Metacognition and Crossmodal Correspondences Between Auditory Attributes and Saltiness in a Large Sample Study

Qian Janice Wang*¹, Steve Keller², and Charles Spence³

¹Department of Food Science, Aarhus University, Aarhus, Denmark

²Studio Resonate at SXM media, Oakland, USA

*³Crossmodal Research Laboratory, Department of
Experimental Psychology, Oxford University, Oxford, UK*

DATE: MAY 2021

CORRESPONDENCE TO: Dr. Qian Janice Wang, Department of Food Science, Aarhus University, Agro Food Park 48, Aarhus, 8200, Denmark.

E-mail: qianjanice.wang@food.au.dk

Abstract

Mounting evidence demonstrates that people make surprisingly consistent associations between auditory attributes and a number of the commonly-agreed basic tastes. However, the sonic representation of (association with) saltiness has remained rather elusive. In the present study, a crowd-sourced online study ($N = 1,819$ participants) was conducted to determine the acoustical/musical attributes that best match saltiness, as well as people's confidence levels in their choices. Based on previous literature on crossmodal correspondences involving saltiness, thirteen attributes were selected to cover a variety of temporal, tactile, and emotional associations. The results revealed that saltiness was associated most strongly with a long decay time, high auditory roughness, and a regular rhythm. In terms of emotional associations, saltiness was matched with negative valence, high arousal, and minor mode. Moreover, significantly higher average confidence ratings were observed for those saltiness-matching choices for which there was a majority agreement, suggesting that individuals were more confident about their own judgments when it matched with the group response, therefore providing support for the so-called 'consensuality principle'. Taken together, these results help to uncover the complex interplay of mechanisms behind seemingly surprising crossmodal correspondences between sound properties and taste.

Keywords: crossmodal correspondences; metacognition; consensuality principle; saltiness

1. Introduction

While it may seem surprising, there are robust connections between what we hear (musically) and what we taste. Previous studies have demonstrated that certain basic sonic attributes (such as pitch or articulation) are generally associated with specific basic tastes, no matter whether they are presented as taste words (e.g., Knoferle et al., 2015) or else as tastants (e.g., Wang et al., 2016). Moreover, soundtracks embodying these correspondences have been shown to demonstrate the phenomenon of *sonic seasoning*, whereby the sensory aspects of what we taste/eat can be altered predictably by listening to such soundtracks (see Spence et al., 2019, for a recent review of the emerging literature on this topic). Pairing foods with soundtracks corresponding to sweetness has already shown itself to be a promising avenue for sugar reduction (see Wang et al., 2019, for a review). Similar efforts can be considered for salt reduction (Thomas-Danguin et al., 2019). However, designing salty music has proven difficult, as previous research into the associations between the gustatory attributes of salt and the sonic attributes of music has yielded somewhat indistinct responses. For instance, while research suggests that sweetness is regularly associated with higher pitch and bitterness with lower pitch, saltiness falls more on a continuum, with some studies suggesting pitches closer to middle C (i.e., C4) and others suggesting lower pitches (see **Table 1**). It is possible, of course, that researchers have simply yet to uncover the auditory dimension with the clearest (or most robust) associations with saltiness. To this end, our aim in the present study was to uncover more nuanced auditory correspondences with saltiness. In so doing, we hope to pave the way for the creation of more effective salty soundtracks, and, from a theoretical perspective, discover what the variety of correspondences can tell us about the nature of crossmodal correspondences.

1.1 Overview of salt-based correspondences

Crossmodal correspondences involving saltiness have been studied over the years (see Table 1). The majority of the research has tended to focus on visual associations, with only a few attempts to assess any auditory and oral-somatosensory textural associations. It is worth noting that we are limiting this overview of crossmodal correspondences with salty taste to the non-chemical senses, since there are many aroma-saltiness associations thought to be driven by prior exposure and associative learning (e.g., Batenburg & van der Velden, 2011). In contrast, sound-taste

correspondences are typically not driven by direct prior co-exposure, but rely on mediation via a third factor, such as emotional valence or arousal; for instance, both sweet tastes and sweet-corresponding soundtracks tend to have high valence (i.e., they are pleasant), while both sour tastes and sour-corresponding soundtracks have high arousal (i.e., they are exciting) (Wang et al., 2016).

Crossmodal correspondences with saltiness have been identified in visual, tactile, and auditory modalities. In terms of visual attributes, the majority of salty correspondences discovered thus far have been related to colour, specifically with either white or blue, presumably due to people's familiarity with the appearance properties of salt crystals, or its association with conventional descriptions of the sea as being blue (see Spence, Wan, Woods, Velasco, Deng, Youssef, & Deroy, 2015, for a review). In terms of tactile attributes, van Rompay and Groothedde (2019) have reported that saltiness is associated with textural irregularity and roughness, which seems to mimic the texture of salt crystals themselves. Moreover, the visual associations of *angularity* can also be interpreted as a tactile feature, since angularity can be both seen and felt. In terms of auditory qualities, saltiness tends to be associated with middle pitch, with staccato articulation being the only distinctive auditory characteristic. This contrasts with the other widely tested basic tastes (sweet, bitter, and sour), where there are clearer auditory correspondences with values at polar extremes, at least in terms of pitch and volume (see Knöferle & Spence, 2012, for a review). We therefore aimed to look towards existing correspondences in order to find inspiration for auditory attributes that might match with saltiness.

