

RUNNING HEAD: SOUNDS SPICY

**Sounds spicy: Enhancing the evaluation of piquancy
by means of a customised crossmodally congruent soundtrack**

Qian (Janice) Wang¹, Steve Keller², & Charles Spence¹

¹ *Crossmodal Research Laboratory, Oxford University, Oxford, UK*

² *iV Audio Branding, Nashville, Tennessee, USA*

DATE: DECEMBER 2016

RESUBMITTED TO: *FOOD QUALITY AND PREFERENCE*

CORRESPONDENCE TO: Qian (Janice) Wang, Crossmodal Research Laboratory,
Department of Experimental Psychology, University of Oxford, South Parks Road,
Oxford, OX1 3UD, U.K. TEL: +44-1865-271307 FAX: +44-1865-310447; E-MAIL:
qian.wang@psy.ox.ac.uk

ABSTRACT

The aim of the present research was to uncover the auditory parameters that correspond to the experience of spiciness/piquancy in food, and to assess whether such crossmodal correspondences have perceptual consequences when it comes to evaluating the spiciness of actual foods tested in a naturalistic environment. An online study (Experiment 1) was conducted first in order to determine the acoustical/musical parameters that best match spiciness. The results were used to compose a spicy soundscape that was incorporated into the subsequent experiments. Next, a between-participants study (Experiment 2) was conducted to test the effect of different background sound conditions on participants' expected and actual ratings of a novel restaurant dish. Four sound conditions were used in testing: the aforementioned spicy soundtrack, a sweet soundtrack, white noise, and silence. The expected spiciness of the dish was significantly higher in the spicy soundscape group as compared to the other groups. However, no significant differences were observed in the actual taste ratings. A contributing factor to this later result may have been the large disparity between the participants' expectations of spiciness and the actual (mild) spiciness of the test dish itself. To follow-up, a study (Experiment 3) was conducted with a spicier food sample and the same sound conditions. Here, the results revealed that the rated spiciness of the food sample was significantly higher in the spicy soundtrack condition than in the other sound conditions. Finally, a study using both mild and hot salsa (Experiment 4) demonstrated an interaction effect between the sound condition and stimuli spiciness level, consistent with the assimilation-contrast model of consumer expectation disconfirmation. These results therefore demonstrate that a soundscape with auditory attributes corresponding to spiciness can enhance the perception of spiciness in foods, likely via a mechanism of inducing sensory expectations.

KEYWORDS: CROSSMODAL CORRESPONDENCES; SPICINESS; EXPECTATION DISCONFIRMATION; ASSIMILATION; CONTRAST;

Introduction

Crossmodal correspondences consist of associations that the majority of people tend to share between seemingly-unrelated attributes (or dimensions) of experience in different sensory modalities (see Spence, 2011, for a review). The majority of the research that has been conducted over the last century has involved audio-visual correspondences, but there is an emerging interest in those crossmodal correspondences that are related to food/flavour (see Spence et al., 2010, for flavour and colour correspondences; Deroy et al., 2013, for odour-sound correspondences, and Slocombe et al., 2015, for touch-taste correspondences). Specifically, a growing list of correspondences have now been demonstrated between sonic properties and different flavor 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, 2010; 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). The focus of the present work concerns 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 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. People have also been shown to choose spicier sauces in rooms that are more brightly-lit, as compared to in rooms that are dimly-lit (Xu & Labroo, 2014). The first aim of the present research, then, was to uncover correspondences between basic auditory attributes and spiciness/piquancy.

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).

According to Shermer and Levitan (2014), the red colouring used in their study to modify the salsa samples generated sensory expectations of spiciness, which in turn enhanced the perceived spiciness of the salsas. Our hypothesis, then, is that auditory stimuli corresponding to spiciness might also enhance spiciness 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 adjust their sensory perceptions to try and conform to their initial taste expectations (see Lelièvre et al., 2009). When it comes to sounds and basic tastes, the research that has been published to date 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., 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. Therefore, the second aim of the research reported here is to verify whether spicy-congruent soundtracks might enhance spiciness ratings by acting on participants' sensory expectations.

