between auditory stimuli (both in terms of auditory features and musical pieces) and a range of taste/flavour attributes including basic tastes, creaminess, spiciness, even temperature (see Appendix A). Moreover, the studies illustrated that soundtracks crossmodally congruent with specific taste/flavours can modify the evaluation of a variety of foods and beverages.

More importantly, this thesis highlights the multiple pathways by which taste/flavour-congruent soundtracks might influence taste/flavour perception. Mirroring the role of food-related auditory cues (such as the sizzle of the steak or the crunch of an apple), such soundtracks could shape our sensory expectations before tasting, as well as shift our focus to a specific taste/flavour during eating. In addition, these soundtracks

might induce certain emotions in the listeners, which, in turn, can influence their affective response to food/beverages. Overall, the results presented in this thesis demonstrate the role of crossmodal correspondences – those often surprising associations between basic features of different sensory modalities that do not provide redundant identifying information about the same object/concept (see Section 1.1) – in facilitating the integration of information from different senses, especially when it comes to food/beverage perception, that most important influence in determining the organization of the brain (see Young, 1968).

REFERENCES

- Agarwal, P., Karnick, H., & Raj, B. (2013). *A comparative study of Indian and western music forms*. Paper presented at the International Society for Music Information Retrieval Conference (ISMIR), Curitiba, Brazil.
- Alaoui-Ismaili, O., Robin, O., Rada, H., Dittmar, A., & Vernet-Maury, E. (1997). Basic emotions evoked by odorants: Comparison between autonomic responses and self-evaluation. *Physiology & Behavior*, **62**, 713-720.
- Alsius, A., Navarra, J., Campbell, R., & Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Current Biology*, **15**, 839-843.
- Alsius, A., Navarra, J., & Soto-Faraco, S. (2007). Attention to touch weakens audiovisual speech integration. *Experimental Brain Research*, **183**, 399-404.
- Amerine, M. A., Pangborn, R. M., & Roessler, E. B. (1965). *Principles of sensory evaluation of food*, London, UK: Academic Press.
- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, **22**, 261-295.
- Anderson, R. E. (1973). Consumer dissatisfaction: The effect of disconfirmed expectancy on perceived product performance. *Journal of Marketing Research*, **10**, 38-44.
- Arabie, P., & Moskowitz, H. R. (1971). The effects of viscosity upon perceived sweetness. *Perception & Psychophysics*, **9**, 410-412.
- Arvisenet, G., Guichard, E., & Ballester, J. (2016). Taste-aroma interaction in model wines: Effect of training and expertise. *Food Quality and Preference*, **52**, 211-221.

- Ashkenazi, A., & Marks, L. E. (2004). Effect of endogenous attention on detection of weak gustatory and olfactory flavors. *Perception & Psychophysics*, **66**, 596-608.
- Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and Cognition*, **17**, 1016-1031.
- Bachorik, J. P., Bangert, M., Loui, P., Lark, K., Berger, J., Rowe, R., & Schlaug, G. (2009). Emotion in motion: Investigating the time-course of emotional judgments of musical stimuli. *Music Perception*, **26**, 355-364.
- Barbiere, J. M., Vidal, A., & Zellner, D. A. (2007). The color of music: Correspondence through emotion. *Empirical Study of the Arts*, **25**, 193-208.
- Bartoshuk, L. M. (1991). Sweetness. History, preference, and genetic variability. *Food Technology*, **45**, 108, 110, 112-113.
- Bartoshuk, L. M., Duffy, V. B., & Miller, I. J. (1994). PTC/PROP tasting: Anatomy, psychophysics, and sex effects. *Physiology & Behavior*, **56**, 1165-1171.
- Bartoshuk, L. M., Fast, K., & Snyder, D. J. (2005). Differences in our sensory worlds: Invalid comparisons with labeled scales. *Current Directions in Psychological Science*, **14**, 122-125.
- Bartoshuk, L. M., Rifkin, B., Marks, L. E., & Bars, P. (1986). Taste and aging. *Journal of Gerontology*, **41**, 51-57.
- Basu, T. (2016). These music clips are supposed to change the way your coffee tastes. *New York Magazine*, **January 27th**. Retrieved from <http://nymag.com/scienceofus/2016/01/drink-your-coffee-and-listen-to-these-two-clips.html> on 10/05/2016.

- Battacharya, J., & Lindsen, J. P. (2016). Music for a brighter world: Brightness judgment bias by musical emotion. *PLoS ONE*, **11**:e0148959.
- Beeli, G., Esslen, M., & Jancke, L. (2005). Synaesthesia: When coloured sounds taste sweet. *Nature*, **434**, 38.
- Belkin, K., Martin, R., Kemp, S. E., Gilbert, A. N. (1997). Auditory pitch as a perceptual analogue to odor quality. *Psychological Science*, **8**, 340-342.
- Ben-Artzi, E., & Marks, L. E. (1995). Visual-auditory interaction in speeded classification: Role of stimulus difference. *Perception and Psychophysics*, **57**, 1151-1162.
- Bernstein, I. H., & Edelstein, B. A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology*, **87**, 241-247.
- Bharitkar, S., & Kyriakakis, C. (2006). *Immersive audio signal processing*. New York, NY: Springer-Verlag.
- Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience*, **29**, 13165-13171.
- Biggs, L., Juravle, G., & Spence, C. (2016). Haptic exploration of plateware alters the perceived texture and taste of food. *Food Quality and Preference*, **50**, 129-134.
- 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**, 97-107.
- Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, **2**, 382-387.

- Bolognini, N., Frassinetti, F., Serino, A., & Ladavas, E. (2005). “Acoustical vision” of below threshold stimuli: Interaction among spatially converging audiovisual inputs. *Experimental Brain Research*, **160**, 273-282.
- Botvinick, M., & Cohen, J. (1998). Rubber hands ‘feel’ touch that eyes see. *Nature*, **391**, 756.
- Brand, G., & Brisson, R. (2012). Lateralisation in wine olfactory threshold detection: Comparison between experts and novices. *Laterality*, **17**, 583-596.
- Bredie, W. L. P., Tan, H. S. G., & Wendin, K. (2014). A comparative study on facially expressed emotions in response to basic tastes. *Chemosensory Perception*, **7**, 1-9.
- Bremner, A., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K., & Spence, C. (2013). Bouba and Kiki in Namibia? A remote culture make similar shape-sound matches, but different shape-taste matches to Westerners. *Cognition*, **126**, 165-172.
- Bresin, R. (2005). What is the color of that music performance? *Proceedings of the International Computer Music Conference – ICMC 2005*, 367-370.
- Bronner, K., Frieler, K., Bruhn, H., Hirt, R., & Piper, D. (2012). What is the sound of citrus? Research on the correspondences between the perception of sound and flavour. In E. Cambouropoulos, C. Tsougras, P. Mavromatis, & K. Pasteriadis (Eds.), *Proceedings of the ICMPC – ESCOM 2012 Joint Conference*. Thessaloniki, Greece.
- Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., & Liotti, M. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *NeuroImage*, **58**, 250-258.
- Cadena, R. S., & Bolini, H. M. A. (2011). Time–intensity analysis and acceptance test for traditional and light vanilla ice cream. *Food Research International*, **44**, 677-683.

- Cain, W. S., & Murphy, C. L. (1980). Interaction between chemoreceptive modalities of odour and irritation. *Nature*, **284**, 255-257.
- Cardello, A. V. (2007). Measuring consumer expectations to improve food product development. In H. J. H. MacFie (Ed.), *Consumer-led food product development* (pp. 223-261). Cambridge, UK: Woodhead Publishing.
- Cardello, A. V., & Sawyer, F. (1992). Effects of disconfirmed consumer expectations of food acceptability. *Journal of Sensory Studies*, **7**, 253-277.
- Carlsmith, J. M., & Aronson, E. (1963). Some hedonic consequences of the confirmation and disconfirmation of expectancies. *Journal of Abnormal and Social Psychology*, **66**, 151-156.
- Castriota-Scanderbeg, A., Hagberg, G. E., Cerasa, A., Committeri, G., Galati, G., Patria, F., et al. (2005). The appreciation of wine by sommeliers: A functional magnetic resonance study of sensory integration. *NeuroImage*, **25**, 570-578.
- Catanzaro, D., Chesbro, E. C., & Velkey, A. J. (2013). Relationship between food preferences and PROP taster status of college students. *Appetite*, **68**, 124-131.
- Caterina, M. J., Schumacher, M. A., Tominaga, M., Rosen, T. A., Levine, J. D., & Julius, D. (1997). The capsaicin receptor: A heat-activated ion channel in the pain pathway. *Nature*, **389**, 816-824.
- Cavanaugh, L. A., MacInnis, D. J., & Weiss, A. M. (2016). Perceptual dimensions differentiate emotions. *Cognition and Emotion*, **30**, 1430-1445.

- Chalé-Rush, A., Burgess, J. R., & Mattes, R. D. (2007). Multiple routes of chemosensitivity to free fatty acids in humans. *American Journal of Physiology: Gastrointestinal and Liver Physiology*, **292**, G1206-G1212.
- Chang, S., & Cho, Y. S. (2015). Polarity correspondence effect between loudness and lateralized response set. *Frontiers in Psychology*, **6**:683.
- Chen, K., Zhou, B., Chen, S., He, S., & Zhou, W. (2013). Olfaction spontaneously highlights visual saliency map. *Proceedings of the Royal Society*, **280**:20131729.
- Chen, X., Gabitto, M., Peng, Y., Ryba, N. J. P., & Zucker, C. S. (2011). A gustotopic map of taste qualities in the mammalian brain. *Science*, **333**, 1262-1266.
- Chen, Y.-C., & Spence, C. (2010). When hearing the bark helps to identify the dog: Semantically-congruent sounds modulate the identification of masked pictures. *Cognition*, **114**, 389-404.
- Chen, Y.-C., & Spence, C. (2011a). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, **37**, 1554-1568.
- Chen, Y.-C., & Spence, C. (2011b). The crossmodal facilitation of visual object representations by sound: evidence from the backward masking paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, **37**, 1784-1802.
- Cheskin, L. (1972). *Marketing success: How to achieve it*. Boston, MA: Cahnern Books.
- Chiandetti, C., & Vallortigara, G. (2011). Chicks like consonant music. *Psychological Science*, **22**, 1270-1273.

- Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception*, **41**, 339-353.
- Chollet, S., & Valentin, D. (2000). Le degré d'expertise a-t-il une influence sur la perception olfactive? Quelques éléments de réponse dans le domaine du vin [Does the level of expertise influence olfactory perception? Some answers from the domain of wine]. *Annals of Psychology*, **100**, 11-36.
- Christensen, C. M. (1980). Effects of solution viscosity on perceived saltiness and sweetness. *Perception & Psychophysics*, **28**, 347-353.
- Clapperton, J. F., & Piggott, J. R. (1979). Flavor characterization by trained and untrained assessors. *Journal of the Institute of Brewing*, **85**, 275-277.
- Clark, C. C., & Lawless, H. T. (1994). Limiting response alternatives in time-intensity scaling: An examination of the halo dumping effect. *Chemical Senses*, **19**, 583-594.
- Clydesdale, F. M., Gover, R., Philipsen, D. H., & Fugardi, C. (1992). The effect of color on thirst quenching, sweetness, acceptability and flavor intensity in fruit punch flavored beverages. *Journal of Food Quality*, **15**, 19-38.
- Collier, G. L. (1996). Affective synesthesia: Extracting emotion space from simple perceptual stimuli. *Motivation and Emotion*, **20**, 1-32.
- Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, **82**, 407-428.
- Connolly, K. (2014). Multisensory perception as an associative learning process. *Frontiers in Psychology*, **5**:1095.

- Cook, D. J., Hollowood, T. A., Linforth, R. S. T., & Taylor, A. J. (2003). Oral shear stress predicts flavour perception in viscous solutions. *Chemical Senses*, **28**, 11-23.
- Cramb, A. (2008). Why wine tastes better with music. *The Telegraph*, **May 13th**.
<http://www.telegraph.co.uk/news/uknews/1952381/Why-wine-tastes-better-with-music.html>.
- Crawshaw, A. (2012). *How musical emotion may provide clues for understanding the observed impact of music on gustatory and olfactory perception in the context of wine-tasting*. Unpublished manuscript.
- Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, **2**, 263-275.
- Crisinel, A.-S., Cosser, S., King, S., Jones, R., Petrie, J., & Spence, C. (2012). A bittersweet symphony: Systematically modulating the taste of food by changing the sonic properties of the soundtrack playing in the background. *Food Quality and Preference*, **24**, 201-204.
- Crisinel, A.-S., Jacquier, C., Deroy, O., & Spence, C. (2013). Composing with cross-modal correspondences: Music and smells in concert. *Chemosensory Perception*, **6**, 45-52.
- Crisinel, A.-S., & Spence, C. (2009). Implicit association between basic tastes and pitch. *Neuroscience Letters*, **464**, 39-42.
- Crisinel, A.-S., & Spence, C. (2010a). A sweet sound? Food names reveal implicit associations between taste and pitch. *Perception*, **39**, 417-425.

- Crisinel, A.-S., & Spence, C. (2010b). As bitter as a trombone: Synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Attention, Perception, & Psychophysics*, **72**, 1994-2002.
- Crisinel, A.-S., & Spence, C. (2011a). A fruity note: Crossmodal associations between odours and musical notes. *Chemical Senses*, **37**, 151-158.
- Crisinel, A.-S., & Spence, C. (2011b). Crossmodal associations between flavoured milk solutions and musical notes. *Acta Psychologica*, **138**, 155-161.
- Crisinel, A.-S., & Spence, C. (2012). The impact of pleasantness ratings on crossmodal associations between food samples and musical notes. *Food Quality and Preference*, **24**, 136-140.
- Croijmans, I., & Majid, A. (2016). Not all flavor expertise is equal: The language of wine and coffee experts. *PLoS ONE*, **11**:e0155845.
- Crollen, V., Grade, S., Pesenti, M., & Dormal, V. (2013). A common metric magnitude system for the perception and production of numerosity, length, and duration. *Frontiers in Psychology*, **4**:449.
- Dalton, P., Santangelo, V., & Spence, C. (2009). The role of working memory in auditory selective attention. *The Quarterly Journal of Experimental Psychology*, **62**, 2126-2132.
- de Araujo, I. E. T., & Rolls, E. T. (2004). The representation in the human brain of food texture and oral fat. *Journal of Neuroscience*, **24**, 3086-3093.
- de Araujo, I. E. T., Rolls, E. T., Kringelbach, M. L., McGlone, F., & Phillips, N. (2003). Taste-olfactory convergence, and the representation of the pleasantness of flavour, in the human brain. *European Journal of Neuroscience*, **18**, 2059-2068.