Table 1. A representative selection of studies demonstrating crossmodal interactions/correspondences between saltiness and other non-chemical sensory attributes (i.e., not olfactory associations).

Authors	N (participants)	Modality	Attribute(s)	Evocation of salty stimulus	Crossmodal association (CA) / Crossmodal influence (CI)
Crisinel & Spence, 2010	34	Audition	Pitch (average: C4) Instrument type (brass)	Word	CA
Simner et al., 2010	65	Audition	Vowel height (lower)	Aqueous solution	CA
Mesz et al., 2011	9 (Experiment 1) 57 (Experiment 2)	Audition	Pitch (low) Articulation (staccato)	Word	CI (Experiment 1) music improvisation

			Loudness (average) Chord consonance (average) Melody consonance (average)		changed based on taste word theme CA (Experiment 2)
Wang et al., 2015	100 (Experiment 1) 50 (Experiment 2)	Audition	5 salty soundtracks (see Table 1 in paper)	Word	CA
Wang et al., 2016	33 (Experiment 1) 33 (Experiment 2)	Audition	Pitch (average, C4)	Aqueous solution	CA
Guetta & Loui, 2017	160 (Experiment 1) 50 (Experiment 3)	Audition	Articulation (staccato) Tempo (fast)	Word (Experiment 1) Flavoured chocolate (Experiment 3)	CA
Padulo et al., 2018.	89	Audition	5 salty soundtracks from Wang et al. 2015	Food images	CA
van Rompay & Groothedde, 2019	180 (90 per condition)	Touch	Surface texture (irregularity, roughness)	Potato chips	CI, chips from rough cup rated 20% saltier than smooth cup
O'Mahony, 1983	51	Vision	Colour (white)	Word	CA
Heller, 1999	2000	Vision	Colour (white, grey, blue)	Word	CA
Koch & Koch, 2003	45	Vision	Colour (white)	Word	CA
Tomasik-Krótki & Dtrojny, 2009	519	Vision	Colour (blue)	Word	CA
Harrar et al., 2011	60	Vision	Colour (blue)	Popcorn	CI, sweet popcorn from blue bowl rated saltier than in white bowl
Wan et al., 2014	452	Vision	Colour (white)	Word	CA
Woods & Spence, 2016	201	Vision	Colour (white)	Word	CA
Woods et al., 2016	100 (Experiment 1) 100 (Experiment 2)	Vision	Colour (white, white with blue background)	Word	CA
Saluja & Stevenson, 2018	50	Vision	Colour (blue)	Aqueous solution	CA

Wongthahan et al., 2020	25 (20 consumers, 5 chefs)	Vision	Colour (darker brown compared to lighter brown)	Soy sauce	CI, soy sauce colour intensity is correlated with perceived saltiness
Fateminia et al., 2020	112 children (2-6 years)	Vision	Colour (brown)	Word	CA
Velasco et al., 2015a	26 (Experiment 1) 100 (Experiment 2) 65 (Experiment 3) 20 (Experiment 4)	Vision	Shape (angularity)	Word (Experiments 1-3) Aqueous solution (Experiment 4)	CA
Velasco et al., 2015b	101 (Experiment 1)	Vision	Typeface (angularity)	Word	CA
Velasco et al., 2016	30	Vision	Shape (angularity)	Aqueous solution	CA

1.2 Accounting for crossmodal correspondences with salty tastes

The search for underlying mechanisms behind crossmodal correspondences has been an active field of research in recent years. In general, four plausible explanations have been offered to account for crossmodal correspondences: (1) structural or intensity matching, (2) hedonic matching (or more generally, emotional mediation), (3) semantic matching, and (4) statistical co-occurrences (Knöferle & Spence, 2012; Spence, 2019). As pointed out in Section 1.1, most of the established crossmodal correspondences with saltiness can potentially be linked to the sensorial properties of salt crystals themselves, such as its colour or shape (see Table 1), which would point to an associative learning explanation in terms of the statistical co-occurrence of the attributes (explanation 4). However, when it comes to sound-taste correspondences, the challenge is that tastants themselves do not tend to make sounds (compared to say, sound-temperature associations, where one can learn to associate sizzling sounds with hot foods). Some researchers have ventured to suggest that associative learning might be at play for sound-taste correspondences via the stereotypical orofacial gestures we instinctively make when encountering different tastants. For example, the bitterness-low pitch correspondence may be explained by the fact that, as newborns, we tend to stick out our tongues outwards and downwards when tasting bitterness, resulting in a lower-frequency sound when exhaling (Knöferle & Spence, 2012; Spence, 2012). However, this

theory does not work so well when applied to saltiness, because it has been reported that infants do not respond to salty taste in the first few months (Desor et al., 1975).

Therefore, when considering sound-taste correspondences, one often looks to mediated accounts, such as emotion mediation (explanation 3). For tastes with strong emotional associations, such as sweetness, bitterness, or sourness, the same emotions seem to mediate taste-sound correspondences; for instance, the music that is normally chosen to match with sweetness tends to be pleasant, just like the taste itself (Wang et al., 2016). With saltiness, however, it is less clear how emotions might influence any salt-related crossmodal correspondences, since, according to previous research (Wang et al., 2016, Figure 6), saltiness does not appear to have such strong emotional associations. That said, there may be some more indirect semantic associations between saltiness and sadness, since 1) tears are salty, and 2) the colour blue is associated with both saltiness and sadness (Adams & Osgood, 1973). Note also that our preference for salt (and minerals) depends on need state/balance in a way that is not true for other tastes, which raises the question of whether crossmodal correspondences related to salt might also be specifically state-dependent. In other words, salty foods are preferred when our bodies need salt (e.g. after exercising), but are not preferred right after meals (Finlayson et al., 2007). This implies that saltiness might be associated with pleasant sounds when one is in need of salt, but associated with unpleasant sounds otherwise.