First, we adapted the method from Knoeferle et al. (2015) for constructing music that corresponds to one of the basic tastes. Namely, we conducted a study in which the participants had to answer a series of questions where they would choose from multiple musical segments that differed in terms of only one auditory parameter. In Experiment 1, an online study was conducted in order to explore the extent to which individuals would consistently associate different musical parameters with spiciness. A spicy soundscape was then composed, based on the findings from Experiment 1, and was used in Experiments 2 - 4 in order to determine whether sound can influence participants' expected and actual taste ratings of real food.

Experiment 1

Methods

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).

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). We also included three questions 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>.

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 musical attribute tested. For each question, the participants had to choose the sound clip that best matched spicy foods (see Figure 1 for question format). In the test itself, it was further clarified that spicy meant piquant foods like chilli 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.

Please select the sound clip that matches best with spicy foods - as in food that is "hot", like in chili pepper or hot sauce.

☐ Clip 432

  432 

☐ Clip 827

  827 

☐ Clip 916




  916 

Figure 1. Sample question from Study 1.

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.

Data Analysis

A chi-squared goodness of fit test was conducted to assess whether the distribution of choices for each question were random. We also analysed the results based on musical expertise and spiciness liking.

Results

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 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 with 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.

				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 1. Results of participants' soundtrack-spiciness matches, 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 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 2. Results of participants' soundtrack-spiciness matches, 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 no auditory features where participants had different choices based on spicy liking (alpha level = .0038, corrected for multiple tests).

Discussion

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 1 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 (see Piqueras-Fiszman & Spence, 2015, for a review of the impact of expectations on sensory perception). In Experiment 2, we set out to test this hypothesis empirically.

Experiment 2

Methods

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).

Auditory stimuli

A spicy soundtrack was composed by iV Audio Branding based on the attributes found to be significantly associated with spiciness from Experiment 1. A sweet soundtrack was also composed by iV, with high pitch, legato articulation, and consonant harmony (Crisinel & Spence, 2010; 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 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).

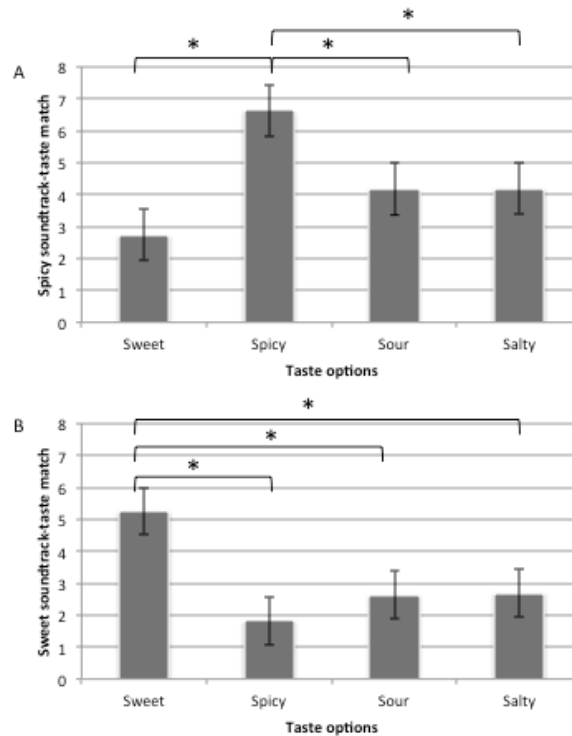


Figure 2. Results of participants' ratings of sound-taste matching (on a scale of 0-10) for A) the spicy soundtrack and B) the sweet soundtrack composed by iV Audio Branding. Error bars indicate the standard error of the means. * indicate statistical significant comparisons ($p < .0005$ in this case).

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

Food 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 3). The dish consisted of a piece of ancho chilli dusted butternut squash tempura with brie, pear, mushrooms and salad leaves, topped with a sour-spicy dressing, spicy pepitas, and spiced chocolate sauce.



Figure 3. The dish served at Etch restaurant during Experiment 2. The dish itself is a novel creation that is likely to be unfamiliar to diners.

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.

Results

The average expected ratings for the dish are shown in Figure 4. A multivariate analysis of variance (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 ($F(3,175)=0.33$, $p=0.81$), flavour intensity ($F(3,175)=1.28$, $p=0.28$), or liking ($F(3,175)=0.11$, $p=0.96$). 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 4. Participants' mean ratings of the expected taste of the dish in Experiment 2, in all four sound conditions. 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. 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 ($F(12, 522)=0.93$, $p=.51$, Pillai's Trace=0.06).