- de Gelder, B., and Vroomen, J. (2000). The perception of emotions by ear and by eye. *Cognition and Emotion*, **14**, 289-311.
- de Lange, C. (2012). Feast for the senses: Cook up a master dish. *New Scientist*.
<http://www.newscientist.com/article/mg21628962.200-feast-for-the-senses-cook-up-a-master-dish.html>. Accessed 21 December 2016.
- Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, **10**, 204-211.
- Delaney, H. D., & Maxwell, S. E. (1981). On using analysis of covariance in repeated measures designs. *Multivariate Behavioral Research*, **16**, 105-123.
- 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**, 103-128.
- Dellacherie, D., Roy, M., Hugueville, L., Peretz, I., & Samson, S. (2011). The effect of musical experience on emotional self-reports and psychophysiological responses to dissonance. *Psychophysiology*, **48**, 337-349.
- Delon-Martin, C., Plailly, J., Fonlupt, P., Veyrac, A., & Royet, J.-P. (2013). Perfumers' expertise induces structural reorganization in olfactory brain regions. *NeuroImage*, **68**, 55-62.
- Delwiche, J. F. (2004). The impact of perceptual interactions on perceived flavor. *Food Quality and Preference*, **15**, 137-146.

- Demattè, M. L., Pojer, N., Endrizzi, I., Corollaro, M. L., Betta, E. Aprea, E., ... Gasperi, F. (2014). Effects of the sound of the bite on apple perceived crispness and hardness. *Food Quality and Preference*, **38**, 58-64.
- Deroy, O., Crisinel, A.-S., & Spence, C. (2013). Crossmodal correspondences between odors and contingent features: Odors, musical notes, and geometrical shapes. *Psychonomic Bulletin & Review*, **20**, 878-896.
- Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic Bulletin & Review*, **20**, 643-664.
- Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, **92**, 53-78.
- Droit-Volet, S., Ramos, D., Bueno, J. L., & Bigand, E. (2013). Music, emotion, and time perception: The influence of subjective emotional valence and arousal? *Frontiers in Psychology*, **4**:417.
- Durgin, F. H., Evans, L., Dunphy, N., Klostermann, S., & Simmons, K. (2007). Rubber hands feel the touch of light. *Psychological Science*, **18**, 152-157.
- Eitan, Z., & Rothschild, I. (2010). How music touches: Musical parameters and listeners' audiotactile metaphorical mappings. *Psychology of Music*, **39**, 449-467.
- Ekman, P. (2012). An argument for basic emotions. *Cognition and Emotion*, **6**, 169-200.
- Endo, H., Ino, S., & Fujisaki, W. (2016). The effect of a crunchy pseudo-chewing sound on perceived texture of softened foods. *Physiology & Behavior*, **167**, 324-331.

- Epel, E., Lapidus, R., McEwen, B., & Brownell, K. (2001). Stress may add bite to appetite in women: A laboratory study of stress-induced cortisol and eating behaviour. *Psychoneuroendocrinology*, **26**, 37-49.
- Erlebacher, A., & Sekuler, R. (1971). Response frequency equalization: A bias model for psychophysics. *Perception & Psychophysics*, **9**, 315-320.
- Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, **7(5)**:7.
- Ethofer, T., Anders, S., Erb, M., Droll, C., Royen, L., Saur, R., et al. (2006). Impact of voice on emotional judgment of faces: An event-related fMRI study. *Human Brain Mapping*, **27**, 707-714.
- Evans, K. K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, **10(1)**:6
- Fairhurst, M. T., Pritchard, D., Ospina, D., & Deroy, O. (2015). Bouba-Kiki in the plate: Combining crossmodal correspondences to change flavour experience. *Flavour*, **4**:22.
- Fallows, D. "Dolce (i)". *Grove Music Online*. Oxford Music Online. Oxford University Press. Web. 02 July 2015. <http://www.oxfordmusiconline.com/subscriber/article/grove/music/07936>.
- Fernández-Prieto, I., Navarra, J. & Pons, F. (2015). How big is this sound? Crossmodal association between pitch and size in infants. *Infant Behavior and Development*, **38**, 77-81.
- Fielden, C. (2009). *Exploring the world of wines and spirits*. London, UK: Wine & Spirit Education Trust.

- Finucane, M. M., Stevens, G. A., Cowan, M. J., Danaei, G., Lin, J. K., Paciorek, C. J., et al. (2011). National, regional, and global trends in body-mass index since 1980: Systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants. *Lancet*, **377**, 557-567.
- Fishman, Y. I., Volkov, I. O., Noh, M. D., Garell, P. C., Bakken, H., Arezzo, J. C., Howard, M.A., & Steinschneider, M. (2001). Consonance and dissonance of musical chords: Neural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology*, **86**, 2761-2788.
- Fleming, A. (2014). How sound affects the taste of our food. *The Guardian*, **March 11th**. Retrieved from <http://www.theguardian.com/lifeandstyle/wordofmouth/2014/mar/11/sound-affects-taste-food-sweet-bitter> on 10/05/2016.
- Foster, P. S., Roosa, K. M., Drago, V., Branch, K., Finney, G., & Heilman, K. M. (2013). Recall of word lists is enhanced with increased spreading activation. *Aging, Neuropsychology, and Cognition*, **20**, 553-566.
- Foster, P. S., Wakefield, C., Pryjmak, S., Roosa, K. M., Branch, K. K., Drago, V., ... Ruff, R. (2016). Spreading activation in nonverbal memory networks. *Brain Informatics*, **28**, doi:10.1007/s40708-016-0058-y.
- Frassinetti, F., Pavani, F., & Ladavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, **147**, 332-343.
- Gal, D. (2012). A mouth-watering prospect: Salivation to material reward. *Journal of Consumer Research*, **38**, 1022-1029.

- Gallace, A., Boschini, E., & Spence, C. (2011). On the taste of “Bouba” and “Kiki”: An exploration of word-food associations in neurologically normal participants. *Cognitive Neuroscience*, **2**, 34-46.
- Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, **68**, 1191-1203.
- Gardner, H. (1974). Metaphors and modalities: How children project polar adjectives onto diverse domains. *Child Development*, **45**, 84-91.
- Garneau, N. L., Nuessle, T. M., Sloan, M. M., Santorico, S. A., Coughlin, B. C., & Hayes, J. E. (2014). Crowdsourcing taste research: Genetic and phenotypic predictors of bitter taste perception as a model. *Frontiers in Integrative Neuroscience*, **8**:33.
- Gawel, R. (1997). The use of language by trained and untrained experienced wine tasters. *Journal of Sensory Studies*, **12**, 267-284.
- Gerdes, A. B. M., Wieser, M. J., & Alpers, G. W. (2014). Emotional pictures and sounds: A review of multimodal interactions of emotion cues in multiple domains. *Frontiers in Psychology*, **5**:1351.
- Geronazzo, M., Avanzini, F., & Grassi, M. (2015). Absence of modulatory action on haptic height perception with musical pitch. *Frontiers in Psychology*, **6**:1369.
- Gescheider, G. A. (1997). *Psychophysics: The fundamentals*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.

- Giovanni, M. E., & Pangborn, R. M. (1983). Measurement of taste intensity and degree of liking of beverages by graphic scales and magnitude estimation. *Journal of Food Science*, **48**, 1175-1182.
- Goodstein, E. S., Bohlscheid, J. C., Evans, M., & Ross, C. F. (2014). Perception of flavor finish in model white wine: A time-intensity study. *Food Quality and Preference*, **36**, 50-60.
- Gottfried, J. A. (2010). Central mechanisms of odour object perception. *Nature Reviews Neuroscience*, **11**, 628-641.
- Grassi, M., Pastore, M., & Lemaitre, G. (2013). Looking at the world with your ears: how do we get the size of an object from its sound? *Acta Psychologica*, **143**, 96-104.
- Gray, W. B. (2007). Music to drink wine by: Vintner insists music can change wine's flavors. *San Francisco Chronicle*, **November 11**. <http://www.sfgate.com/wine/article/Music-to-drink-wine-by-Vintner-insists-music-can-3235602.php> on 13/08/2015.
- Green, B. G., & Frankmann, S. P. (1988). The effect of cooling on the perception of carbohydrate and intensive sweeteners. *Physiology & Behavior*, **43**, 515-519.
- Green, B. G., & Lawless, H. T. (1991). The psychophysics of somatosensory chemoreception in the nose and mouth. In T. V. Getchell, L. M. Bartoshuk, R. L. Doty, & J. B. Snow (Eds.), *Smell and taste in health and disease* (pp. 235-253). New York, NY: Raven Press.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, **74**, 1461-1480.

- Guest, S., Grabenhorst, F., Essick, G., Chen, Y., Young, M., McGlone, F., ... Rolls, E. T. (2007). Human cortical representation of oral temperature. *Physiology & Behavior*, **92**, 975-984.
- Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive coding of visual spatial frequency and auditory amplitude-modulation rate. *Current Biology*, **22**, 383-388.
- Gvion, L., & Trostler, N. (2008). From spaghetti and meatballs through Hawaiian pizza to sushi: The changing nature of ethnicity in American restaurants. *Journal of Popular Culture*, **41**, 950-974.
- Hagemuller, F., Rossler, W., Wittwer, A., & Haker, H. (2014). Juicy lemons for measuring basic empathic resonance. *Psychiatry Research*, **219**, 391-396.
- Hall, M. J., Bartoshuk, L. M., Cain, W. S., & Stevens, J. C. (1975). PTC taste blindness and the taste of caffeine. *Nature*, **253**, 442-443.
- Harrar, V., Piqueras-Fiszman, B., & Spence, C. (2011). There's more to taste in a coloured bowl. *Perception*, **40**, 880-882.
- Harrar, V., Smith, B., Deroy, O., & Spence, C. (2013). Grape expectations: How the proportion of white grape in Champagne affects the ratings of experts and social drinkers in a blind tasting. *Flavour*, **2**:25.
- Hayes, A. E. (2013). *Introduction to mediation, moderation, and conditional analysis*. New York, NY: Guildford Publications.

- Hein, G., Doehrmann, O., Müller, N. G., Kaiser, J., Muckli, L., & Naumer, M. J. (2007). Object familiarity and semantic congruency modulate responses in cortical audiovisual integration areas. *Journal of Neuroscience*, **27**, 7881-7887.
- Helmholtz, H. L. F. (1954). *On the sensations of tone as a physiological basis for the theory of music*. New York, NY: Dover, (Original work published 1863).
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, **33**, 61-135.
- Hetherington, M. M. (1996). Sensory-specific satiety and its importance in meal termination. *Neuroscience and Biobehavioral Reviews*, **20**, 113-117.
- Hollingworth, H. L., & Poffenberger, A. T. (1917). *The sense of taste*. New York, NY: Moffat Yard.
- Holt-Hansen, K. (1968). Taste and pitch. *Perceptual & Motor Skills*, **27**, 59-68.
- Holt-Hansen, K. (1976). Extraordinary experiences during cross-modal perception. *Perceptual & Motor Skills*, **43**, 1023-1027.
- Hovland, C. I., Harvey, O. J., & Sherif, M. (1957). Assimilation and contrast effects in reactions to communication and attitude change. *Journal of Abnormal Social Psychology*, **55**, 244-252.
- Howard, I. P., & Templeton, W. B. (1966). *Human Spatial Orientation*. New York, NY: Wiley.
- Hughson, A. L., & Boakes, R. A. (2002). The knowing nose: The role of knowledge in wine expertise. *Food Quality and Preference*, **13**, 463-472.

- Iordanescu, L., Grabowecky, M., Franconeri, S., Theeuwes, J., & Suzuki, S. (2010). Characteristic sounds make you look at the targets faster in visual search. *Attention, Perception, & Psychophysics*, **72**, 1736-1741.
- Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sound facilitates visual search. *Psychonomic Bulletin & Review*, **15**, 548-554.
- Ishii, R., & O'Mahony, M. (1987). Taste sorting and naming: Can taste concepts be misrepresented by traditional psychophysical labelling systems? *Chemical Senses*, **12**, 37-51.
- Jackson, C. V. (1953). Visual factors in auditory localization. *Quarterly Journal of Experimental Psychology*, **5**, 52-65.
- Janssens, P. L. H. R., Hursel, R., Martens, E. A. P., & Westerterp-Plantenga, M. S. (2013). Acute effects of capsaicin on energy expenditure and fat oxidation in negative energy balance. *PLoS ONE*, **8**:e67786.
- Johnson, J. L., & Clydesdale, F. M. (1982). Perceived sweetness and redness in colored sucrose solutions. *Journal of Food Science*, **47**, 747-752.
- Jonides, J., & Mack, R. (1984). On the cost and benefit of cost and benefit. *Psychological Bulletin*, **96**, 29-44.
- Kantono, K., Hamid, N., Sheperd, D., Yoo, M. J. Y., Carr, B. T., & Grazioli, G. (2015). The effect of background music on food pleasantness ratings. *Psychology of Music*, **44**, 1111-1125.

- Kantono, K., Hamid, N., Shepherd, D., Yoo, M. J., Grazioli, G., & Carr, B. T. (2016). Listening to music can influence hedonic and sensory perceptions of gelati. *Appetite*, **100**, 244-255.
- Kappes, S. M., Schmidt, S. J., & Lee, S.-Y. (2006). Color halo/horns and halo-attribute dumping effects within descriptive analysis of carbonated beverages. *Journal of Food Science*, **71**, S590-S595.
- Karalus, M., Pontet, C., & Vickers, Z. (2010). *Experimentally created intensity scales for the five basic tastes: Sweet, sour, salty, bitter, and umami*. Unpublished results.
- Kerr, A. C. (1961). *The physiological regulation of salivary secretions in man*. Oxford, UK: Pergamon.
- King, B. M., & Duineveld, C. A. A. (1999). Changes in bitterness as beer ages naturally. *Food Quality and Preference*, **10**, 315-324.
- Klapetek, A., Ngo, M. K., & Spence, C. (2012). Do crossmodal correspondences enhance the facilitatory effect of auditory cues on visual search? *Attention, Perception, & Psychophysics*, **74**, 1154-1167.
- Knapton, S. (2014). To cut down on sugar just change the background music. *The Telegraph*, **June 14th**. Retrieved from <http://www.telegraph.co.uk/news/science/science-news/10898123/To-cut-down-on-sugar-just-change-the-background-music.html> on 10/05/2016.
- Knight, T. (2012). Bacon: The slice of life. In C. Vega, J. Ubbink, & E. van der Linden (Eds.), *The kitchen as laboratory: Reflections on the science of food and cooking* (pp. 73-82). New York, NY: Columbia University Press.