Another consideration is how the strength of crossmodal correspondences, both in terms of how confident an individual might feel about the correspondence and in terms of the consensus of the mapping across individuals, might be driven by different mechanisms. For instance, are emotional associations less stable or consistent than statistical ones (Spence, 2020a)? The strength of correspondences can be measured in several ways (Rader & Tellegen, 1987): test-retest reliability, the degree of general agreement across participants (consensus), or self-reported confidence (metacognition) in the matching task. According to the consensuality principle, subjective confidence in one's answers is typically correlated with group consensus (Koriat 2008, 2011). The use of metacognitive measures in crossmodal correspondences is a relatively new development, as Chen and his colleagues recently demonstrated how the combination of consensus and confidence measures can be used to quantify the strength of associations in the case of sound-shape correspondences (Chen et al., 2019). It should be stressed here that consensuality principle in the

cases of crossmodal correspondences is especially interesting, as there is no correct response per se when it comes to correspondences, which makes consensuality more surprising than in instances where there is a correct answer.

To this end, we decided to conduct an online study in order to explore the extent to which individuals would consistently associate different acoustic and musical attributes with saltiness. We adapted the method from Knoeferle et al. (2015) and Wang et al. (2017), whereby participants answered a series of questions where, for each question, they chose from sound clips differing only in terms of a single auditory attribute. The participants were also asked to give their confidence levels for each response.

2. Methods

2.1 Participants

A convenience sample of 1819 participants (948 women, 858 men, 13 unspecified) aged between 16 to 88 years of age (Mean = 50.39, SD = 15.86) took part in an online study. No participant reported hearing impairments. The crowdsourced participants were recruited from listeners of a major streaming music company via newsletter and social media outlets in April 2019. While the recruitment was achieved by means of convenience sampling, a power analysis was performed to verify the sample size. Considering $\alpha = 0.05$, and power of 0.9, 1266 participants would have been required in order to detect a small effect size ($w = 0.1$) for a Chi squared goodness of fit test with three multiple choice options (see Table 2). The study was approved by the Central University Research Ethics Committee of Oxford University (R61059) and were performed in accordance with relevant guidelines and regulations to the Declaration of Helsinki. A participant information and consent form was built into the first page of the online study, and all participants gave their informed consent before proceeding to the study itself.

2.2 Auditory stimuli

A composer was commissioned to create a series of short sound clips, designed to reflect different variations within a set of 13 auditory attributes (all clips can be heard at <https://soundcloud.com/janicewang09/sets/saltyccstudy2020>). Eleven of these targeted specific musical attributes: articulation, tempo, length of attack, length of decay, consonance, harmonic

complexity, randomness, syncopation, roughness, reverberation, and mode; two more attributes targeted emotional representation of valence and arousal. Sets of short sound clips (2-10 seconds) were composed to insure that the only variation between each clip in the set was in one specific musical attribute (see Table 2 for specifications).

The choice of auditory attributes was informed by the previous literature. One group of auditory attributes was selected to reflect previously known crossmodal correspondences, most notably from touch (irregularity, roughness) and vision (angularity), which are both driven by the tactile features of salt crystals. The auditory attributes of randomness and syncopation were selected to reflect irregularity. Harmonic complexity, distortion, and consonance were selected to reflect the notion of roughness, since dissonant intervals and harmonies give the impression of roughness (Vassilakis, 2001). Moreover, articulation was selected to reflect the visual/tactile notion of angularity, because the staccato(legato) articulation results in a more “pointed”(“rounded”) sound and is associated with angular(round) shapes (Reinoso Carvalho et al., 2017).

Another group of attributes was selected to reflect the lingering temporal experience of tasting salt (Obrist et al., 2014, Figure 5). These include tempo, length of attack, length of decay, and reverberation.

Finally, one group of attributes was selected to capture the emotional experience of tasting salt (Obrist et al., 2014). This included mode (major/minor) as well as two soundtrack excerpts designed to explicitly assess the association of saltiness with emotional dimensions of valence and arousal. The music excerpts for the emotion questions were taken from the Emotion Film Music database (Eerola & Vuoskoski, 2010). In the valence question, two excerpts were chosen with similar levels of arousal but different levels of valence (positive valence = clip #335, valence = 6.80, arousal = 2.80; negative valence = clip #231, valence = 2.00, arousal = 2.80). In the arousal question, two excerpts were chosen with similar levels of valence but different levels of arousal (high arousal = clip #246, valence = 5.50, arousal = 6.33; low arousal = clip #275, valence = 5.50, arousal = 1.50).

Table 2. Musical attributes investigated in Study 1 with definitions and operationalisations, arranged by the nature of their association with saltiness. Eleven of the thirteen attributes were assessed via specifically composed sound clips, the remaining two attributes (valence, arousal) were selected from the Emotion Film Music database.

