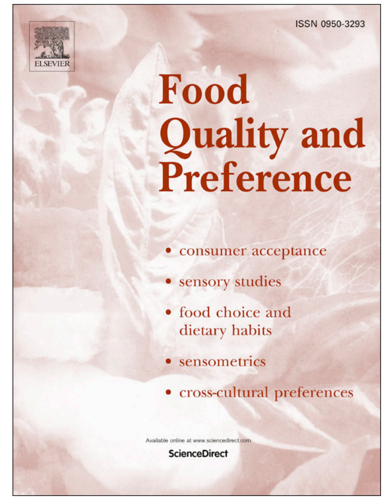
Journal Pre-proofs




Crossmodal interactions between audition and taste: A systematic review and narrative synthesis

David Guedes, Margarida Vaz Garrido, Elsa Lamy, Bernardo Pereira Cavalheiro, Marília Prada

PII: S0950-3293(23)00050-2
DOI: https://doi.org/10.1016/j.foodqual.2023.104856
Reference: FQAP 104856

To appear in: *Food Quality and Preference*

Received Date: 13 December 2022
Revised Date: 17 March 2023
Accepted Date: 19 March 2023

Please cite this article as: Guedes, D., Vaz Garrido, M., Lamy, E., Pereira Cavalheiro, B., Prada, M., Crossmodal interactions between audition and taste: A systematic review and narrative synthesis, *Food Quality and Preference* (2023), doi: https://doi.org/10.1016/j.foodqual.2023.104856

**Crossmodal interactions between audition and taste: A systematic review and narrative synthesis**

David Guedes[1,2], Margarida Vaz Garrido[1], Elsa Lamy[2], Bernardo Pereira Cavalheiro[1], Marília Prada[1]

[1] Iscte - Instituto Universitário de Lisboa, CIS_Iscte, Portugal

[2] MED - Mediterranean Institute for Agriculture, Environment and Development & CHANGE - Global Change and Sustainability Institute, University of Évora, Évora, Portugal



*Correspondence concerning this article should be addressed to D. Guedes, Iscte - Instituto Universitário de Lisboa, Av. das Forças Armadas, Room 2W08, 1649-026, Lisboa, Portugal.

E-mail: dhfgs@iscte-iul.pt

**Abstract**

Taste perception results from integrating all the senses. In the case of audition, research shows that people can associate certain auditory parameters (e.g., pitch) with basic tastes. Likewise, the surrounding sonic environment (e.g., noise, music) may influence individuals' evaluation of the taste attributes of foods and drinks. This paper presents the first pre-registered systematic examination of the literature on the crossmodal interactions between audition and taste. For that purpose, four indexing services (EBSCOhost, SCOPUS, Web of Science, and PubMed) were searched using three sets of keywords on the crossmodal interactions between audition and basic tastes. Empirical, quantitative studies with healthy subjects in field, lab, or online settings were considered for inclusion. A total of 2484 records (*n* = 1481 after removing duplicates) were subject to abstract and title screening, followed by a full-text screening (*n* = 79). Sixty articles, reporting 94 eligible studies, were reviewed. Results suggest that taste may be crossmodally associated with a) pitch and musical instruments; b) words, nonwords, and speech sounds; and c) music and soundtracks. Moreover, the reviewed evidence supports the employment of auditory stimuli in the context of taste modulation, specifically in the case of a) familiar music; b) custom soundtracks, and c) noise, tones, and soundscapes. Overall, this review provides a comprehensive outlook on the multisensory interactions between audition and taste. The results show that audition has a relevant contribution to taste perception with important implications for how foods and drinks are perceived. The theoretical and practical implications of these findings are discussed.

Keywords: audition, taste, multisensory perception, crossmodal correspondences, sonic seasoning, systematic review

## 1. Introduction

Taste perception plays a central role in determining food preference and choice. Since early in ontogeny, infants show a preference for foods with sweet tastes, while bitterness is a common reason for rejection (Mennella & Bobowski, 2015). Later in life, basic taste sensations (sweetness, bitterness, saltiness, sourness, and umami) are still relevant determinants of preference. For instance, salt and sugar are currently consumed in excessive amounts on a global scale, despite the negative health consequences (World Health Organization, 2012, 2015). Likewise, sensitivity to basic tastes has been previously associated with food preference and choice in adults (Chamoun et al., 2019; Proserpio et al., 2016).