- Knoeferle, K. M., Woods, A., Käppler, F., & Spence, C. (2015). That sounds sweet: Using crossmodal correspondences to communicate gustatory attributes. *Psychology & Marketing*, **32**, 107-120.
- Knöferle, K. M., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic Bulletin & Review*, **19**, 992-1006.
- Knöferle, K. M., Knöferle, P., Velasco, C., & Spence, C. (2016). Multisensory brand search: How the meaning of sounds guides consumers' visual attention. *Journal of Experimental Psychology: Applied*, **22**, 196-210.
- Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *Trends in Cognitive Sciences*, **11**, 16-22.
- Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York, NY: Liveright.
- Kontukoski, M., Luomala, H., Mesz, B., Sigman, M., Trevisan, M., Rotola-Pukkila, M., et al. (2015). Sweet and sour: Music and taste associations. *Nutrition and Food Science*, **45**, 357-376.
- Krishna, A., Morrin, M., & Sayin, E. (2014). Smellizing cookies and salivating: A focus on olfactory imagery. *Journal of Consumer Research*, **41**, 18-34.
- Kron, A., Schul, Y., Cohen, A., & Hassin, R. R. (2010). Feelings don't come easy: Studies on the effortful nature of feelings. *Journal of Experimental Psychology: General*, **139**, 520-534.

- Labroo, A. A., Dhar, R., & Schwartz, N. (2008). Of frog wines and frowning watches: Semantic priming, perceptual fluency, and brand evaluation. *Journal of Consumer Research*, **34**, 819-831.
- Ladefoged, P., & Johnson, K. (2011). *A course in phonetics* (6th Ed.). Boston, MA: Wadsworth Publishing Company.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical report A-8*. University of Florida, Gainesville, FL.
- Lavie, N. (1995) Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 451-468.
- Lavie, N., & Cox, S. (1997) On the efficiency of attentional selection: Efficient visual search results in inefficient rejection of distraction. *Psychological Science*, **8**, 395-398.
- Lee, B. Y., Bartsch, S. M., Mui, Y., Haidari, L. A., Spiker, M. L., & Gittelsohn, J. (2017). A systems approach to obesity. *Nutrition Reviews*, **75**, 94-106.
- Lee, L., Frederick, S., & Ariely, D. (2006). Try it, you'll like it: The influence of expectation, consumption, and revelation on preferences for beer. *Psychological Science*, **17**, 1054-1058.
- Lee, W. E., III, & Pangborn, R. M. (1986). Time-intensity: The temporal aspects of sensory perception. *Food Technology*, **40**, 71-78. 82.
- Lelièvre, M., Chollet, S., Abdi, H., & Valentin, D. (2009). Beer-trained and untrained assessors rely more on vision than on taste when they categorize beers. *Chemosensory Perception*, **2**, 143-153.

- Lewkowicz, D. J., & Minar, N. J. (2014). Infants are not sensitive to synesthetic cross-modality correspondences. A comment to Walker et al. (2010). *Psychological Science*, **25**, 832-834.
- Lim, J., Fujimaru, T., & Linscott, T. D. (2014). The role of congruency in taste-odor interactions. *Food Quality & Preference*, **34**, 5-13.
- Lim, J., & Johnson, M. (2012). The role of congruency in retronasal odor referral to the mouth. *Chemical Senses*, **37**, 515-521.
- Lim, J., & Johnson, M. B. (2011). Potential mechanisms of retronasal odor referral to the mouth. *Chemical Senses*, **36**, 283-289.
- Lin, I. Y. (2010). The interactive effect of Gestalt situations and arousal seeking tendency on customers' emotional responses: Matching color and music to specific servicescapes. *Journal of Services Marketing*, **24**, 294-304.
- Linscott, T. D., & Lim, J. (2016). Retronasal odor enhancement by salty and umami tastes. *Food Quality & Preference*, **48**, 1-10.
- Litt, A., & Shiv, B. (2012). Manipulating basic taste perception to explore how product information affects experience. *Journal of Consumer Psychology*, **22**, 55-66.
- Logeswaran, N., & Bhattacharya, J. (2009). Crossmodal transfer of emotion by music. *Neuroscience Letters*, **455**, 129-133.
- Ludwig, V. U., Adachi, I., & Matzuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (*Pan troglodytes*) and humans. *Proceedings of the National Academy of Sciences USA*, **108**, 20661-20665.
- Luria, A. R. (1968). *The mind of a mnemonist*. Cambridge, MA: Harvard University Press.

- Macdonald, J. S. P., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics*, **73**, 1780-1789.
- MacGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, **264**, 746-748.
- Macht, M. (2008). How emotions affect eating: A five-way model. *Appetite*, **50**, 1-11.
- Maeda, F., Kanai, R., & Shimojo, S. (2004). Changing pitch induced visual motion illusion. *Current Biology*, **14**, R990-R991.
- Marks, L. E. (1987). On cross-modal similarity: Auditory-visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, **13**, 384-394.
- Marks, L. E. (1989). On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 586-602.
- Marks, L. E. (2004). Cross-modal interactions in speeded classification. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The Handbook of Multisensory Processes* (pp. 85-106). Cambridge, MA: MIT Press.
- Marks, L. E., Ben-Artzi, E., & Lakatos, S. (2003). Cross-modal interactions in auditory and visual discrimination. *International Journal of Psychophysiology*, **50**, 125-145.
- Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development*, **52**, 1-102.
- Marks, L. E., Stevens, J. C., Bartoshuk, L. M., Gent, J. F., Rifkin, B., & Stone, V. K. (1988). Magnitude-matching: The measurement of taste and smell. *Chemical Senses*, **13**, 63-87.

- Marks, L. E., & Wheeler, M. E. (1998). Attention and the detectability of weak taste stimuli. *Chemical Senses*, **23**, 19-29.
- Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, **28**, 903-923.
- Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face. *Psychonomic Bulletin and Review*, **3**, 215-221.
- Mattingley, J. B., Payne, J. M., & Rich, A. N. (2006). Attentional load attenuates synaesthetic priming effects in grapheme-colour synaesthesia. *Cortex*, **42**, 213-221.
- McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology*, **20**, 1035-1041.
- McElrea, H., & Standing, F. (1992). Fast music causes fast drinking. *Perceptual and Motor Skills*, **75**, 362.
- Meillon, S., Urbano, C., & Schlich, P. (2009). Contribution of the Temporal Dominance of Sensations (TDS) method to the sensory description of subtle differences in partially dealcoholized red wines. *Food Quality and Preference*, **20**, 490-499.
- Meiselman, H., & Dzendolet, E. (1967). Variability in gustatory quality identification. *Perception & Psychophysics*, **2**, 496-498.
- Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 69-79.
- Melara, R. D., & O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, **116**, 323-336.

- Mesz, B., Sigman, M., & Trevisan, M. A. (2012). A composition algorithm based on crossmodal taste-music correspondences. *Frontiers in Human Neuroscience*, **6**:71.
- Mesz, B., Trevisan, M. A., & Sigman, M. (2011). The taste of music. *Perception*, **40**, 209-219.
- Mitchell, D. (2012). A time and a place for a peach: Taste trends in contemporary cooking. *The Senses and Society*, **7**, 135-152.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, **132**, 297-326.
- Morrot, G., Brochet, F., & Dubourdieu, D. (2001). The color of odors. *Brain & Language*, **79**, 309-320.
- Mroczo-Wąsowicz, A. (2016). The unity and complexity of flavour perception: multisensory phenomenology and processing in human-food interactions. In A. Nijholt, C. Velasco, G. Huisman, & K. Karunanayaka (Eds.), *Proceedings of the 1st Workshop on Multi-sensorial Approaches to Human-Food Interaction (MHFI '16)* (Article 6). New York, NY: ACM.
- Mudd, S. A. (1963). Spatial stereotypes of four dimensions of pure tone. *Journal of Experimental Psychology*, **66**, 347-352.
- Nederkorn, C., de Wit, T., Smulders, F., & Jansen, A. (2001). Experimental comparison of different techniques to measure saliva. *Appetite*, **37**, 251-252.
- Nehrkorn, J., Tanimoto, H., Herz, A. V. M., & Yarali, A. (2015). A model for non-monotonic intensity coding. *Royal Society Open Science*, **2**, 150120.

- Ng, M., Lawlor, J. B., Chandra, S., Chaya, C., Hewson, L., & Hort, J. (2012). Using quantitative descriptive analysis and temporal dominance of sensations analysis as complementary methods for profiling commercial blackcurrant squashes. *Food Quality and Preference*, **25**, 121-134.
- Ngo, M., Misra, R., & Spence, C. (2011). Assessing the shapes and speech sounds that people associate with chocolate samples varying in cocoa content. *Food Quality and Preference*, **22**, 567-572.
- Nishizawa, M, Jiang, W., & Okajima, K. (2016). Projective-AR system for customizing the appearance and taste of food. *Proceedings of 18th International Conference on Multimodal Interaction*, **6**, doi: 10.1145/3001959.3001966.
- North, A. C. (2012). The effect of background music on the taste of wine. *British Journal of Psychology*, **103**, 293-301.
- North, A., & Hargreaves, D. (2008). *The social and applied psychology of music*. Oxford, UK: Oxford University Press.
- North, A. C., Hargreaves, D. J., & McKendrick, J. (1999). The influence of in-store music on wine selections. *Journal of Applied Psychology*, **84**, 271-276.
- North, A. C., Shilcock, A., & Hargreaves, D. J. (2003). The effect of musical style on restaurant customers' spending. *Environment and Behavior*, **35**, 712-718.
- Nunes, C. A., & Pinheiro, A. C. M. (2015). *Sensomaker User Guide*. Downloaded from <http://ufla.br/sensomaker/>.
- O'Doherty, J. P., Deichmann, R., Critchley, H. D., & Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, **33**, 815-826.

- Obrist, M., Comber, R., Subramanian, S., Piqueras-Fiszman, B., Velasco, C., & Spence, C. (2014). Temporal, affective, and embodied characteristics of taste experiences: A framework for design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, 2853-2862.
- Odgaard, E. C., Arieh, Y., & Marks, L. E. (2003). Cross-modal enhancement of perceived brightness: Sensory interaction versus response bias. *Perception & Psychophysics*, **65**, 123-132.
- Olson, J. C., & Dover, P. (1976). Effects of expectation creation and disconfirmation on belief elements of cognitive structure. In *Advances in Consumer Research* (B. B. Anderson, Ed.), Association for Consumer Research, **3**, 168-175.
- O'Mahony, M. (1973). A note on the reliability of subjective judgement in discerning when the mouth is clear of salt taste stimuli. *Perception & Psychophysics*, **14**, 437-439.
- O'Mahony, M., Goldenberg, M., Stedmon, J., & Alford, J. (1979). Confusion in the use of the taste adjectives 'sour' and 'bitter'. *Chemical Senses and Flavour*, **4**, 301-318.
- Orchard-Mills, E., Alais, D., & Van der Burg, E. (2013). Amplitude-modulated auditory stimuli influence selection of visual spatial frequencies. *Journal of Vision*, **13(3)**:6.
- Orchard-Mills, E., Van der Burg, E., & Alais, D. (2016). Crossmodal correspondence between auditory pitch and visual elevation affects temporal ventriloquism. *Perception*, **45**, 409-424.
- Owen, D. H., & Machamer, P. K. (1979). Bias-free improvement in wine discrimination. *Perception*, **8**, 199-209.

- Palmer, S. E., Langlois, T. A., & Schloss, K. B. (2015). Music-to-color associations of single-line piano melodies in non-synesthetes. *Multisensory Research*, **29**, 157-193.
- Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music-color associations are mediated by emotion. *Proceedings of the National Academy of Sciences of the USA*, **110**, 8836-8841.
- Pangborn, R. M. (1968). Parotid flow stimulated by the sight, feel and odor of lemon. *Perceptual & Motor Skills*, **27**, 1340-1342.
- Pangborn, R. M., Witherly, S. A., & Jones, F. (1979). Parotid and whole-mouth secretion in response to viewing, handling, and sniffing food. *Perception*, **8**, 339-346.
- Parise, C., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience Letters*, **442**, 257-261.
- Parise, C., & Spence, C. (2013). Audiovisual cross-modal correspondences in the general population. In J. Simner & E. Hubbard (Eds.), *The Oxford handbook of synaesthesia* (pp.790-815). Oxford, UK: Oxford University Press.
- Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 6104-6108.
- Parise, C. V., & Spence, C. (2009). “When birds of a feather flock together”: Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS ONE*, **4**, e5664.
- Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: An IAT study. *Experimental Brain Research*, **220**, 319-333.

- Parr, W. V., Heatherbell, D., & White, K. G. (2002). Demystifying wine expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses*, **27**, 747-755.
- Parr, W. V., White, K. G., & Heatherbell, D. A. (2003). The nose knows: Influence of colour on perception of wine aroma. *Journal of Wine Research*, **14**, 79-101.
- Patching, G. R., & Quinlan, P. T. (2002). Garner and congruence effects in the speeded classification of bimodal signals. *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 755-775.
- Pazart, L., Comte, A., Magnin, E., Millot, J.-L., & Moulin, T. (2014). An fMRI study on the influence of sommeliers' expertise on the integration of flavor. *Frontiers in Behavioral Neuroscience*, **8**:358.
- Pedersen, A.M., Bardow, A., Beier Jensen, S., & Nauntofte, B. (2002). Saliva and gastrointestinal functions of taste, mastication, swallowing and digestion. *Oral Diseases*, **8**, 117-129.
- Peng, Y., Gillis-Smith, S., Jin, H., Tränkner, D., Ryba, N. J. P., & Zuker, C. S. (2015). Sweet and bitter taste in the brain of awake behaving animals. *Nature*, **527**, 512-515.
- Peterson, J. M., Bartoshuk, L. M., & Duffy, V. B. (1999). Intensity and preference for sweetness is influenced by genetic taste variation. *Journal of the American Dietetic Association*, **99**, A28.
- Pineau, N., Cordelle, S., & Schlich, P. (2003). Temporal Dominance of Sensations: A new technique to record several sensory attributes simultaneously over time. In *5th Pangborn Symposium*. July 20-24th, Boston, USA.