Nature of association	Attribute	Definition	Number of levels/sound clips	Operationalisation
Tactile features	Articulation	The degree to which notes are separated or connected.	3	A five note series played three ways ranging from legato to staccato.
Tactile features	Harmonic complexity	The number and quality of notes playing simultaneously at a given time.	3	An eight note melody either played alone, or harmonised with one (thirds) or two (thirds and fifths) simultaneous notes.
Tactile features	Randomness	The lack of regular rhythmic pattern.	3	A three-second shaker rhythm pattern ranging from very regular to random.
Tactile features	Syncopation	The deliberate upsetting of the meter or pulse of a composition by means of a temporary shifting of the accent to a weak beat or an off-beat.	3	An eight-second shaker rhythm ranging from very regular to syncopated.
Tactile features	Roughness	The buzzing, rattling auditory sensation accompanying narrow harmonic intervals.	3	A three-note synth arpeggio with increasing levels of auditory roughness, implemented by adding distortion to the notes.
Tactile features	Consonance	A combination of notes producing a feeling of stability.	3	A three-note chord ranging in three levels of consonance, from C major (very consonant) to a cluster chord containing two minor seconds (very dissonant).
Temporal	Tempo	The speed of the rhythm of a composition.	3	A drum loop at three tempi ranging from 100, 120, 140 beats per minute.
Temporal	Length of attack	The initial part of the sound envelope of a note, from start to reaching full volume.	3	A single note with three levels of attack length at 0 ms, 200 ms, and 800 ms.
Temporal	Length of decay	The time it takes for a note to gradually disappear from the audible spectrum.	3	A single note with three levels of decay length at 0 ms, 700 ms, 2,000 ms.
Temporal	Reverberation	The persistence of a sound after the source stops emitting it.	3	Four woodblock hits with increasing levels of reverberation.
Emotion	Mode	A series of notes into which the octave is divided.	2	Three-note arpeggio in C major or minor scale.
Emotion	Valence	Emotional dimension corresponding to pleasantness	2	Five-second excerpts of clip 335 and 231 from the Emotion Film Music database
Emotion	Arousal	Emotional dimension spanning calm-excitement.	2	Five-second excerpts of clip 335 and 231 from the Emotion Film Music database

2.3 Procedure

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

After the sound check, the participants answered 13 questions, one for each auditory attribute tested, in which they chose the sound clip that best matched with saltiness. They also rated their confidence on a visual analogue scale (VAS) anchored by “not at all” on one end and “extremely confident” on the other (see Figure 1 for question format). For articulation, tempo, length of attack, length of decay, consonance, harmonic complexity, randomness, syncopation, roughness, and reverberation, there were three sound clips to choose from for each attribute. For the attributes of modality, valence, and arousal, there were two 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. The 13 questions were presented to the participants in random order.

Please select the sound clip that matches best with saltiness:

Clip 812

Clip 134

Clip 229

How confident are you with your choice?

Not at all Moderately confident Extremely confident

NEXT

Figure 1. Sample question from the online study.

After the sound selection questions, the participants reported their gender, age, and the extent to which they agreed with the following five statements on five-point scales. Salt liking was measured by agreement with the statement “I like the taste of salt”; salt usage was measured by the statements “I tend to use the salt shaker when eating out” and “I tend to use the salt shaker when eating at home”; salt conscientiousness was expressed by these two statements: “I watch the amount of salt I consume”, and “Whenever possible, I buy the reduced salt option”.

The study lasted for approximately 20 minutes.

2.4 Data Analysis

A chi-square goodness of fit test was conducted to assess whether the distribution of choices for each question were random. As there were thirteen auditory attributes in total, all chi square values were corrected by Bonferroni correction ($\alpha = .0038$). After checking for data normality, confidence ratings were analysed via a one-way Analysis of Variance (ANOVA) across different auditory attributes. Since some sound attributes were tested where participants were given two alternatives (i.e., ones with direct emotional associations), while other sound attributes included three choices, we could not directly compare confidence ratings across all thirteen sound attributes (Li & Ma, 2020). Instead, we performed the ANOVA analyses separately for the sound attributes with two alternatives versus those with three alternatives.

To assess whether individual confidence levels were higher when they happened to fall in line with the consensual response, two-way ANOVA analyses was performed on confidence ratings between major and non-major choices across two-alternative and three-alternative auditory attributes. As part of all ANOVA analyses, where assumption of sphericity had been violated, Greenhouse-Geisser or Huynh-Feldt degree of freedom corrections were applied based on Greenhouse-Geisser estimates of sphericity. Furthermore, all post-hoc pairwise comparisons were Bonferroni-corrected.

Responses from the salt attitude questions were averaged into one general salt attitude scale (with score points from 1 to 5), taking into account people's self-reported salt liking, salt usage, and salt conscientiousness. A high score on the general salt attitude scale reflects high salt liking/usage and low salt conscientiousness, whereas a low score on the general salt attitude reflects low salt liking/usage and high salt conscientiousness (the salt conscientiousness scale was reversed before the average was taken). Age, gender, and salt attitude were taken into account as individual differences between participants' salt-sound matches in a correlation analysis using Spearman's rho.

3. Results

3.1 Salt-sound matches

The results of salt-sound associations for all participants are visualised in Figure 2. A chi-square test of goodness of fit was calculated for each auditory attribute to determine which of them were matched to saltiness at a level that was significantly different from chance (see Table S1 in Supplementary materials). All 13 auditory attributes had significantly non-random distributions of answers. However, the distribution of responses for emotional valence, length of sound decay, emotional arousal, roughness, mode, and rhythm randomness had the greatest level of deviation from random guessing, as determined by the most frequent choice being selected at least 1.5 times as much as the next most frequent choice. More specifically, saltiness was associated with negative valence, minor mode, long sound decay duration, high arousal, rough sonic texture, and regular rhythm.