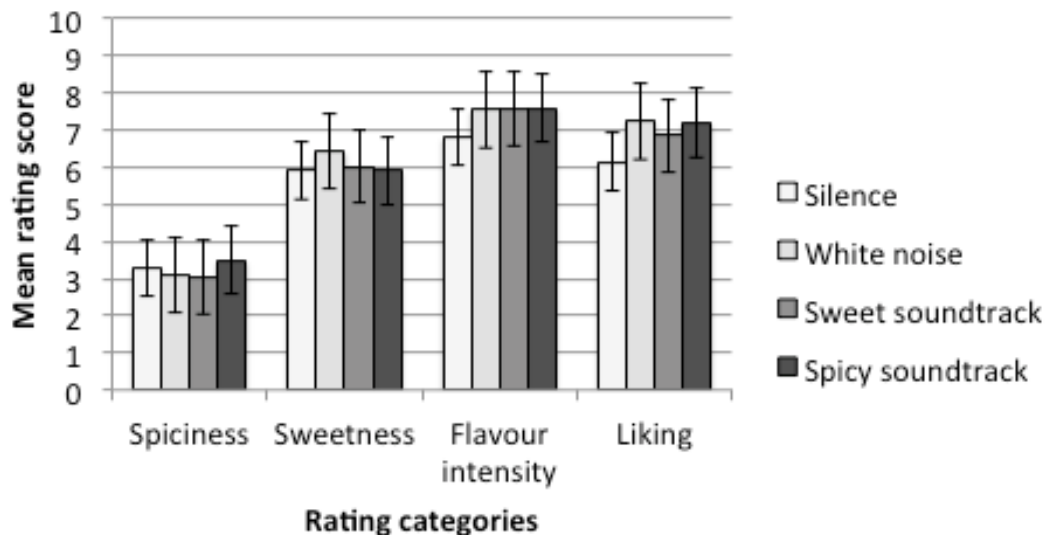


Figure 5. Results of participants' ratings of the actual taste of the dish, in all four sound conditions. Error bars represent the standard error of the means.

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.

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 food-related ratings.

Discussion

The results of Experiment 2 revealed that the spicy soundtrack significantly increased the expected spiciness of the dish 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 2 was simply not spicy enough. Intriguingly, in Shermer and Levitan's (2014) recent 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 are 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 to a control condition). In Experiment 2, 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 the level that 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 to 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

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.

and closer to the expected level of heat (on average 6 out of 10), we would have observed a difference between the spicy soundtrack and other sound conditions. To validate whether this model can explain our results, we conducted Experiment 3 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 we should expect to see positive correlations between music pleasantness and dish pleasantness ratings. We did not, however, observe a significant correlation in this study between music pleasantness and expected ($r_{137}=0.07$, $p=0.41$) or actual ($r_{137}=0.11$, $p=0.19$) dish liking. Therefore, our results suggest that some mechanism other than sensation transference is at work in terms of the soundtracks' influence on participants' ratings.

Experiment 3

In Experiment 2, while changes were observed to the expected spiciness in the spicy soundtrack condition, no changes to the actual taste ratings were found. To explore whether this was due to the taste stimuli being too mild, Experiment 3 was conducted to evaluate whether foods with piquancy levels closer to what is expected can be influenced by the spicy soundtrack.

Methods

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).

Auditory stimuli

The same sound stimuli were used as in Experiment 2 (spicy soundtrack, sweet soundtrack, white noise, silence). 60-second excerpts were made by taking the first 60 seconds of audio from the sound stimuli.

Food 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 the present study. 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).

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, for the non-silent conditions.) In order to make sure that 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.

Results

The average ratings for the salsa under the different sound conditions are shown in Figure 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 the spiciness rating was 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 ($M=5.72$, $SD=1.57$, $p=.021$).