- Pineau, N., Schlich, P., Cordelle, S., Mathonnière, C., Issanchou, S., Imbert, A., Rogeaux, M., Etiévant, P., & Köster, E. (2009). Temporal dominance of sensations: Construction of the TDS curves and comparison with time-intensity. *Food Quality and Preference*, **20**, 450-455.
- Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, **41**, 983-993.
- Pinheiro, A. C. M., Nunes, C. A., & Viotoris, V. (2013). Sensomaker: A tool for sensorial characterization of food products. *Ciência e Agrotecnologia*, **37**, 199-201.
- Piqueras-Fiszman, B., Alcaide, J., Roura, E., & Spence, C. (2012). Is it the plate or is it the food? Assessing the influence of the color (black or white) and shape of the plate on the perception of the food placed on it. *Food Quality and Preference*, **24**, 205-208.
- Piqueras-Fiszman, B., & Spence, C. (2012). The influence of the feel of product packaging on the perception of the oral-somatosensory texture of food. *Food Quality & Preference*, **26**, 67-73.
- Piqueras-Fiszman, B., & Spence, C. (2015). Sensory expectations based on product-extrinsic food cues: An interdisciplinary review of the empirical evidence and theoretical accounts. *Food Quality & Preference*, **40**, 165-179.
- Pitteri, M., Marchetti, M., Priftis, K., & Grassi, M. (2017). Naturally together: Pitch-height and brightness as coupled factors for eliciting the SMARC effect in non-musicians. *Psychological Research*, **81**, 243-254.

- Poellinger, A., Thomas, R., Lio, P., Lee, A., Makris, N., Rosen, B. R., & Kwong, K. K. (2001). Activation and habituation in olfaction – an fMRI study. *NeuroImage*, **13**, 547-560.
- Prescott, J., Allen, S., & Stephens, L. (1993). Interactions between oral chemical irritation taste and temperature. *Chemical Senses*, **18**, 389-404.
- Prescott, J., & Stevenson, R. J. (1995). Effects of oral chemical irritation, on tastes and flavors in frequent and infrequent users of chili. *Physiology & Behavior*, **58**, 1117-1127.
- Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, **132**, 416-442.
- Ratcliffe, V. F., Taylor, A. M., & Reby, D. (2016). Cross-modal correspondences in non-human mammal communication. *Multisensory Research*, **29**, 49-91.
- Reinoso Carvalho, F., Van Ee, R., Rychtarikova, M., Touhafi, A., Steenhaut, K., Persoone, D., Spence, C., & Leman, M. (2015a). Does music influence the multisensory tasting experience? *Journal of Sensory Studies*, **30**, 404-412.
- Reinoso Carvalho, F., Van Ee, R., Rychtarikova, M., Touhafi, A., Steenhaut, K., Persoone, D., & Spence, C. (2015b). Using sound-taste correspondences to enhance the subjective value of tasting experiences. *Frontiers in Psychology*, **6**:1309.
- Reinoso Carvalho, F., Van Ee, R., & Touhafi, A. (2013). T.A.S.T.E. Testing Auditory Solutions Towards the improvement of the Tasting Experience. In *Proceedings of 10th international symposium on computer music multidisciplinary research*. Marseille, France: L.M.A. Publications.

- Reinoso Carvalho, F., Van Ee, R., Touhafi, A., Steenhaut, K., Leman, M. & Rychtarikova, M., (2015c). Assessing multisensory tasting experiences by means of customized sonic cues. In *Proceedings of Euronoise 2015*. Maastricht, The Netherlands.
- Reinoso Carvalho, F., Wang, Q. (J.), Van Ee, R., & Spence, C. (2016). The influence of soundscapes on the perception and evaluation of beers. *Food Quality and Preference*, **52**, 32-41.
- Reiter, S., Rodriguez, C. C., Sun, K., & Stopfer, M. (2015). Spatiotemporal coding of individual chemicals by the gustatory system. *Journal of Neuroscience*, **35**, 12309-12321.
- Rich, A. N., & Mattingley, J. B. (2002). Anomalous perception in synaesthesia: A cognitive neuroscience perspective. *Nature Reviews Neuroscience*, **3**, 43-52.
- Rigg, M. G. (1937). An experiment to determine how accurately college students can interpret intended meanings of musical compositions. *Journal of Experimental Psychology*, **21**, 223-229.
- Robin, O., Rousmans, S., Dittmar, A., & Vernet-Maury, E. (2003). Gender influence on emotional responses to primary tastes. *Physiology & Behavior*, **78**, 385-393.
- Robinson, J. O. (1970). The misuse of taste names by untrained observers. *British Journal of Psychology*. **61**, 375-378.
- Rock, I., & Victor, J. (1964). Vision and touch: an experimentally created conflict between the two senses. *Science*, **143**, 594-596.
- Rolls, E. T. (2015). Taste, olfactory, and food reward value processing in the brain. *Progress in Neurobiology*, **15**, 1-27.

- Rosenstein, D., & Oster, H. (1988). Differential facial responses to four basic tastes in newborns. *Child Development*, **59**, 1555-1568.
- Rousmans, S., Robin, O., Dittmar, A., & Vernet-Maury, E. (2000). Autonomic nervous system responses associated with primary tastes. *Chemical Senses*, **25**, 709-718.
- Royet, J.-P., Plailly, J., Saive, A.-L., Veyrac, A., & Delon-Martin, C. (2013). The impact of expertise in olfaction. *Frontiers in Psychology*, **4**:928.
- Rozin, P. (1982). “Taste-smell confusions” and the duality of the olfactory sense. *Perception & Psychophysics*, **31**, 397-401.
- Rudmin, F., & Cappelli, M. (1983). Tone-taste synesthesia: A replication. *Perceptual & Motor Skills*, **56**, 118.
- Salgado-Montejo, A., Alvarado, J. A., Velasco, C., Salgado, C. J., Hasse, K., & Spence, C. (2015). The sweetest thing: The influence of angularity, symmetry, and the number of elements on shape-valence and shape-taste matches. *Frontiers in Psychology*, **6**:1382.
- Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, **99**, 143-165.
- Schiffstein, H. N. J., & Tanudjaja, I. (2004). Visualizing fragrances through colors: The mediating role of emotions. *Perception*, **33**, 1249-1266.
- Seo, H.-S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G., & Hummel, T. (2010). Cross-modal integration between odors and abstract symbols. *Neuroscience Letters*, **478**, 175-178.
- Seo, H.-S., & Hummel, T. (2011). Auditory–olfactory integration: Congruent or pleasant sounds amplify odor pleasantness. *Chemical Senses*, **36**, 301-309.

- Seo, H.-S., Lohse, F., Luckett, C. R., & Hummel, T. (2014). Congruent sound can modulate odor pleasantness. *Chemical Senses*, **39**, 215-228.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear: Sound induced visual flashing. *Nature*, **408**, 788.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, **14**, 147-152.
- Shankar, M. U., Levitan, C., & Spence, C. (2010). Grape expectations: The role of cognitive influences in color-flavor interactions *Consciousness & Cognition*, **19**, 380-390.
- Shannon, I. L. Suddick R. P. & Dowd, Jr, E. J. (1974). *Saliva: Composition and secretion*. In H. M. Myers (Ed.), *Monographs in oral science* (Vol. 2, pp. 1-103). Basel, Switzerland: Karger.
- Sherman, A. (2011). Perfect pairings: Programming wine with music. *NPR Music*. **November 22**. <http://www.npr.org/sections/deceptivecadence/2011/11/21/142590064/perfect-pairings-programming-wine-with-music>.
- Shermer, D. Z., & Levitan, C. A. (2014). Red hot: The crossmodal effect of color intensity on perceived piquancy. *Multisensory Research*, **27**, 207-223.
- Silvestrini, N., Piguet, V., Cedraschi, C., & Zentner, M. R. (2011). Music and auditory distraction reduce pain: Emotional or attentional effects? *Music and Medicine*, **3**, 264-270.
- Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex*, **15**, 1602-1608.

- Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. *Perception*, **39**, 553-569.
- Singh, M. (2014). Mood, food, and obesity. *Frontiers in Psychology*, **5**:925.
- Slocombe, B. G., Carmichael, D. A., & Simner, J. (2016). Cross-modal tactile-taste interactions in food evaluations. *Neuropsychologia*, **88**, 58-64.
- Small, D. M. (2010). Taste representation in the human insula. *Brain Structure and Function*, **214**, 551-561.
- Small, D. M., Jones-Gotman, M., Zatorre, R. J., Petrides, M., & Evans, A. C. (1997). Flavor processing: More than the sum of its parts. *Neuroreport*, **8**, 3913-3917.
- Small, D. M., & Prescott, J. (2005). Odor/taste integration and the perception of flavor. *Experimental Brain Research*, **166**, 345-357.
- Small, D. M., Voss, J., Mak, Y. E., Simmons, K. B., Parrish, T., & Gitelman, D. (2004). Experience-dependent neural integration of taste and smell in the human brain. *Journal of Neurophysiology*, **92**, 1892-1903.
- Small, D. M., Zald, D. H., Jones-Gotman, M., Zatorre, R. J., Pardo, J. V., Frey, S., & Petrides, M. (1999). Human cortical gustatory areas: A review of functional neuroimaging data. *NeuroReport*, **10**, 7-14.
- Smith, B. C. (Ed.). (2007). *Questions of taste: The philosophy of wine*. Oxford, UK: Oxford University Press.
- Smith, C. (2010). Astringency and harmony in tannins. *Wines & Vines*.
<http://www.winesandvines.com/template.cfm?section=columns> August 2015.

- Smith, E. L., Grabowecky, M., & Suzuki, S. (2007). Auditory-visual crossmodal integration in perception of face gender. *Current Biology*, **17**, 1680-1685.
- Sobel, N., Prabhakaran, V., Zhao, Z., Desmond, J. E., Glover, G. H., Sullivan, E. V., & Gabrieli, J. D. E. (2000). Time course of odorant-induced activation in the human primary olfactory cortex. *Journal of Neurophysiology*, **83**, 537-551.
- Sokolowsky, M., & Fischer, U. (2012). Evaluation of bitterness in white wine applying descriptive analysis, time-intensity analysis, and temporal dominance of sensations analysis. *Analytica Chimica Acta*, **732**, 46-52.
- Spence, C. (2007). Audiovisual multisensory integration. *Acoustical Science & Technology*, **28**, 61-70.
- Spence, C. (2010). Crossmodal spatial attention. *Annals of the New York Academy of Sciences (The Year in Cognitive Neuroscience)*, **1191**, 182-200.
- Spence, C. (2011a). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, **73**, 971-995.
- Spence, C. (2011b). Mouth-watering: The influence of environmental and cognitive factors on salivation and gustatory/flavour perception. *Journal of Texture Studies*, **42**, 157-171.
- Spence, C. (2011c). The multisensory perception of touch. In F. Bacci & D. Melcher (Eds.), *Art and the senses* (pp. 85-106). Oxford, UK: Oxford University Press.
- Spence, C. (2012a). Auditory contributions to flavour perception and feeding behaviour. *Physiology & Behavior*, **107**, 505-515.

- Spence, C. (2012b). Managing sensory expectations concerning products and brands: Capitalizing on the potential of sound and shape symbolism. *Journal of Consumer Psychology*, **22**, 37-54.
- Spence, C. (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences*, **1296**, 31-49.
- Spence, C. (2014a) Assessing the influence of shape and sound symbolism on the consumer's response to chocolate. *New Food Magazine*, **17**, 59-62.
- Spence, C. (2014b). Noise and its impact on the perception of food and drink. *Flavour*, **3**:9.
- Spence, C. (2014c). Orienting attention: A crossmodal perspective. In A. C. Nobre & S. Kastner (Eds.), *The Oxford handbook of attention* (pp.446-471). Oxford, UK: Oxford University Press.
- Spence, C. (2015a). Cross-modal perceptual organization. In J. Wagemans (Ed.), *The Oxford Handbook of Perceptual Organization* (pp. 649-664). Oxford, UK: Oxford University Press.
- Spence, C. (2015b). Eating with our ears: Assessing the importance of the sounds of consumption to our perception and enjoyment of multisensory flavour experiences. *Flavour*, **4**:3.
- Spence, C. (2015c). Multisensory flavor perception. *Cell*, **161**, 24-35.
- Spence, C. (2015d). Music from the kitchen. *Flavour*, **4**:25.
- Spence, C. (2015e). On the psychological impact of food colour. *Flavour*, **4**:21.
- Spence, C. (2016a). Oral referral: On the mislocalization of odours to the mouth. *Food Quality and Preference*, **50**, 117-128.

- Spence, C. (2016b). Sound bites and digital seasoning. In V. Sorensen (Ed.), *Sound, media art, and the metaverse* (pp. 9-15). [ICMA Array (Lindborg, P.-M., & Styles, S. Eds.), Vol 2016, Special Issue: Proceedings of Si15, Singapore, August 2015].
- Spence, C. (2016c). The neuroscience of flavor. In B. Piqueras-Fiszman & C. Spence (Eds.), *Multisensory flavor perception: From fundamental neuroscience through to the marketplace* (pp. 235-248). Oxford, UK: Elsevier.
- Spence, C. (2017). The neuroscience of flavour. In N. Levent & I. D. Mihalache (Eds.), *Food and museums* (pp. 57-70). London, UK: Bloomsbury Academic.
- Spence, C., & Chen, Y.-C. (2012). Intramodal and crossmodal perceptual grouping. In B. E. Stein (Ed.), *The New Handbook of Multisensory Processing* (pp. 265-282). Cambridge, MA: MIT Press.
- Spence, C., & Deroy, O. (2012). On the shapes of tastes and flavours. *Petits Propos Culinaires*, **97**, 75-108.
- Spence, C., & Deroy, O. (2013a). How automatic are crossmodal correspondences? *Consciousness and Cognition*, **22**, 245-260.
- Spence, C., & Deroy, O. (2013b). On why music changes what (we think) we taste. *i-Perception*, **4**, 137-140.
- Spence, C., & Deroy, O. (2013c). Tasting shapes: A review of four hypotheses. *Theoria et Historia Scientiarum*, **10**, 207-238.
- Spence, C., Harrar, V., & Piqueras-Fiszman, B. (2012). Assessing the impact of the tableware and other contextual variables on multisensory flavour perception. *Flavour*, **1**:7.