Figure 2. Results of participants' soundtrack-saltiness matches, for all 13 auditory attributes. Those attributes for which there were three possible alternatives are shown in yellow-orange colours, and those attributes for which there

were only two possible alternatives are shown in shades of red. All attributes have non-random distributions of answers, according to a chi-square test of goodness of fit and corrected by Bonferroni correction ($\alpha = .0038$).

3.2 Confidence ratings

Participants' average confidence ratings of sound-saltiness matchings are shown in Figure 3. An one-factor ANOVA revealed that confidence ratings were significantly different amongst the sound attributes for which two alternatives were provided ($F(1.99, 3616.95) = 127.66, p < .001, \eta_p^2 = 0.066$), with the confidence level being highest for valence and lowest for mode ($p < .001$ for all comparisons). Similarly, confidence ratings were significantly different for the attributes where three alternatives were provided ($F(8.26, 15022.08) = 61.53, p < .001, \eta_p^2 = 0.033$), with the confidence level being highest for tempo and randomness ($p < .05$ compared to all other attributes except roughness), and lowest for attack duration ($p < .001$ compared to all other attributes).

Figure 3. Average saltiness matching confidence level for sound attributes for which there were A) two alternatives and B) three alternatives. Error bars indicate standard error of the mean. Different letters indicate significant differences between confidence levels.

In terms of the relationship between individual confidence and group consensus, results showed a similar pattern for the sound attributes where two alternatives were given (mode, emotional valence, emotional arousal) and for those where three alternatives were given. For the two-alternative sound attributes, the mean confidence judgment for majority responses ($M = 58.16$, $SE = 0.48$) was significantly higher than for the non-majority responses ($M = 54.29$, $SE = 0.68$, $F(1,5451) = 27.17$, $p < .001$, $\eta_p^2 = 0.005$). Likewise, for the three-alternative sound attributes, the mean confidence judgment for majority responses ($M = 56.83$, $SE = 0.27$) was significantly higher than the confidence judgment for the non-majority responses ($M = 53.06$, $SE = 0.27$, $F(1, 18170) = 69.64$, $p < .001$, $\eta_p^2 = 0.004$).

Table 3 presents mean confidence judgments for the majority and non-majority responses for each of the thirteen attributes. For 10 of the 13 attributes, when the majority response was chosen, it was also endorsed with significantly higher confidence compared to when the non-majority response was chosen.

Table 3. Average confidence ratings for major and non-major saltiness-matching choices each sound attribute. Bold attribute name indicates a significant difference between the confidence ratings between when the participant makes a major choice (e.g., agreement with group consensus) versus a non-major choice (e.g., non-agreement with group consensus).

<i>Attributes with two alternatives</i>				
Attribute	Group consensus (χ^2)	Average major choice confidence (SE)	Average non-major choice confidence (SE)	<i>p</i>
Mode	129.32	53.57 (0.85)	48.77 (1.11)	.001
Arousal	176.74	60.62 (0.83)	55.08 (1.15)	< .001
Valence	300.23	61.74 (0.80)	59.01 (1.24)	.064
<i>Attributes with three alternatives</i>				
Attribute	Group consensus (χ^2)	Average major choice confidence (SE)	Average non-major choice confidence (SE)	<i>p</i>

Syncopation	23.75	54.84 (1.11)	52.92 (0.85)	.169
Harmonic complexity	36.74	53.47 (1.07)	49.19 (0.87)	.002
Reverberation	38.76	56.78 (1.08)	53.05 (0.86)	.007
Tempo	72.14	60.69 (1.05)	56.48 (0.88)	.002
Attack	97.70	49.78 (1.05)	47.10 (0.88)	.050
Consonance	102.98	57.21 (0.89)	55.23 (0.89)	.146
Articulation	109.96	51.45 (1.01)	54.64 (0.90)	.018
Randomness	150.79	61.43 (0.99)	55.68 (0.92)	< .001
Roughness	172.01	53.49 (1.04)	53.49 (0.93)	< .001
Decay	295.63	57.75 (0.93)	49.14 (0.98)	< .001

3.3 Individual differences

An overview of participants' age and general attitude toward salt are shown in Figure 4.

Figure 4. Density graphs of participants' (A) age (in years) and (B) general salt attitude (1 = low salt liking/usage and high salt conscientiousness, 5 = high salt liking/usage and low salt conscientiousness) showing an overview of participant demographics.

Correlation analysis of participants' demographic data with salt-sound matching revealed that age was significantly correlated with tempo ($r_s = .12, p < .001$), decay ($r_s = .16, p < .001$), consonance ($r_s = -.07, p = .002$), complexity ($r_s = .09, p < .001$), roughness ($r_s = -.09, p < .001$), valence ($r_s = .09, p < .001$), and arousal ($r_s = .15, p < .001$). In other words, the older participants tended to match saltiness with sound clips featuring higher tempo, longer decay duration, more consonance, greater complexity, less roughness, more positive valence, and higher arousal (i.e., more exciting). Gender was significantly correlated with articulation ($r_s = .07, p = .003$), with female participants being more likely to match saltiness with more staccato articulation.