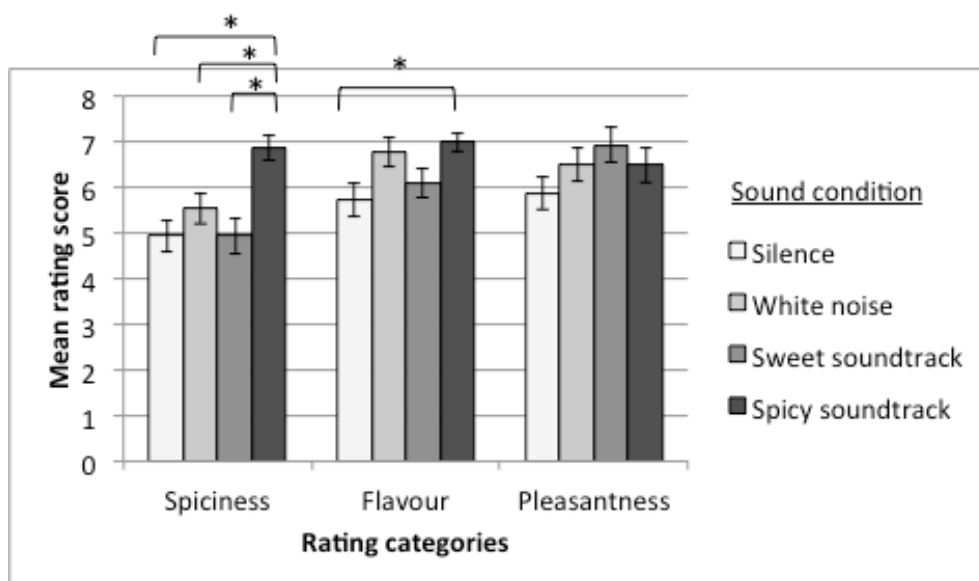


Figure 6. Results of participants' ratings of the taste of salsa in Experiment 3, in all four sound conditions. Error bars represent the standard error of the means. Asterisks denote statistical significance (* $p<.05$).

Discussion

The results of Experiment 3 revealed that the spiciness rating of the salsa samples was significantly higher while listening 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 2 and 3 is 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 2). 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, we did not see a similar enhancement effect 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 3 is that we did not use a balanced Latin Square design, and that, with 18 participants and 4 soundtrack sequences, we did not have an equal number of participants for each sequence. Given the relatively small number of participants ($N=18$) and the uneven balancing, it is possible that our results could have been subject to carryover effects of capsaicin.

Experiment 4

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

Methods

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).

Auditory stimuli

Since Experiment 3 showed that only the spicy soundtrack enhanced perceived spiciness compared to all other sound conditions, we used the spicy soundtrack and silence as a control condition in this study. The 60 second spicy soundtrack was the same as in Experiment 3. 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.

Food 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.

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 number) 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 in 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.

Results

The average ratings for the salsa with different heat levels and under different sound conditions are shown in Figure 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 covariates⁴ 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 intensity ($F(1,36)=.91$, $p=.48$, $\eta^2=.014$). 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 to be 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 ($p=.018$)

⁴ We applied the Delaney-Maxwell method of mean-centering the covariate values to prevent the covariates from altering the main effects of the repeated measures factors (Delaney & Maxwell, 1981).

than the Doritos Mild Salsa (all p-values in post hoc comparison tests have been Bonferroni corrected).