Despite its major contribution to eating enjoyment, a great deal of what people think of when they refer to taste is actually flavor, that is, the combination of gustatory and olfactory sensations. An illustrative example is the common reporting of taste loss when olfactory perception is compromised (namely due to COVID-19 infections), even when taste function is intact (Le Bon et al., 2021). The classic confusion between taste and olfaction may be explained by the fact that these senses (and, to a lesser extent, also the trigeminal modality) are combined to form a unitary perception of flavor (Auvray & Spence, 2008; Stevenson, 2014). This binding process is associated with why some people may refer to certain odors, such as vanilla, as “sweet” and why, in turn, some of these odors may lead to increased perceived taste intensities (Stevenson et al., 1999). Indeed, sweetness enhancement by the addition of aromas is a well-established sensory trick for improving the acceptance of low-sugar products (Bertelsen et al., 2020, 2021).

To a certain degree, all the senses contribute to flavor perception. However, an important distinction should be made between the senses that are constitutive and those that are merely modulatory of flavor perception. While this distinction is still a matter of debate, there seems to be some agreement that taste and olfaction are intuitive examples of the constitutive senses, whereas vision and audition are generally reserved a role as modulatory senses (Skrzypulec, 2021; Spence, 2015a). In the visual modality, the colors of foods and drinks (e.g., Calvo et al., 2001) or of the plateware/glassware (Piqueras-Fiszman et al., 2012) in which they are served may account for different tasting experiences (for a review, see Spence, 2019). As far as audition is concerned, there are several ways by which hearing may influence how people perceive food (Spence et al., 2019). Yet, the role of this sensory modality is not always acknowledged in the context of eating. It has been previously argued that audition has been neglected in the context of multisensory research and that sound may have a much more significant impact on eating than what it has been given credit for (Spence, 2016). In this review, we systematically examine the existing evidence on the crossmodal links between audition and taste.

### 1.1. Audition and taste perception

Some foods produce singular sounds, such as the crunch of biting an apple or the crackling sound of a spoon on a *crème brûlée*. To understand how sonic cues may impact the sensory experience, one seminal experiment had participants taste potato chips while listening to their own mastication sounds through headphones, either unaltered or manipulated for volume and frequency (Zampini & Spence, 2004). The results suggested that chips were perceived as fresher and crisper when listening to the sound with amplified frequency and/or volume compared to when the sound was unaltered.

Apart from the sounds of foods themselves, environmental sounds (e.g., the soundscape of a busy cafeteria vs. that of a Michelin-star dining room) may also affect eating behavior differently. For instance, listening to music during eating is associated with longer meals and higher food intake (Stroebele & de Castro, 2006). Loud background music, in particular, may lead to increased consumption of soft and alcoholic drinks (Guéguen et al., 2008; McCarron & Tierney, 1989). Listening to fast-tempo music can make participants drink faster (McElrea & Standing, 1992) and eat more quickly, accounting either for shorter eating times (Mathiesen et al., 2020) or a larger number of bites per minute (Roballey et al., 1985).

This bulk of research highlights that audition is implicated in eating and influences how we behave towards food. Another line of inquiry has shown that audition may also play a role in how the taste experience unfolds (Spence, 2012, 2015b, 2016; Yan & Dando, 2015). One of the first pieces of evidence of associations between audition and taste came from Holt-Hansen's (1968,1976) seminal experiments, in which distinct beer varieties were matched to different sound pitches. More recently, other examples of systematic crossmodal associations have been documented that link taste with various sounds (e.g., music, speech sounds) or sonic attributes (e.g., frequency, tempo). Just as individuals seem able to describe vanilla or caramel odors as "sweet", it appears that a similar ability may also exist for auditory stimuli, such as music pieces (e.g., Guedes et al., 2022). These consistent mappings between attributes of stimuli pertaining to different sensory modalities (such as audition and gustation) are known as "crossmodal correspondences" (Knöferle & Spence, 2012).