In terms of salt attitude, there was a significant correlation with age ($r_s = -.11, p < .001$), where older participants had lower salt liking/usage and higher salt usage conscientiousness compared to younger participants. There were no significant correlations between overall salt attitude and salt-sound matching for any of the sound attributes.

4. Discussion

Overall, the auditory attributes that were associated most strongly with saltiness were based on emotional associations (negative valence, minor mode, and high arousal), long decay durations, regular rhythm, and a high degree of auditory roughness. These attributes encompass a combination of attributes from tactile, temporal, and emotional associations. This is comparable with previous results demonstrating that odour-based crossmodal matches span semantic, perceptual, and hedonic mechanisms (Spence, 2019, 2020b; Stevenson et al., 2012).

The fact that emotional auditory attributes had strong matches with saltiness may be surprising, given our introductory remarks concerning the lack of emotional associations for saltiness. This could be due to two factors. One, participants were given a forced choice between two attributes for all the emotional questions without having a neutral third option, making it difficult to quantify the intensity of the emotional associations. Two, crossmodal correspondences could also depend

on the context of the question and the mental state of the participant. Thinking about saltiness in general, apart from the context of a dish, could have resulted in overly negative associations. At the same time, hunger state could have also modulated people's emotional, at least reward, associations with saltiness (Finlayson et al., 2007).

When it comes to confidence ratings, we observed a consensuality effect whereby there was a positive correlation between confidence ratings and the degree of group consensus about salty-sound matches, and with participants in general being confident in their choices when it was in agreement with the group consensus. Moreover, the relative confidence displayed by the participants speak to the nature of the sound-saltiness correspondences. Those attributes associated with higher levels of confidence are those based on emotional associations (e.g. valence, arousal) as well as those characteristics strongly associated with the tactile elements of actual salt crystals (e.g., roughness, randomness). Looking back to other taste-based correspondences, colour-taste correspondences exhibit clearer cross-population agreement as compared to sound-taste correspondences, plausibly because colour-taste is based on associative learning, whereas sound-taste relies on more mediated correspondences via shared emotions or semantic associations (excluding intensity/magnitude-based correspondences which stem from structural associations; Spence et al., 2019; Spence & Levitan, 2021).

We also observed some individual differences in terms of how participants made salt-sound matches, with older participants associating saltiness with a notably longer decay time, higher arousal, and faster tempo. This is especially interesting when combined with the fact that older participants also had lower salt liking/usage compared to younger participants. However, it should be noted that, where it might be expected that those who particularly like/dislike salt might choose liked/disliked sonic attributes to match with saltiness, we did not observe any differences in crossmodal matches between salt and sound attributes based on one's attitude towards salt.

While having a large sample size (see Table 1 for comparisons of sample sizes from previous studies), the present study has several limitation stemming from the fact that this was an online study conducted with North American participants. Although this crowd-sourced study has avoided the typical target population of university students, mainly by surveying a large sample of users of a major streaming music company, our participants were nevertheless WEIRD (Western and educated, from industrialised, rich, and democratic countries, Henrich et al., 2010a, b). Our

study was also limited by the fact that we asked participants to match only the concept of “saltiness” with sounds, without controlling for what exactly they imagined. Presumably, matching sounds with a mental image of salt crystals might be very different from imaging the taste of salt water or tears, in terms of which tactile-related auditory attributes would be involved. Furthermore, the degree of saltiness could also influence people’s matching choice; for instance, something that is just salty enough would have different emotional associations compared to a tasting experience that is too salty (Bredie et al., 2014). Another limitation in the study involves our operationalisation of salt usage. This was confined to two questions related to the tendency for participants to use the salt shaker, which might be confounded by different dietary usage habits. Questions related to salt use in general while cooking or eating, not limited to the salt shaker itself, might have given us a more generalisable view of people’s attitude towards salt.

In conclusion, the present large-sample study has uncovered new sound-saltiness correspondences, as well as highlighted the consensuality principle underlying such correspondences. Methodologically speaking, the study demonstrates the relative ease of conducting crowdsourced studies online (Woods et al., 2015), especially in these increasingly isolated times (given the Covid-19 pandemic). Beyond any theoretical contributions, the results of the present study potentially have practical relevance for those sonic practitioners aiming to promote healthier purchasing and food consumption behaviour. By identifying new salt-corresponding auditory elements, it becomes possible to create soundtracks that are more easily associated with saltiness. Soundtracks corresponding with different tastes and cultural associations have been shown to influence both explicit food choice and implicit eye movement (Peng-Li et al., 2020ab), both in the laboratory and in field studies (Biswas et al., 2019). Therefore, a salty soundtrack might be used to encourage the purchase of more savoury products (compared to, say, relatively less healthy sweet products in the context of a snack bar). Moreover, in terms of eating behaviour, sonic seasoning has already been proven to reliably enhance taste, mouthfeel, and trigeminal attributes such as sweetness, creaminess, and spiciness (Reinoso Carvalho et al., 2017; Wang et al., 2017, 2019). Therefore, the judicious use of salty-congruent soundtracks may be able to draw people’s attention selectively to the salty elements in the food, thereby heightening the perceived level of salt without while reducing salt consumption.

ACKNOWLEDGEMENTS

This work was financially supported by an Aarhus University Research Foundation starting grant awarded to Q.J.W.