- Spence, C., Levitan, C. A., Shankar, M. U., & Zampini, M. (2010). Does food color influence taste and flavor perception in humans? *Chemical Perception*, **3**, 68-84.
- Spence, C., Michel, C., & Smith, B. (2014). Airplane noise and the taste of umami. *Flavour*, **3**:2.
- Spence, C., & Piqueras-Fiszman, B. (2014). *The perfect meal: The multisensory science of food and dining*. Oxford, UK: John Wiley & Sons.
- Spence, C., Richards, L., Kjellin, E., Huhnt, A.-M., Daskal, V., Scheybeler, A., Velasco, C., & Deroy, O. (2013). Looking for crossmodal correspondences between classical music & fine wine. *Flavour*, **2**:29.
- Spence, C., & Santangelo, V. (2009). Capturing spatial attention with multisensory cues: A review. *Hearing Research*, **258**, 134-142.
- Spence, C., & Shankar, M. U. (2010). The influence of auditory cues on the perception of, and responses to, food and drink. *Journal of Sensory Studies*, **25**, 406-430.
- Spence, C., Shankar, M. U., & Blumenthal, H. (2011). 'Sound bites': Auditory contributions to the perception and consumption of food and drink. In F. Bacci & D. Melcher (Eds.), *Art and the senses* (pp. 207-238). Oxford, UK: Oxford University Press.
- Spence, C., Smith, B., & Auvray, M. (2015a). Confusing tastes and flavours. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 247-274). Oxford, UK: Oxford University Press.
- Spence, C., Velasco, C., & Knoeferle, K. (2014a). A large sample study on the influence of the multisensory environment on the wine drinking experience. *Flavour*, **3**:8.

- Spence, C., Velasco, C., Vanne, M., & Hopia, A. (2014b). Can you taste the music? In the *5D cookbook* (pp. 74-77). Finland. www.5Dcookbook.fi.
- Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., & Deroy, O. (2015b). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour*, **4**:23.
- Spence, C., & Wang, Q. (J.) (2015a). Sensory expectations elicited by the sounds of opening the packaging and pouring a beverage. *Flavour*, **4**:35.
- Spence, C., & Wang, Q. (J.) (2015b). Wine & music (I): On the crossmodal matching of wine & music. *Flavour*, **4**:34.
- Spence, C., & Wang, Q. (J.) (2015c). Wine & music (II): Can you taste the music? Modulating the experience of wine through music and sound. *Flavour*, **4**:33.
- Spence, C., & Wang, Q. (J.) (2015d). Wine & music (III): So what if music influences taste? *Flavour*, **4**:35.
- Spence, C., Youssef, J., & Deroy, O. (2015c). Where are all the synaesthetic chefs? *Flavour*, **4**:29.
- Staats, A. W., & Hammond, O. W. J. (1972). Natural words as physiological conditioned stimuli. Food-word-elicited salivation and deprivation effects. *Journal of Experimental Psychology*, **96**, 206-208.
- Stafford, L. D., Agobiani, E., & Fernandes, M. (2013). Perception of alcohol strength impaired by low and high volume distraction. *Food Quality and Preference*, **28**, 470-474.

- Stafford, L. D., Fernandes, M., & Agobiani, E. (2012). Effects of noise and distraction on alcohol perception. *Food Quality & Preference*, **24**, 218-224.
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. Cambridge, MA: MIT Press.
- Steiner, J. E., Glaser, D., Hawilo, M. E., & Berridge, K. C. (2001). Comparative expression of hedonic impact: Affective reactions to taste by human infants and other primates. *Neuroscience and Biobehavioral Reviews*, **25**, 53-74.
- Stevens, J. C., & Marks, L. E. (1965). Cross-modality matching of brightness and loudness. *Proceedings of the National Academy of Sciences*, **54**, 407-411.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, **64**, 153-181.
- Stevenson, R. J. (2009). *The psychology of flavour*. Oxford, UK: Oxford University Press.
- Stevenson, R. J. (2012). The role of attention in flavour perception. *Flavour*, **1**:2.
- Stevenson, R. J. (2014). Flavor binding: Its nature and cause. *Psychological Bulletin*, **140**, 487-510.
- Stevenson, R. J., & Boakes, R. A. (2004). Sweet and sour smells: Learned synaesthesia between the senses of taste and smell. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processing* (pp. 69-83). Cambridge, MA: MIT Press.
- Stevenson, R. J., Boakes, R. A., & Prescott, J. (1998). Changes in odor sweetness resulting from implicit learning of a simultaneous odor-sweetness association: An example of learned synesthesia. *Learning and Motivation*, **29**, 113-132.

- Stevenson, R. J., Oaten, M. J., & Mahmut, M. K. (2011). The role of attention in the localization of odors to the mouth. *Attention, Perception, & Psychophysics*, **73**, 247-258.
- Suda, M., Morimoto, K., Obata, A., Koizumi, H., & Maki, A. (2008). Emotional responses to music: Towards scientific perspectives on music therapy. *NeuroReport*, **19**, 75-78.
- Suzuki, Y., & Takeshima, H. (2004). Equal-loudness-level contours for pure tones. *Journal of the Acoustical Society of America*, **116**, 918-933.
- Szybillo, G. J., & Jacoby, J. (1974). Intrinsic versus extrinsic cues as determinants of perceived product quality. *Journal of Applied Psychology*, **59**, 74-78.
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, **14**, 400-410.
- Titchener, E. B. (1909). *A textbook of psychology*. New York, NY: Macmillan.
- Tu, Y., Yang, S., & Ma, C. (2016). The taste of plate: How the spiciness of food is affected by the color of the plate used to serve it. *Journal of Sensory Studies*, **31**, 50-60.
- Tuorila, H., Cardello, A. V., & Leshner, L. L. (1994). Antecedents and consequences of expectations related to fat-free and regular-fat foods. *Appetite*, **23**, 247-263.
- Turek, R. (1996). *The elements of music: Concepts and applications, Volume 1*. Akron, OH: McGraw-Hill.
- Van den Stock, J., Grèzes, J., & de Gelder, B. (2008). Human and animal sounds influence recognition of body language. *Brain Research*, **1242**, 185-190.

- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7*, 487-494.
- 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*, 1277-1284.
- Van Dillen, L. F., & Koole, S. L. (2007). Clearing the mind: A working memory model of distraction from negative feelings. *Emotion, 7*, 715-723.
- Van Opstal, F., & Verguts, T. (2013). Is there a generalized magnitude system in the brain? Behavioral, neuroimaging, and computational evidence. *Frontiers in Psychology, 4*:435.
- Velasco, C., Jones, R., King, S., & Spence, C. (2013a). Assessing the influence of the multisensory environment on the whisky drinking experience. *Flavour, 2*:23.
- Velasco, C., Jones, R., King, S., & Spence, C. (2013b). The sound of temperature: What information do pouring sounds convey concerning the temperature of a beverage. *Journal of Sensory Studies, 28*, 335-345.
- Velasco, C., Woods, A., Deroy, O., & Spence, C. (2015). Hedonic mediation of the crossmodal correspondence between taste and shape. *Food Quality & Preference, 41*, 151-158.
- Velasco, C., Woods, A., Liu, J., & Spence, C. (2016). Assessing the role of taste intensity and hedonics in taste-shape correspondences. *Multisensory Research, 29*, 209-221.
- Visalli, M. (2016). *Results of the analysis of the wine-cheese data collected during the social program*. Presentation at EuroSense 2016, Dijon, France.
- Walker, L., Walker, P., & Francis, B. (2012). A common scheme for cross-sensory correspondences across stimulus domains. *Perception, 41*, 1186-1192.

- Walker, P. (2012a). Cross-sensory correspondences and cross talk between dimensions of connotative meaning: Visual angularity is hard, high-pitched, and bright. *Attention, Perception, & Psychophysics*, **74**, 1792-1809.
- Walker, P. (2012b). Cross-sensory correspondences and naive conceptions of natural phenomena. *Perception*, **41**, 620-622.
- Walker, P. (2016a). Cross-sensory correspondences: A theoretical framework and their relevance to music. *Psychomusicology: Music, Mind, and Brain*, **26**, 103-116.
- Walker, P. (2016b). Cross-sensory correspondences and symbolism in spoken and written language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **42**, 1339-1361.
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, **21**, 21-25.
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2014). Preverbal infants are sensitive to cross-sensory correspondences: Much ado about the null results of Lewkowicz and Minar (2014). *Psychological Science*, **25**, 835-836.
- Walker, P., Scallon, G., & Francis, B. (2016). Cross-sensory correspondences: Heaviness is dark and low-pitched. *Perception*, doi:[10.1177/0301006616684369](https://doi.org/10.1177/0301006616684369).
- Walker, P., & Smith, S. (1984). Stroop interference based on the synaesthetic qualities of auditory pitch. *Perception*, **13**, 75-81.

- Walker, P., & Walker, L. (2012). Size-brightness correspondence: Crosstalk and congruity among dimensions of connotative meaning. *Attention, Perception, & Psychophysics*, **74**, 1226-1240.
- Walker, P., Walker, L., & Francis, B. (2015). The size–brightness correspondence: Evidence for crosstalk among aligned conceptual feature dimensions. *Attention, Perception, & Psychophysics*, **77**, 2694-2710.
- Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics*, **42**, 491-502.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quality. *Trends in Cognitive Sciences*, **7**, 483-488.
- Wan, X., Velasco, C., Michel, M., Mu, B., Woods, A. T., & Spence, C. (2014). Does the type of receptacle influence the crossmodal association between colour and flavour? A cross-cultural comparison. *Flavour*, **3**:3.
- Wan, X., Woods, A. T., Jacquot, M., Knoeferle, K., Kikutani, M., & Spence, C. (2016). The effects of receptacle on the expected flavor of a colored beverage: Cross-cultural comparison among French, Japanese, and Norwegian consumers. *Journal of Sensory Studies*, **31**, 233-244.
- Wan, X., Woods, A. T., Seoul, K. H., Butcher, N., & Spence, C. (2015). When the shape of the glass influences the flavour associated with a coloured beverage: Evidence from consumers in three countries. *Food Quality and Preference*, **39**, 109-116.
- Wang, G. J., Volkow, N. D., Telang, F., Jayne, M., Ma, J., Rao, M., ... Fowler, J. S. (2004). Exposure to appetitive food stimuli markedly activates the human brain. *NeuroImage*, **21**, 1790-1797.

- Wang, Q. J. (2013). *Music, mind, and mouth: Exploring the interaction between music and flavour perception*. MSc thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Wang, Q. (J.), Keller, S., & Spence, C. (2017). The sound of spiciness: Enhancing the evaluation of piquancy by means of a customized crossmodally congruent soundtrack. *Food Quality and Preference*, **58**, 1-9.
- Wang, Q. (J.), & Spence, C. (2015a). Assessing the effect of musical congruency on wine tasting in a live performance setting. *i-Perception*, **6**:3.
- Wang, Q. (J.), & Spence, C. (2015b). Assessing the influence of the multisensory atmosphere on the taste of vodka. *Beverages*, **1**, 204-217.
- Wang, Q. (J.), & Spence, C. (2016). ‘Striking a sour note’: Assessing the influence of consonant and dissonant music on taste perception. *Multisensory Research*, **29**, 195-208.
- Wang, Q. (J.), Wang, S., & Spence, C. (2016). “Turn up the taste”: Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and pitch. *Chemical Senses*, **29**, 345-356.
- Wang, Q. (J.), Woods, A., & Spence, C. (2015). “What’s your taste in music?” A comparison of the effectiveness of various soundscapes in evoking specific tastes. *i-Perception*, **6**:6.
- Watanabe, T., Kawada, T., & Iwal, K. (1987). Enhancement by capsaicin of energy metabolism in rats through secretion of catecholamine from adrenal medulla. *Agriculture and Biological Chemistry*, **51**, 75-79.

- Watt, R., & Quinn, S. (2007). Some robust higher-level percepts for music. *Perception*, **36**, 1834-1848.
- Wesson, D. W., & Wilson, D. A. (2010). Smelling sounds: Olfactory-auditory sensory convergence in the olfactory tubercle. *Journal of Neuroscience*, **30**, 3013-3021.
- Wesson, D. W., & Wilson, D. A. (2011). Sniffing out the contributions of the olfactory tubercle to the sense of smell: Hedonics, sensory integration, and more? *Neuroscience and Biobehavioral Reviews*, **35**, 655-668.
- Wheeler, E. (1938). *Tested sentences that sell*. New York, NY: Prentice & Co. Hall.
- White, K. D. (1977). Salivation: A review and experimental investigation of major techniques. *Psychophysiology*, **14**, 203-212.
- Wilson, S. (2003). The effect of music on perceived atmosphere and purchase intentions in a restaurant. *Psychology of Music*, **31**, 93-112.
- Winkielman, P., Schwarz, N., Fazendeiro, T., & Reber, R. (2003). The hedonic marking of processing fluency: Implications for evaluative judgment. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 189-217). Mahwah, NJ: Lawrence Erlbaum.
- Wise, P. M., & Brelin, P. A. S. (2011). Relationships among taste qualities assessed with response-context effects. *Chemical Senses*, **36**, 581-587.
- Wolpert, R. S. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. nonmusicians. *Music Perception*, **8**, 95-105.