One important implication of such links between the auditory and gustatory modalities is the potential use of sound to modulate how people subjectively perceive the taste of foods and beverages in real-world contexts. Sounds may contribute to creating taste expectations, directing attention toward specific sensory attributes, or influencing thinking and feeling processes that change how individuals experience (or report experiencing) the taste of foods and drinks (Wang, 2017). While research in these topics is becoming increasingly prolific, no systematic effort has been attempted to map all the possible connections between taste and audition and integrate the existing evidence under one overarching review. Thus, the current work aims at i) providing an updated outlook on a rapidly growing body of literature; ii) addressing the issue of bias of traditional reviews by following a pre-registered systematic protocol; iii) mapping the diverse crossmodal interactions between audition and taste; and iv) identifying research gaps and future directions in the field.

This paper examines the crossmodal role of audition in taste perception from two perspectives. First, we review the evidence regarding the crossmodal correspondences between audition and taste, that is, studies examining the subjective associations people make between specific auditory stimuli (e.g., music) or sonic parameters (e.g., pitch) and basic tastes (sweetness, bitterness, sourness, saltiness, umami). Second, we examine how hearing may actually impact taste perception by reviewing experimental evidence testing the effects of exposure to different sound conditions on the perception of taste in foods and beverages.

## 2. Method

The study was approved by the ethical review board of Iscte – Instituto Universitário de Lisboa (Approval #117/2020).

### 2.1. Literature search

The study was pre-registered in PROSPERO and can be accessed here. A systematic search was conducted based on four indexing services (EBSCOhost, SCOPUS, Web of Science, and PubMed) in April 2020 and updated in July 2022. The search strategy was developed using the PICOS and SPIDER tools and included three sets of keywords regarding crossmodal or multimodal interactions (cross-modal* OR crossmodal* OR multi-sensor* OR multisensor* OR multimodal* OR multi-modal*) between audition (audition OR auditory OR sound* OR music* OR nois* OR sonic* OR hearing) and basic tastes (sweet* OR sugar* OR bitter* OR sour OR sourness OR salt* OR umami OR tast* OR flavor* OR flavour* OR gustat* OR acid* OR in-mouth). The search spanned titles, abstracts, and keywords in the four databases, and results were limited to peer-reviewed publications in four languages (English, French, Spanish, and Portuguese).

### 2.2. Inclusion and exclusion criteria

Empirical, quantitative studies in field, lab, or online settings were considered for inclusion. Given the phenomena of interest, studies selected for inclusion had to clearly relate audition and basic taste. For the sake of parsimony, only the five basic tastes were targeted (sweetness, saltiness, bitterness, sourness or acidity, and umami). Although some scholars make a case for a sixth taste of fat, the ongoing nature of the debate cautioned against its inclusion in this review (Besnard et al., 2015; Keast & Costanzo, 2015). The search did not include flavor variables (e.g., aromas) and oral-somatosensory attributes (e.g., texture, pungency). For the auditory domain, no *a priori* exclusion criteria were defined. Opinion/commentary, conference papers, and clinical studies (e.g., synesthesia) were excluded. Review papers were not considered for data extraction but were scanned to identify additional references. There were no further exclusion criteria regarding participants' characteristics (e.g., age, sex, body mass).