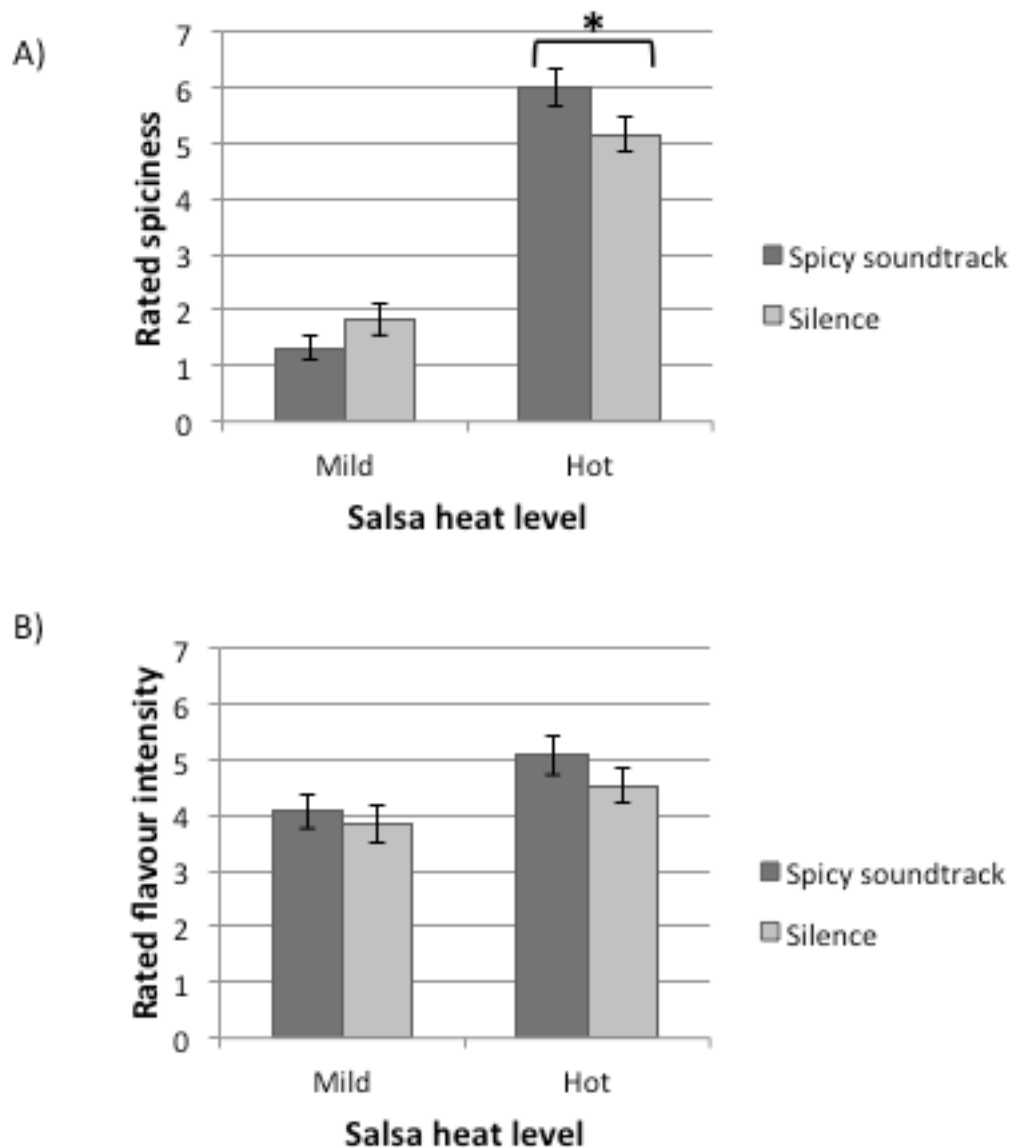


Figure 7. Participants' ratings of the spiciness (A) and flavour intensity (B) of salsa in Experiment 4, for two salsa heat levels and two sound conditions (spicy soundtrack and silence). Error bars represent the standard error of the means. Asterisks denote statistical significance (* $p < .05$).

Discussion

The results of Experiment 4 revealed that the ratings of the spiciness of the salsa samples was 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 flavour, and assimilation occurs if the actual experienced flavour 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 flavour, however, contrast can occur and the resulting percept is in fact biased in the opposite direction of 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. We did not collect spiciness sensitivity information for the current study, although it is plausible that anyone who would volunteer for 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 spice sensitivity and assess the effect of the spicy soundtrack on mildly spicy foods.

An alternative interpretation for the effect observed for the mild salsa is that the salsa 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 (around the same loudness as a vacuum cleaner), might have distracted the participant from paying attention to the mild spiciness of the salsa.

General Discussion

Overall, the research reported here demonstrates, for the first time, the existence of crossmodal correspondences between spiciness/piquancy and sound attributes – high pitch, fast tempo, and high levels of distortion. These results also provide evidence that such correspondences can be used to modify people’s evaluation of the expected and actual spiciness of foods.

One question that remains to be answered is the relative contribution of perceptual and cognitive effects of such crossmodally congruent soundtracks. Experiment 2 demonstrated that participants’ expectations were influenced by the spicy soundtrack, which shows that the influence of music begins before participants even taste the food. Therefore, the soundtrack could give rise to sensory expectations which might then modify perception, perhaps by drawing people’s attention towards specific tastes experienced in the eating process (Deliza & MacFie, 1996; Piqueras-Fiszman & Spence, 2015). The results of Experiments 2-4 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 reality is small (Shermer & Levitan, 2014; Woods et al., 2011). Furthermore, the results of Experiment 4 go against an attentional account of the findings – for if the soundtracks focused participants’ attention on spiciness, we would have seen an enhancement in actual spiciness ratings regardless of the spiciness level.