AUTHOR CONTRIBUTIONS

Q.J.W. and S.K. conceived the study. Q.J.W. and S.K. collected the data. Q.J.W. analysed the results with discussions and input from C.S. Q.J.W. wrote the initial manuscript and all authors contributed to revising the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

- Adams, F. M., & Osgood, C. E. (1973). A cross-cultural study of the affective meanings of color. *Journal of Cross-Cultural Psychology*, **4**(2), 135-156.
- Batenburg, M., & van der Velden, R. (2011). Saltiness enhancement by savory aroma compounds. *Journal of Food Science*, **76**, S280-S288.
- Biswas, D., Lund, K., & Szocs, C. (2019). Sounds like a healthy retail atmospheric strategy: Effects of ambient music and background noise on food sales. *Journal of the Academy of Marketing Science*, **47**, 37-55.
- 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.
- Chen, Y.-C., Huang, P.-C., Woods, A., & Spence, C. (2019). I know that “Kiki” is angular: The metacognition underlying sound–shape correspondences. *Psychonomic Bulletin & Review*, **26**, 261-268.
- Crisinel, A.-S., & Spence, C. (2010a). As bitter as a trombone: Synesthetic correspondences in nonsynesthetes between tastes/ flavors and musical notes. *Attention, Perception, & Psychophysics*, **72**(7), 1994-2002.
- Crisinel, A.-S., & Spence, C. (2010b). A sweet sound? Exploring implicit associations between basic tastes and pitch. *Perception*, **39**, 417-425.
- Desor, J. A., Maller, O., & Andrews, K. (1975). Ingestive responses of human newborns to salty, sour, and bitter stimuli. *Journal of Comparative and Physiological Psychology*, **89**, 966-970.
- Eerola, T., & Vuoskoski, J. K. (2010). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, **39**, 18-49.
- Fateminia, M., Ghotbabadi, T. D., & Azad, K. M. (2020). Perceptions of the taste of colors in children and adults. *Color Research and Application*, **45**, 743-753.

- Finlayson, G., King, N., & Blundell, J. E. (2007). Is it possible to dissociate ‘liking’ and ‘wanting’ for foods in humans? A novel experimental procedure. *Physiology & Behavior*, **90**, 36-42.
- Gifford, S. R., Clydesdale, F. M., & Damon, R. A., Jr. (1987). The psychophysical relationship between color and salt concentrations in chicken flavored broths. *Journal of Sensory Studies*, **2**, 137-147.
- Guetta, R., & Loui, P. (2017). When music is salty: The crossmodal associations between sound and taste. *PLoS ONE*, **12**(3):e0173366.
- Harrar, V., Piqueras-Fiszman, B., & Spence, C. (2011). There’s more to taste in a coloured bowl. *Perception*, **40**, 880-882.
- Heller, E. (1999). *Wie farben wirken. Farbpsychologie, farbsymbolik, kreative farbgestaltung* [How colour works. Colour psychology, colour symbolism, working creatively with colour]. Hamburg: Reinbek.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). The weirdest people in the world? *Behavioral and Brain Sciences*, **33**, 61-135.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). Most people are not WEIRD. *Nature*, **466**, 29.
- Höchenberger, R., & Ohla, K. (2019). A bittersweet symphony: Evidence for taste-sound correspondences without effects on taste quality-specific perception. *Journal of Neuroscience Research*, **97**, 267-275.
- 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.
- Koch, C., & Koch, E. C. (2003). Preconceptions of taste based on color. *Journal of Psychology*, **137**, 233-242.
- Koriat, A. (2008). Subjective confidence in one’s answers: The consensuality principle. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **34**, 945-959.
- Koriat, A. (2011). Subjective confidence in perceptual judgments: A test of the self-consistency model. *Journal of Experimental Psychology: General*, **140**, 117-139.
- Li, H.-H., & Ma, W. J. (2020). Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis. *Nature Communications*, **11**, 2004. <http://doi.org/10.1038/s41467-020-15581-6>.
- Mesz, B., Trevisan, M., & Sigman, M. (2011). The taste of music. *Perception*, **40**, 209-219.
- 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.
- O’Mahony, M. (1983). Gustatory responses to nongustatory stimuli. *Perception*, **12**, 627-633.

- Padulo, C., Tommasi, L., & Brancucci, A. (2018). Implicit association effects between sound and food images. *Multisensory Research*, **31**, 779-791. doi: 10.1163/22134808-20181308.
- Peng-Li, D., Byrne, D. V., Chan, R. C. K., & Wang, Q. J. (2020a). The effects of ethnically congruent music on eye-movements and food choice – A cross-cultural comparison between East and West. *Foods*, **9**, 1109.
- Peng-Li, D., Byrne, D. V., Chan, R. C. K., & Wang, Q. J. (2020b). The influence of taste-congruent soundtracks on visual attention and food choice: A cross-cultural eye-tracking study in Chinese and Danish consumers. *Food Quality & Preference*, **85**, 103962.
- Rader, C. M., & Tellegen, A. (1987). An investigation of synesthesia. *Journal of Personality and Social Psychology*, **52**, 981-987.
- 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.
- Saluja, S., & Stevenson, R. J. (2018). Cross-modal associations between real tastes and colors. *Chemical Senses*, **43**, 475-480.
- Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mapping across gustation and audition. *Perception*, **39**, 553-569.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, **73**, 971-995.
- Spence, C. (2012). 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. (2019). On the relationship(s) between colour and taste. *Experimental Psychology*, **66**, 99-111.
- Spence, C. (2020a). Assessing the role of emotional mediation in explaining crossmodal correspondences involving musical stimuli. *Multisensory Research*, **33**, 1-29.
- Spence, C. (2020b). Olfactory-colour crossmodal correspondences in art, science, & design. *Cognitive Research: Principles & Implications (CRPI)*, **5**:52.
- Spence, C., & Levitan, C. A. (2021). Explaining crossmodal correspondences between colours and tastes. *i-Perception*, **12**(2):1-28. DOI: 10.1177/20416695211018223.
- Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., & Deroy, O. (2015). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour*, **4**:23.
- Stevenson, R. J., Rich, A., & Russell, A. (2012). The nature and origin of cross-modal associations to odours. *Perception*, **41**, 606-619.
- Thomas-Danguin, T., Guichard, E., & Salles, C. (2019). Cross-modal interactions as a strategy to enhance salty taste and to maintain liking of low-salt food: A review. *Food & Function*, **10**, 5269-5281. doi:10.1039/c8fo02006j.
- Tomasik-Krótki, J., and Strojny, J. (2008). Scaling of sensory impressions. *Journal of Sensory Studies*, **23**, 251-266.