- Wolter, S., Dudschig, C., & Kaup, B. (2016). Reading sentences describing high- or low-pitched auditory events: Only pianists show evidence for a horizontal space-pitch association. *Psychological Research*, doi:10.1007/s00426-016-0812-z.
- Woods, A. T., Lloyd, D. M., Kuenzel, J., Poliakoff, E., Dijksterhuis, G. B., & Thomas, A. (2011). Expected taste intensity affects response to sweet drinks in primary taste cortex. *Neuroreport*, **22**, 365-369.
- Woods, A. T., Poliakoff, E., Lloyd, D. M., Dijksterhuis, G. B., & Thomas, A. (2010). Flavour expectation: The effect of assuming homogeneity on drink perception. *Chemosensory Perception*, **3**, 174-181.
- Woods, A. T., Poliakoff, E., Lloyd, D. M., Kuenzel, J., Hodson, R., Gonda, H., Batchlelor, J., Dijksterhuis, G. B., & Thomas, A. (2011). Effect of background noise on food perception. *Food Quality and Preference*, **22**, 42-47.
- Wooley, S. C., & Wooley, O. W. (1973). Salivation to the sight and thought of food: A new measure of appetite. *Psychosomatic Medicine*, **35**, 136-142.
- Xu, A. J., & Labroo, A. A. (2014). Incandescent affect: Turning on the hot emotional system with bright light. *Journal of Consumer Psychology*, **24**, 207-216.
- Yamaguchi, S., & Takahashi, C. (1984). Hedonic functions of monosodium glutamate and four basic taste substances used at various concentration levels in single and complex systems. *Agricultural and Biological Chemistry*, **48**, 1077-1081.
- Yan, K. S., & Dando, R. (2015). A crossmodal role for audition in taste perception. *Journal of Experimental Psychology: Human Perception and Performance*, **41**, 590-596

- Yarmolinsky, D. A., Zuker, C. S., & Ryba, N. J. (2009). Common sense about taste: From mammals to insects. *Cell*, **139**, 234-244.
- Yeoh, J. P. S., & North, A. C. (2010). The effects of musical fit on choice between two competing foods. *Musicae Scientiae*, **14**, 127-138.
- Yeomans, M. R., Tepper, B. J., Rietzschel, J., & Prescott, J. (2007). Human hedonic responses to sweetness: Role of taste genetics and anatomy. *Physiology & Behavior*, **91**, 264-273.
- Yorkston, E., & Menon, G. (2004). A sound idea: Phonetic effects of brand names on consumer judgments. *Journal of Consumer Research*, **31**, 43-51.
- Zampini, M., & Spence, C. (2004). The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *Journal of Sensory Studies*, **19**, 347-363.
- Zampini, M., Wantling, E., Phillips, N., & Spence, C. (2008). Multisensory flavor perception: Assessing the influence of fruit acids and color cues on the perception of fruit-flavored beverages. *Food Quality and Preference*, **19**, 335-343.
- Zellner, D., Geller, T., Lyons, S., Pyper, A., & Riaz, K. (2017). Ethnic congruence of music and food affects food selection but not liking. *Food Quality & Preference*, **56**, 126-129.
- Zellner, D. A., & Kautz, M. A. (1990). Color affects perceived odor intensity. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 391-397.
- Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior and Development*, **21**, 483-492.

APPENDIX A The role of pitch and tempo in sound-temperature crossmodal correspondences

A.1 Introduction

In recent years, a growing body of empirical research has revealed various surprising yet robust crossmodal correspondences between auditory and gustatory stimulus attributes. For instance, people reliably associate a number of musical parameters, such as pitch, tempo, and timbre, with basic tastes (see Section 1.2 for a review). However, correspondences between sound and the oral-somatosensory attributes of the eating/drinking experience (e.g., temperature, texture, viscosity) have mostly been limited to those sounds that are related to the consumption of food products. The sounds associated with the opening of product packaging can, for instance, communicate freshness, while the sounds of a liquid being poured might indicate levels of carbonation and viscosity, or perhaps the shape of the container, not to mention the liquid's temperature (see Spence & Wang, 2015, for a review).

Velasco et al. (2013a) demonstrated that people (N=33) can reliably distinguish hot (82-84°C) from cold water (6-8°C) based on nothing more than the sounds of pouring. Moreover, the perceived temperature of a liquid can be artificially raised simply by enhancing the volume around 200 Hz and decreasing it at around 5-6 kHz (and vice versa to lower the perceived temperature). Such findings suggest that people might associate hot and cold temperatures with different sound frequencies, and what applies to pouring sounds could potentially also be extended to the case of music.

Experiments 15A and 15B assessed whether two basic auditory properties – namely tempo and pitch – are consistently associated with drinks of different temperatures.

First, an online pre-study (Experiment 15A) was performed in order to assess sound-temperature associations with the imagined experiences of drinking water at different temperatures. Next, the main study (Experiment 15B) was conducted using real water samples in order to validate the findings from the pre-study.

A.2 Experiment 15A: Pre-study

A.2.1 Methods

A.2.1.1 Participants

30 participants¹ (15 women, 15 men) between 21-52 years of age ($M=28.2$, $SD=7.4$) took part in the study. The participants gave their informed consent, and reported no hearing impairments. The participants were recruited from Prolific Academic. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

A.2.1.2 Auditory stimuli

The same short melody (see Figure A.1) was manipulated to have 5 levels of both pitch and tempo. I created the melody, in the major key, and was contained within a simple octave. In terms of pitch, the melody was shifted by an octave each time, resulting in melodies ranging from C3 (131 Hz) to C7 (2093 Hz). In terms of tempo, the melodies were presented at 60, 140, 220, 300, and 380 beats per minute. All of the soundtracks were RMS-equalised. One set of soundtracks was produced with GarageBand's Steinway Grand Piano plugin, and another with GarageBand's String

¹ Only 30 participants were recruited given prior research conducted with $N=33$ participants revealed that people were able to tell the difference between the sound of pouring hot versus cold water (Velasco et al., 2013a).

Ensemble plugin (samples of the manipulated melody can be heard at <https://soundcloud.com/janicewang09/sets/sound-of-cold-and-hot>). The soundtracks were presented in a 5x5 grid, with variations in tempo along the horizontal axis and variations in pitch along the vertical axis. Two versions of the grid were produced, one with tempo increasing to the right and pitch decreasing going down, and one with tempo decreasing to the right and pitch increasing going down.



Figure A.1. The original melody used in Experiment 15A and 15B.

A.2.1.3 Design and procedure

The study was programmed on the Qualtrics online survey platform. Before the actual study began, the participants specified their gender, age, and self-rated their musical expertise levels (none, up to 2 years training, 2-10 years training, or 10+ years training). The participants had to answer an audio-based question correctly in order to ensure sound playback was functional and to allow the participant to adjust the volume to a comfortable listening level.


























The participants were first given a practice trial in order to familiarise themselves with the 5x5 sound grid. They were instructed to imagine drinking a glass of warm milk, and to choose one sound from the grid that best matched the experience.

The study consisted of two blocks of six trials each. For each trial, the participants were presented with a line of instruction asking them to imagine drinking either cold, room-temperature, or hot water. Below the instruction, on the same page, they were presented with a 5x5 sound grid with either piano or with string ensemble instrumentation. Each grid was associated with a sound clip and a number, and

participants were asked to input the number of the sound clip that best matched the imagined drinking experience (see Figure A.2). The participants clicked “next” to advance to the next page. The two blocks were identical except for the order of auditory stimuli along each axis. In the first block, the sounds were arranged from fast to slow going from left to right, and low pitch to high pitch from top to bottom. The directions were reversed in the second block. The order in which the questions were asked in each block was randomised.

Imagine you are drinking a cup of **ice cold water**.

Please explore the grid of string instrument sounds below until you find the sound that best matches the experience of drinking cold **water**. Sound files may take a few seconds to load, please be patient.

	fastest tempo	faster tempo	medium tempo	slower tempo	slowest tempo
lowest pitch	1 	6 	11 	16 	21 
lower pitch	2 	7 	12 	17 	22 
medium pitch	3 	8 	13 	18 	23 
higher pitch	4 	9 	14 	19 	24 
highest pitch	5 	10 	15 	20 	25 

Which string sound best matched the experience of drinking cold water?

- | | | | | |
|-------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="radio"/> 1 | <input type="radio"/> 6 | <input type="radio"/> 11 | <input type="radio"/> 16 | <input type="radio"/> 21 |
| <input type="radio"/> 2 | <input type="radio"/> 7 | <input type="radio"/> 12 | <input type="radio"/> 17 | <input type="radio"/> 22 |
| <input type="radio"/> 3 | <input type="radio"/> 8 | <input type="radio"/> 13 | <input type="radio"/> 18 | <input type="radio"/> 23 |
| <input type="radio"/> 4 | <input type="radio"/> 9 | <input type="radio"/> 14 | <input type="radio"/> 19 | <input type="radio"/> 24 |
| <input type="radio"/> 5 | <input type="radio"/> 10 | <input type="radio"/> 15 | <input type="radio"/> 20 | <input type="radio"/> 25 |

Figure A.2. A screenshot of how participants might see a question in Experiment 15A. The participants were presented with an audio grid below the instructions, and were asked to choose the sound clip that best matched the imagined drinking experience by picking the number that corresponds to the number in the audio grid.

The study lasted for approximately 3 minutes and the participants were paid £0.60 for taking part.

A.2.2 Results and discussion

Pitch and tempo information was extracted from participants' melody choices (each of the possible 25 melody choices was associated with 1 out of 5 levels of pitch and tempo). The mean values of pitch and tempo for each temperature are shown in Figure A.3. A RM-MANOVA was conducted with temperature (hot, room-temperature, or cold) and instrumentation (piano, string ensemble) as factors, and with pitch and tempo as measures. The RM-MANOVA revealed a significant main effect of temperature ($F(4,26)=7.50$, $p<.0005$, Wilk's Lambda=0.46), but not of instrument type ($F(2,28)=0.73$, $p=.49$). Follow-up univariate tests with Greenhouse-Geisser corrections revealed that temperature had a significant main effect for both pitch ($F(1.34,38.95)=10.03$, $p=.001$, $\eta^2=0.26$) and tempo ($F(1.49,43.41)=4.58$, $p=.02$, $\eta^2 = 0.14$). The participants' choice of pitch was significantly higher for cold water than for room-temperature ($p<.0005$) or hot water ($p=.02$), and their choice of tempo was significantly higher for cold water than for room-temperature water ($p=.002$).

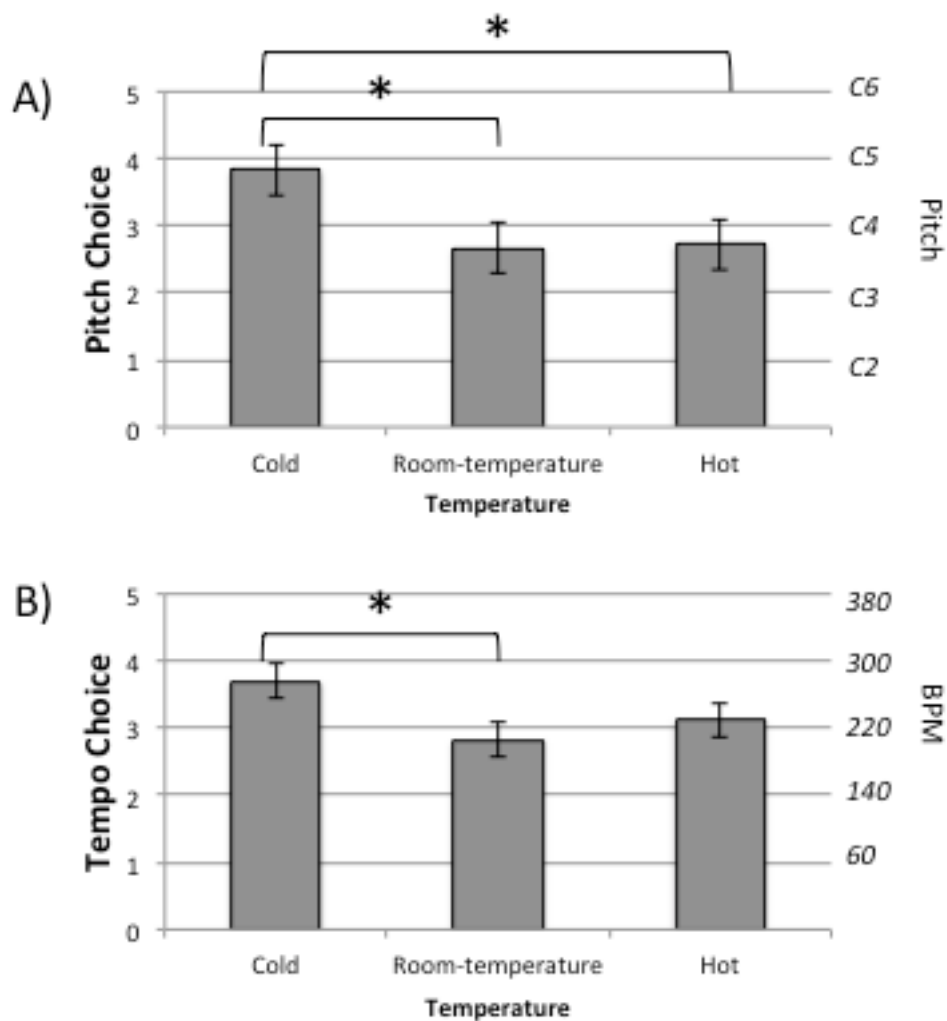


Figure A.3. Mean values of pitch (A) and tempo (B) for each temperature condition in Experiment 15A. Pitch ratings range from 1=C2 to 5=C6. Tempo ratings range from 1 - 80 beats per minute (BPM) to 5 - 330 BPM. Error bars indicate standard error. * indicate statistical significance at $p < .05$.

Overall, there was a significant positive correlation between pitch and tempo choices, $r_{179} = 0.47$, $p < .0005$. More specifically, a positive correlation between pitch and tempo was observed when the participants were asked to imagine drinking cold ($r_{60} = 0.54$, $p < .0005$) and hot water ($r_{60} = 0.30$, $p = .02$), but not when imagining drinking room-temperature water ($r_{60} = 0.20$, $p = .13$). In other words, high pitch tended to be associated with fast tempo, and low pitch with slow tempo, when it came to choosing the best matching melody for imagined cold and hot temperature drinks. This pattern

of results can be seen in the colour-coded participant response frequency grid in Figure A.4, where the tendency for responses to cluster around the pitch-tempo diagonal is greater for the cold water and hot water conditions than for the room-temperature condition.