A potential issue with the current study is that Experiment 2 was conducted using a between-participants design while Experiments 3 and 4 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 4 would contradict this theory). To fully account for this, further experiments should perhaps be conducted using more sophisticated designs that can isolate response bias from true perceptual effects (see Marks et al., 2003; Odgaard et al., 2003; Zellner & Kautz, 1990). Alternatively, however, 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). This might explain the difference in spiciness

ratings in Experiments 3 and 4, where participants experienced multiple sound conditions.

Another aspect for future exploration is the potential cultural associations that people might make between sound and spiciness. After the experiments, some participants commented that the soundtracks reminded them of Latin American music, which they strongly associate with spicy food. In order to separate the effect of culture from sound-spiciness correspondences, it would be interesting to conduct Experiment 1 with a non-Western population who might not have such preconceived cultural associations.

A deeper understanding of the underlying processes is key to understanding how soundtracks can be used in practical settings by food manufacturers and chefs, as well as for those in the healthcare industry looking to enhance flavour perception for those individuals with a palate that is fading (e.g., due to age-related decline) or other health issues.

ACKNOWLEDGEMENTS

CS would like to thank the AHRC grant entitled 'Rethinking the senses' (AH/L007053/1) for supporting this research. We would like to thank chef Deb Paquette for her help in designing the food stimuli for Experiment 2, and the staff at Etch restaurant for helping with logistics. We would also like to thank the entire team at iV Audio Branding for designing the auditory stimuli for all three experiments.

REFERENCES

- Anderson, R. E. (1973). Consumer dissatisfaction: The effect of disconfirmed expectancy on perceived product performance. *Journal of Marketing Research*, **10**, 38-44.
- Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and Cognition*, **17**, 1016-1031.
- Bartoshuk, L. M., Rifkin, B., Marks, L. E., & Bars, P. (1986). Taste and aging. *Journal of Gerontology*, **41**, 51-57.
- 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.
- 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.
- Cheskin, L. (1972). *Marketing success: How to achieve it*. Boston, MA: Cahnern Books.
- 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., & Spence, C. (2010). As bitter as a trombone: Synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Attention, Perception, & Psychophysics*, **72**, 1994-2002.
- 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.

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.

Gallace, A., & Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size. *Perception & Psychophysics*, **68**, 1191-1203.

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.

Knoeferle, K. M., Woods, A., K  ppler, F., & Spence, C. (2015). That sounds sweet: Using crossmodal correspondences to communicate gustatory attributes. *Psychology & Marketing*, **32**, 107-120.

Kn  ferle, K. M., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic Bulletin & Review*, **19**, 992-1006.

Leli  vre, 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.

Marks, L. E., Ben-Artzi, E., & Lakatos, S. (2003). Cross-modal interactions in auditory and visual discrimination. *International Journal of Psychophysiology*, **50**, 125-145.

Mesz, B., Trevisan, M. A., & Sigman, M. (2011). The taste of music. *Perception*, **40**, 209-219.

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. and 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.

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.

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.

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.

Seo, H.-K., 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.

Shermer, D. Z., & Levitan, C. A. (2014). Red hot: The crossmodal effect of color intensity on perceived piquancy. *Multisensory Research*, **27**, 207-223.

Slocombe, B. G., Carmichael, D. A., & Simner, J. (2015). Cross-modal tactile-taste interactions in food evaluations. *Neuropsychologia*, **88**, 58-64.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, **73**, 971-995.

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.

Stevenson, R. J. (2009). *The psychology of flavour*. Oxford, UK: Oxford University Press.

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.

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.

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.

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. *Chemical Senses*, **22**, 365-369.

Woods, A. T., Poliakoff, E., Lloyd, D. M., Bijksterhuis, G. B., & Thomas, A. (2010). Flavour expectation: the effect of assuming homogeneity on drink perception. *Chemical perception*, **3**, 174-181.

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.

Zellner, D. A., & Kautz, M. A. (1990). Color affects perceived odor intensity. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 391-397.