- Van Rompay, T. J. L., & Groothedde, S. (2019). The taste of touch: Enhancing saltiness impressions through surface texture design. *Food Quality and Preference*, **73**, 248-254.
- Vassilakis, P. (2001). Auditory roughness estimation of complex spectra—Roughness degrees and dissonance ratings of harmonic intervals revisited. *The Journal of the Acoustical Society of America*, **110**, 2755. <https://doi.org/10.1121/1.4777600>
- Velasco, C., Woods, A., Deroy, O., & Spence, C. (2015a). Hedonic mediation of the crossmodal correspondence between taste and shape. *Food Quality and Preference*, **41**, 151-158.
- Velasco, C., Woods, A. T., Hyndman, S., & Spence, C. (2015b). The taste of typeface. *i-Perception*, **6**, 1-10.
- 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.
- Wan, X., Woods, A. T., van den Bosch, J., McKenzie, K. J., Velasco, C., & Spence, C. (2014). Cross-cultural differences in crossmodal correspondences between tastes and visual features. *Frontiers in Psychology: Cognition*, **5**:1365.
- Wang, Q. (J.), Keller, S., & Spence, C. (2017). Sounds spicy: Enhancing the evaluation of piquancy by means of a customised crossmodally congruent soundtrack. *Food Quality & Preference*, **58**, 1-9.
- Wang, Q. J., Mielby, L. A., Junge, J. Y., Bertelsen, A. S., Kidmose, U., Spence, C., & Byrne, D. V. (2019). The role of intrinsic and extrinsic sensory factors in sweetness perception of food and beverages: A review. *Foods*, **8**:211.
- 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.
- 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)**:1-23.
- Wongthahan, P., Sae-eaw, A., & Prinyawiwatkul, W. (2020). Sensory lexicon and relationships among brown colour, saltiness perception and sensory liking evaluated by regular users and culinary chefs: A case of soy sauces. *International Journal of Food Science & Technology*, **55**, 2841-2850.
- Woods, A. T., Marmolejo-Ramos, F., Velasco, C., & Spence, C. (2016). Using single colours and colour pairs to communicate basic tastes II; Foreground-background colour combinations. *i-Perception*, **7**:5.
- Woods, A. T., & Spence, C. (2016). Using single colors and color pairs to communicate basic tastes. *i-Perception*, **7(4)**:1-15.
- Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the internet: A tutorial review. *PeerJ*, **3**:e1058.

SUPPLEMENTARY INFORMATION

Table S1. Results of participants' soundtrack-saltiness matches, for all 13 auditory attributes. The chart reveals the distribution of participants' choice of soundtrack that best evoked saltiness for each auditory attribute, with a total of 1819 responses for each attribute. All attributes have non-random distributions of answers, according to a chi-square test of goodness of fit and corrected by Bonferroni correction ($\alpha = .0038$). Within each attribute, the choice with most votes is shown in bold. All attributes with three choices are shown first, sorted by decreasing order of chi-square values, followed by those attributes with two choices.

<i>Attributes with three alternatives</i>						
				χ^2	<i>df</i>	<i>p</i>
Decay	Shortest	Middle	Longest			
	437	430	952	295.63	2	< .001
Roughness	Least rough	Middle	Most rough			
	526	429	864	172.01	2	< .001
Randomness	Least random rhythm	Middle	Most random rhythm			
	842	552	425	150.79	2	< .001
Articulation	Legato	Middle	Staccato			
	562	450	807	109.96	2	< .001
Consonance	Consonant	Middle	Dissonant			
	630	419	770	102.98	2	< .001
Attack	Shortest	Middle	Longest			
	652	416	751	97.70	2	< .001
Tempo	Slowest tempo	Middle	Fastest tempo			
	450	625	744	72.14	2	< .001
Reverberation	Least reverberation	Middle	Most reverberation			
	493	617	709	38.76	2	< .001
Harmonic complexity	Least complex	Middle	Most complex			
	523	571	725	36.74	2	< .001

Syncopation	Least syncopated rhythm	Middle	Most syncopated rhythm			
	639	510	670	23.75	2	< .001
<i>Attributes with two alternatives</i>						
				χ^2	<i>df</i>	<i>p</i>
Valence	Positive	Negative				
	540	1279		300.23	1	< .001
Arousal	High	Low				
	1193	626		176.74	1	< .001
Mode	Major	Minor				
	667	1152		129.32	1	< .001