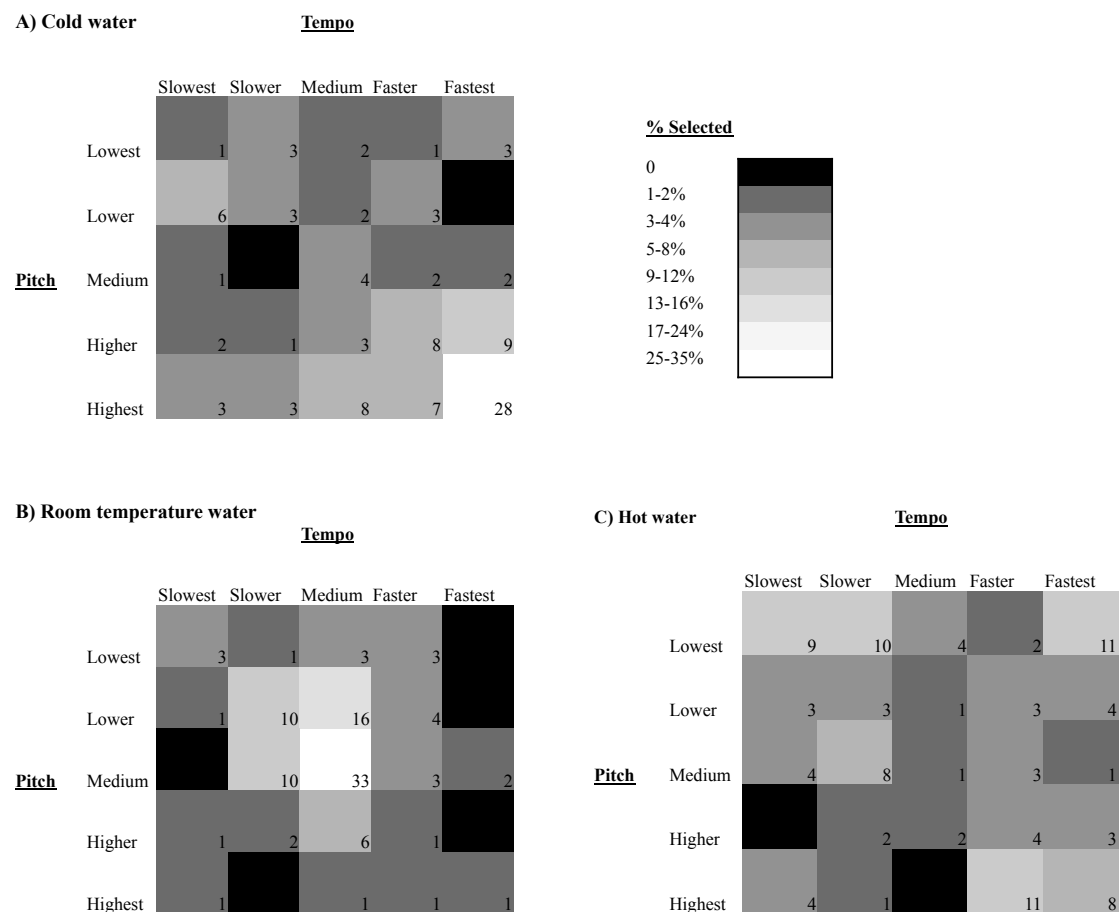


Figure A.4. Frequency information concerning how often participants selected each combination of tempo and pitch on the audio grid in Experiment 15A. Trials from both piano and string instrumentation sounds have been merged. The number in each box indicates the % of total responses.

The imagined experience of drinking cold water was associated with a melody having a higher pitch (around C5 or 523 Hz) as compared to room-temperature and hot water; and associated with a faster tempo (around 300 BPM as compared to room-temperature water. However, exact temperatures were not specified, only general descriptions of “hot” and “cold” that might perhaps have varied for each participant.

Moreover, since the pre-study only involved imagined temperatures, it is possible that the association uncovered there only applies between sound attributes and imagined, not actually experienced, temperature. On the other hand, there is evidence that overlapping brain areas are activated by mental imagery and perception (see Kosslyn et al., 2001, for a review). For instance, patterns of eye movements are similar when inspecting visual mental images and real images (Brandt & Stark, 1997), and the combination of visual and olfactory mental imagery of food can enhance salivatory response (Krishna et al., 2014). In terms of temperature, imagined warmth/coldness from an egocentric perspective (same as in Experiment 15A) has been demonstrated to impact social evaluation in a similar way as real sensory experiences (Macrae et al., 2013).

In the main experiment, I set out to replicate the findings from the pre-study using actual water samples at different temperatures, in order to verify that the results of the pre-study, in fact, apply to orally experienced temperature.

A.3 Experiment 15B: Main experiment

A.3.1 Methods

A.3.1.1 Participants

24 participants (11 women, 13 men) between 19-43 years of age ($M=28.2$, $SD=5.7$) took part in the study. The participants gave their informed consent, and reported no hearing impairments. The participants were recruited from the Oxford University Research Participant Scheme. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

A.3.1.2 Auditory stimuli

The same auditory stimuli grid was used as in the pre-study (Experiment 15A).

A.3.1.3 Temperature stimuli

Distilled water samples were prepared at three different temperatures: cold (5°C), room-temperature (21°C), and hot (45°C). The cold water sample was kept in the refrigerator until needed for the relevant trials. The room-temperature water sample was kept at room-temperature for at least 24 hours in order for the temperature to stabilise. The hot water sample was prepared immediately before it was needed for the relevant trials and produced by mixing boiling and room-temperature water. The temperature of the samples was checked via an infrared thermometer (Benetech GM550 non-contact infrared digital thermometer) before each sample was served. The samples were served in 20 mL portions in 150 mL clear plastic cups.

A.3.1.4 Design and procedure

Participants performed the study seated in a study booth in front of a computer screen. Each participant completed 6 trials, with one of two possible arrangements of pitch and tempo on the audio grid (either fastest tempo on the left and lowest pitch on top, or slowest tempo on the left and highest pitch on top). The procedure was identical to that of Study 1. The only difference being that, for each trial, instead of asking participants to imagine drinking water at different temperatures, they were given the instruction to wait for the experimenter to provide them with a sample of water to taste. The samples were either cold, room-temperature, or hot, but the temperature was not explicitly communicated to the participant. Each sample was prepared by the experimenter immediately before it was consumed by the participant, in order to ensure it was served at the intended temperature. Participants took a 60 second break

between each trial in order to avoid gustatory adaptation effects (Green & Nachtigal, 2012).

The study lasted for approximately 10 minutes and the participants were paid £2.00 for their participation.

A.3.2 Results and discussion

The mean values of pitch and tempo for each temperature are shown in Figure A.5. The RM-MANOVA revealed a significant main effect of temperature ($F(4,20)=8.48$, $p<.0005$, Wilk's Lambda=.37), but not of instrument type ($F(2,22)=1.72$, $p=.20$, Wilk's Lambda=.86). Follow-up univariate tests with Greenhouse-Geisser corrections revealed that temperature had a significant main effect for both pitch ($F(1.29, 29.67)=22.09$, $p<.0005$, $\eta^2=0.49$) and tempo ($F(1.37, 31.59)=8.24$, $p=.004$, $\eta^2 = 0.26$). The participants' choice of pitch was significantly higher for cold water than for room-temperature water, and higher for room-temperature water than for hot water (all comparisons $p<.01$). The participants choice of tempo was significantly higher for cold water than for room-temperature ($p=.025$) and hot water ($p=.015$).

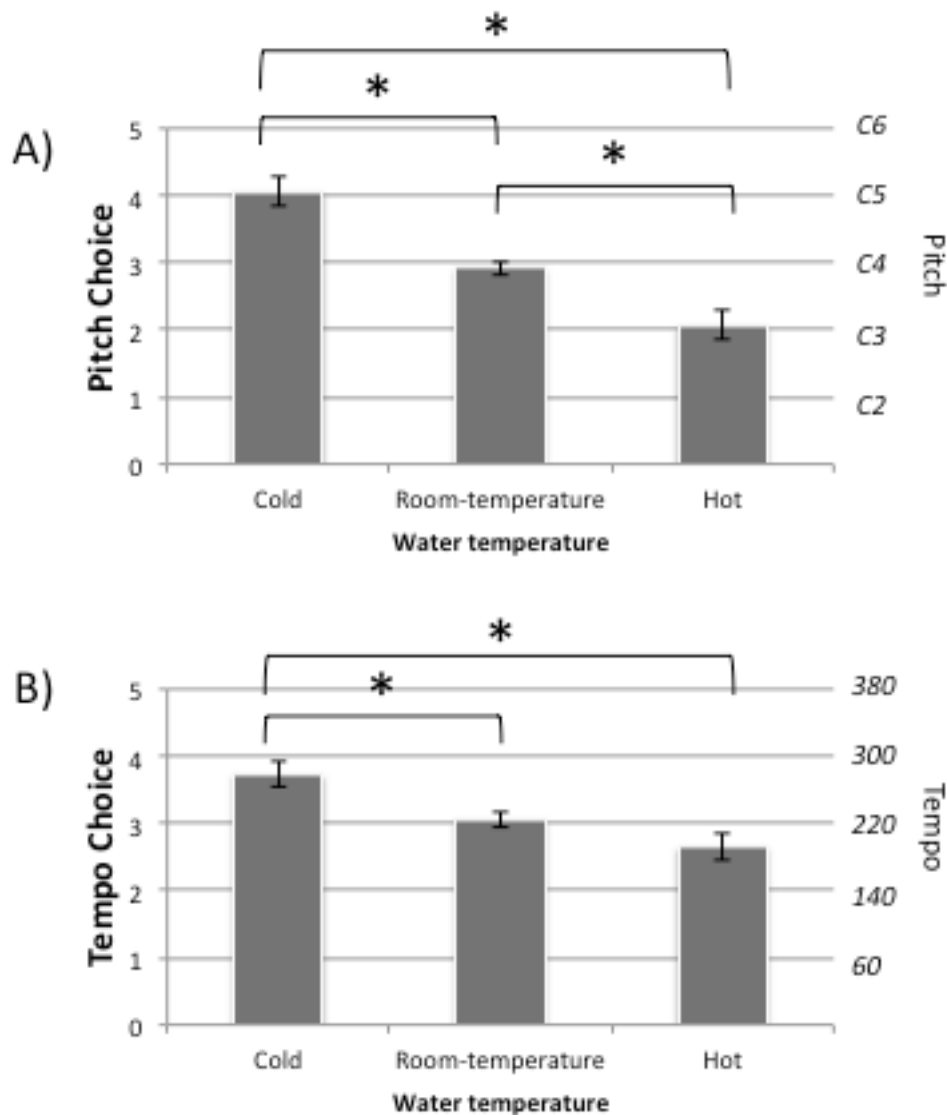


Figure A.5. Mean values of pitch (A) and tempo (B) for each temperature condition in Experiment 15B. Pitch ratings range from 1=C2 to 5=C6. Tempo ratings range from 1 - 80 BPM to 5 - 330 BPM. Error bars indicate standard error. * indicate statistical significance at $p < .05$.

As in the pre-study, there was a significantly positive correlation overall between participants' pitch and tempo choices, $r_{144} = 0.49$, $p < .0005$. Specifically, a positive correlation between pitch and tempo was observed when participants tasted warm water ($r_{48} = 0.53$, $p < .0005$), but not when they tasted cold water ($r_{48} = 0.26$, $p = .08$) or room-temperature water ($r_{48} = -0.16$, $p = .27$). This pattern of results can be seen in the colour-coded participant response frequency grid in Figure A.6, where the tendency

for responses to cluster around the pitch-tempo diagonal is much greater for the warm water than for the cold or room-temperature water conditions. Figure A.6 also demonstrates the consistency in the participants' choices, especially in the room-temperature and warm water scenarios, where just two adjacent squares account for roughly 50% of the total responses. In addition, Figure A.6 also reveals that for the cold water condition, high pitch seemed to be favoured over any specific tempo ranges.

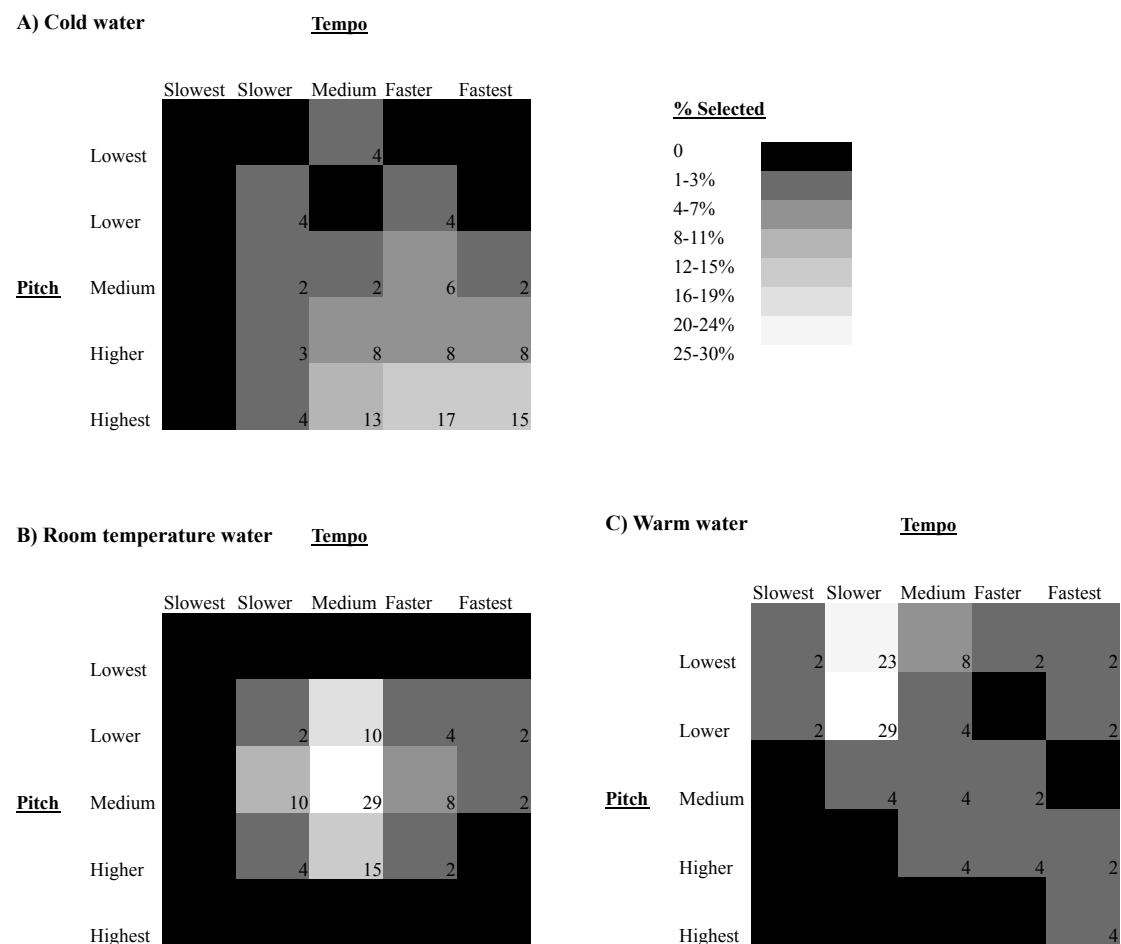


Figure A.6. Frequency information concerning how often participants selected each combination of tempo and pitch on the audio grid in Experiment 15B. Trials from both piano and string instrumentation sounds have been merged. The number in each box indicates % of total responses.

Finally, a comparison of the results from the pre-study and the main experiment was conducted via a RM-MANOVA with temperature and instrumentation as the within-participants factors, study number as a between-participants factor, and with pitch and tempo as response measures. There was no significant effect of study number ($F(2,51) < 1$, n.s.), and no significant interaction effect between study number and beverage temperature ($F(4,49) = 2.37$, $p = .07$, $\eta^2 = 0.16$), or between study number and instrumentation ($F(2,51) = 1.05$, $p = .36$, $\eta^2 = 0.04$).

The results of the main experiment confirmed the relationship between pitch/tempo and temperature shown in the pre-study. Replacing the imagined experience of drinking water samples at different temperatures with real water samples, the results revealed that cold water is associated with higher pitch and faster tempo compared to room-temperature and hot water. Moreover, the results of Experiment 15B revealed that when it came to actual water samples, the hot temperature sample was associated with a lower pitch than the room-temperature sample. This result was not observed in the pre-study, possibly because, when given the prompt to imagine drinking hot water, participants thought of water at a higher temperature than the rather comfortable 45°C that was offered in Experiment 15B. This inconsistency in imagined hot water temperature is shown in Figure A.4C, where there was a tendency for participants to choose either slow tempo and low pitched soundtracks, or fast tempo and high pitched soundtracks.

A.4 General discussion and conclusions

Why, one might ask, might people associate colder temperatures with higher pitch and faster tempo? One potential explanation points to emotional associations. There is evidence that the crossmodal correspondences between sound and smell (Levitan et

al., 2015; see Deroy et al., 2013, for a review) and between sound and taste (Wang et al., 2016) are partially mediated by emotion. Both fast tempo (Van der Zwaag et al., 2011) and high pitch (see Wang et al., 2016, Appendix B) are associated with increased arousal. The experience of drinking cold water might therefore be associated with fast tempo and high pitch because it is deemed arousing and refreshing (see Brunstrom et al., 1996; Sandick et al., 1983). Hot water, on the other hand, may be associated with soothing, calming warm beverages like tea instead. This was especially true for the main study since the hot water was served at 45°C, a comfortable drinking temperature. It would be interesting to ask participants in an online study to associate pitch and tempo with both extremely hot water (around boiling, at 100°C) in addition to a comfortable 45°C. One might expect the dangerously hot (hence arousing) water to be associated with faster tempo and higher pitch than the comfortably warm water. Of course, to truly verify the emotional association hypothesis, a future study would need to be conducted to collect participants' perceived emotion from each beverage sample.

The emotional association hypothesis could also be explained by brain connectivity. At a neuronal level, both oral temperature and audition are represented in the orbitofrontal cortex (Guest et al., 2007; Kadohisa et al., 2004; Verhagen & Engelen, 2006). In addition to unimodal neurons representing oral temperature, the majority of temperature-sensitive neurons are multisensory, associated with combinations of temperature, taste, and viscosity (Kadohisa et al., 2004). On the one hand, fMRI studies have revealed that the same brain regions – the prefrontal cortex and pregenual cingulate cortex – are responsible for processing the pleasantness of oral temperature as well as pleasantness of food flavour (Guest et al., 2007). On the other hand, the orbitofrontal cortex is also responsible for processing emotional responses

to auditory stimuli (Royet et al., 2000). Furthermore, there is evidence that the processing of aesthetic stimuli – whether they be paintings, music, or food – overlaps in the right anterior insula, an area associated with the processing of negative valence stimuli/concepts such as disgust and pain (Brown et al., 2011). Putting all of this together provides further evidence for the emotion-mediation theory underlying the correspondence between temperature and sound attributes.

The association of cold temperature with high pitch and fast tempo may also have a statistical origin, as Velasco et al. (2013a, b) demonstrated that enhancing the pouring sound around the 5-6 kHz range raised the perceived temperature of the liquid being poured. At the same time, people are familiar with the sound of ice cubes tinkling in the glass, whereas a hot drink gives rise to images of low-pitched gurgling bubbling pots. Finally, from an acoustics perspective, it is worth noting that a stringed instrument would sound flatter (i.e., lower pitch) in warmer temperature, as the string expands and loses tension (Tipler & Mosca, 2008). As these environmental observations seem to exist in nature, it would be helpful to conduct the same study with participants from different cultures in order to validate whether these mappings are indeed cross-cultural (see Knoeferle et al., 2015).

Here it is worth pointing out that what has been observed in Experiments 15A and 15B should not be thought of as synaesthesia *per se*. The results of the two experiments demonstrate a consistent general tendency for participants to associate certain temperatures with certain pitches and tempi without an accompanying sensory concurrent. While temperature-sound synaesthesia (where temperature induces a sound concurrent) does exist, it is extremely rare. According to one source, those with temperature-sound synaesthesia have been reported in 0.1% of the population with synaesthesia (Day 2005). Accounts of those synaesthetes usually mention childhood

associations of certain sounds with certain locations, such as the association of dragonflies with hot temperatures or video game music with cold (<http://syndiscovery.com/a-childhood-memory-sound-to-temperature-synaesthesia-2/>).

The fact that temperature has distinct pitch and tempo associations should certainly be of interest to those working in the food and beverage marketing industry. One could, for instance, imagine sonic branding or advertising jingles created to emphasize the cold, refreshing aspects of carbonated drinks, say, or the warming qualities of soup or tea. In addition, an interesting topic for future study would be to assess whether such sound-temperature associations might also apply to “warming” or “cooling” flavours such as cinnamon or menthol (Chartier 2012; Green 1992; Nagata et al., 2005).

In summary, oral temperature (e.g., of a drink) is a multisensory construct that does not only concern the tactile/oral-somatosensory senses, but also vision (Fenko et al., 2010; Ho et al., 2014; Wastiels et al., 2012), smell (Michael et al., 2010), and sound (Velasco et al., 2013). Experiments 15A and 15B demonstrate that beyond the sounds that are made by the beverage in question, more abstract musical parameters of pitch and tempo may also exhibit consistent associations with specific temperatures. One interesting future test would be to assess whether “cold” or “hot” soundtracks might alter the perceived temperature of a food/beverage (e.g., Sester et al., 2013) or even regulate participants’ body temperature (see Takakura et al., 2015, for an example of visual information affecting human thermodynamics).

A.5 References

- Brant, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, **9**, 29-38.
- Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., & Liotti, M. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *NeuroImage*, **58**, 250-258.
- Brunstrom, J. M., MacRae, A. W., & Roberts, B. (1997). Mouth-state dependent changes in the judged pleasantness of water at different temperatures. *Physiology & Behavior*, **61**, 667-669.
- Chartier, F. (2012). *Taste buds and molecules: The art and science of food, wine, and flavor* (translated by Levi Reiss). Hoboken, NJ: John Wiley and Sons.
- Day, S. (2005). Some demographic and socio-cultural aspects of synaesthesia. In L. C. Robertson, & N. Sagviv (Eds.), *Synaesthesia: Perspectives from cognitive neuroscience*. (pp. 11-13). Oxford, UK: Oxford University Press.
- Deroy, O., Crisinel, A.-S., & Spence, C. (2013). Crossmodal correspondences between odors and contingent features: Odors, musical notes, and geometrical shapes. *Psychonomic Bulletin & Review*, **20**, 878-896.
- Engelen, L., De Wijk, R. A., Prinz, J. F., van der Bilt, A., Janssen, A. M., & Bosman, F. (2002). The effect of oral temperature on the temperature perception of liquids and semisolids in the mouth. *European Journal of Oral Sciences*, **110**, 412-416.

- Fenko, A., Schifferstein, H. N. J., & Hekkert, P. (2010). Looking hot or feeling hot: What determines the product experience of warmth? *Materials & Design*, **31**, 1325-1331.
- Green, B. G. (1992). The sensory effects of I-menthol on human skin. *Somatosensory and Motor Research*, **9**, 235-244.
- Green, B. G., & Nachtigal, D. (2012). Somatosensory factors in taste perception: Effects of active tasting and solution temperature. *Physiology & Behavior*, **107**, 488-495.
- Guest, S., Grabenhorst, F., Essick, G., Chen, Y., Young, M., McGlone, F., ... Rolls, E. T. (2007). Human cortical representation of oral temperature. *Physiology & Behavior*, **92**, 975-984.
- Ho, H.-N., Van Doorn, G. H., Kawabe, T., Watanabe, J., & Spence, C. (2014). Colour-temperature correspondences: When reactions to thermal stimuli are influenced by colour. *PLoS ONE*, **9**:e91854.
- Kadohisa, M., Rolls, E. T., & Verhagen, J. V. (2004). Orbitofrontal cortex: Neuronal representation of oral temperature and capsaicin in addition to taste and texture. *Neuroscience*, **127**, 207-221.
- Knöferle, K. M., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic Bulletin & Review*, **19**, 992-1006.
- Knöferle, K. M., Woods, A., Käßler, F., & Spence, C. (2015). That sounds sweet: Using crossmodal correspondences to communicate gustatory attributes. *Psychology & Marketing*, **32**, 107-120.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, **2**, 635-642.

- Krishna, A., Morrin, M., & Sayin, E. (2014). Smellizing cookies and salivating: A focus on olfactory imagery. *Journal of Consumer Research*, **41**, 18-34.
- Levitan, C. A., Charney, S. A., Schloss, K. B., & Palmer, S. E. (2015). The smell of jazz: Crossmodal correspondences between music, odor, and emotion. *Cogsci 2015 Proceedings*, 1326-1331.
- MacRae, C. N., Raj, R. S., Best, S. B., Christian, B. M., & Miles, L. K. (2013). Imagined sensory experiences can shape person perception: It's a matter of visual perspective. *Journal of Experimental Social Psychology*, **49**, 595-598.
- Michael, G. A., Galich, H., Relland, S., & Prud'hon, S. (2010). Hot colors: The nature and specificity of color-induced nasal thermal sensations. *Behavioural Brain Research*, **207**, 418-428.
- Nagata, H., Dalton, P., Doolittle, N., & Breslin, P. A. S. (2005). Psychophysical isolation of the modality responsible for detecting multimodal stimuli: A chemosensory example. *Journal of Experimental Psychology: Human Perception & Performance*, **31**, 101-109.
- Royet, J. P., Zald, D., Versace, R., Costes, N., Lavenne, F., Koenig, O., & Gervais, R. (2000). Emotional responses to pleasant and unpleasant olfactory, visual, and auditory stimuli: A positron emission tomography study. *Journal of Neuroscience*, **20**, 7752-7759.
- Sandick, B. L., Engell, D. B., & Maller, O. (1984). Perception of drinking water temperature and effects for humans after exercise. *Physiology & Behavior*, **32**, 851-855.
- Sester, C., Deroy, O., Sutan, A., Galia, F., Desmarchelier, J.-F., Valentin, D., & Dacremont, C. (2013). "Having a drink in a bar": An immersive approach to

- explore the effects of context on beverage choice. *Food Quality and Preference*, **28**, 23-31.
- Spence, C., & Wang, Q. (J). (2015). Sensory expectations elicited by the sounds of opening the packaging and pouring a beverage. *Flavour*, **4**:35.
- Takakura, J., Nishimura, T., Choi, D., Egashira, Y., & Watanuki, S. (2015). Nonthermal sensory input and altered human thermoregulation: Effects of visual information depicting hot or cold environments. *International Journal of Biometeorology*, **59**, 1453-1460.
- Tipler, P. A., & Mosca, G. (2008). *Physics for scientists and engineers, Vol. 1* (6th Ed.) (pp. 666-670). New York, NY: Worth Publishers.
- Van der Zwaag, M. D., Westerink, J. H. D. M., & van den Broek, E. L. (2011). Emotional and psychophysiological responses to tempo, mode, and percussiveness. *Musicae Scientiae*, **15**, 250-269.
- Velasco, C., Jones, R., King, S., & Spence, C. (2013a). The sound of temperature: What information do pouring sounds convey concerning the temperature of a beverage. *Journal of Sensory Studies*, **28**, 335-345.
- Velasco, C., Jones, R., King, S., & Spence, C. (2013b). "Hot or cold?" On the informative value of auditory cues in the perception of the temperature of a beverage. In K. Bronner, R. Hirt, & C. Ringe (Eds.), *(((ABA))) Audio Branding Academy Yearbook 2012/2013* (pp. 177-187). Baden-Baden: Nomos.
- Verhagen, J. V., & Engelen, L. (2006). The neurocognitive bases of human multimodal food perception: Sensory integration. *Neuroscience and Biobehavioral Reviews*, **30**, 613-650.

- Wang, (Q.) J., Wang, S., & Spence, C. (2016). “Turn up the taste”: Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and pitch. *Chemical Senses*, **41**, 345-356.
- Wastiels, L., Schifferstein, H. N. J., Heylighen, S., & Wouters, I. (2012). Red or rough, what makes materials warmer? *Materials & Design*, **43**, 441-449.
- Watt, R., & Quinn, S. (2007). Some robust higher-level percepts for music. *Perception*, **36**, 1834-1848.