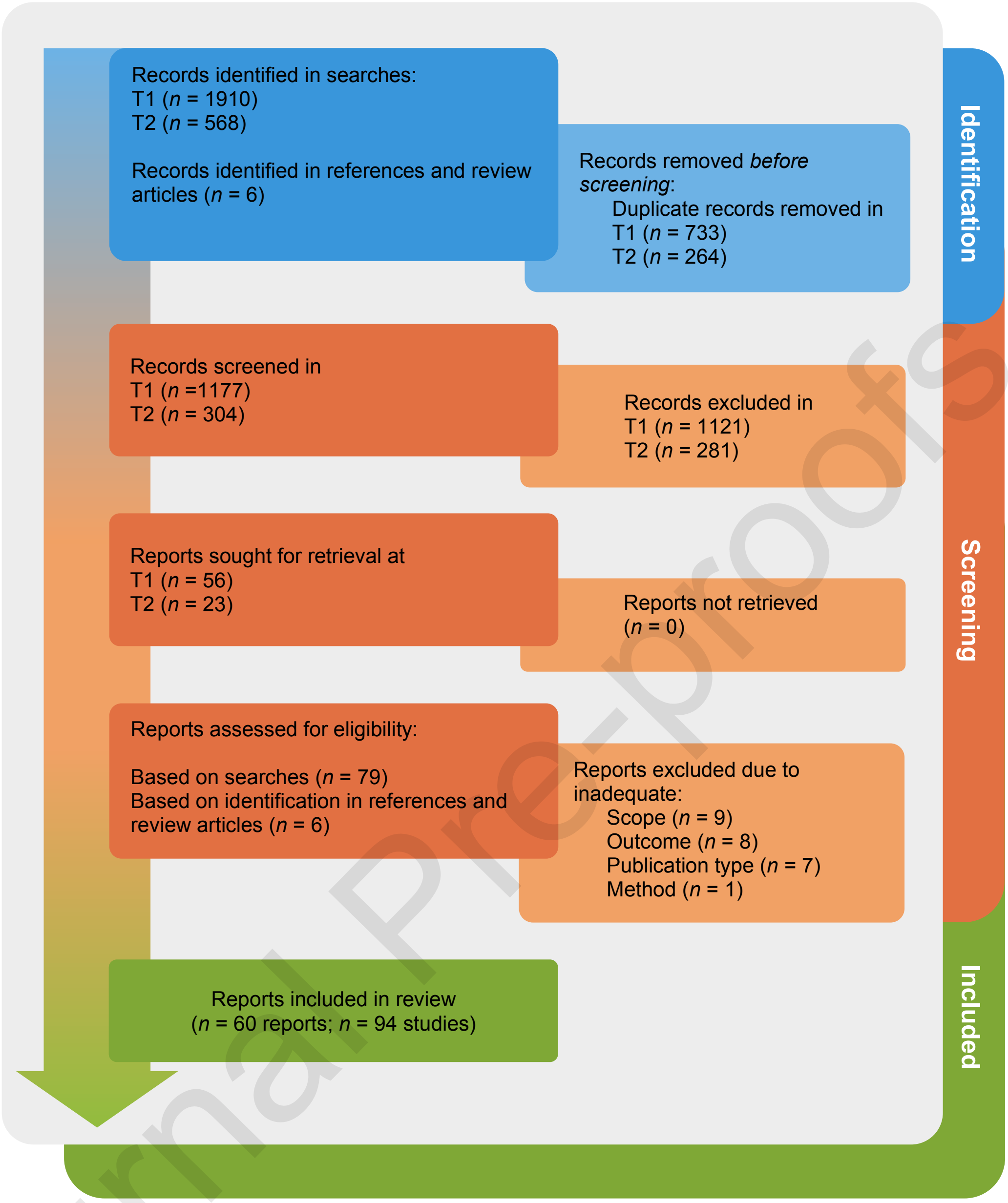
### 2.3. Selection of studies

The first search resulted in 1910 records managed in Endnote version X7. Two rounds of (automatic and manual) removal of duplicates resulted in 1193 records (see Figure 1). The remaining results were exported to the online reference management platform Rayyan (Ouzzani et al., 2016) for the title and abstract screening stage. This platform identified and excluded 16 additional duplicate records, and the remaining ($n$ = 1177) were screened by two independent reviewers. Inter-rater agreement was 99%, which resulted in Cohen's $\kappa$ = 0.88 ($p$ < .001), 95% CI [.823, .945]. Disagreements were resolved by discussion between the two reviewers. The final number of records selected in the title and abstract screening round was 56.

Following the method outlined above, an additional search took place in July 2022. This search resulted in 304 articles for title and abstract screening, of which 23 were subject to full-text screening. Across the two searches, 79 articles were subject to full-text screening, and 57 were retained for data extraction. Three additional records were included by scanning through the references of review papers, totaling 60 articles for review. An overview of the stages of selection and extraction of records is presented in Figure 1.

**Figure 1**

*PRISMA Flow Diagram Representing the Stages of Record Identification, Screening, and Inclusion*


Records identified in searches:
T1 (n = 1910)
T2 (n = 568)
Records identified in references and review articles (n = 6)
Records removed before screening:
Duplicate records removed in
T1 (n = 733)
T2 (n = 264)
Records screened in
T1 (n =1177)
T2 (n = 304)
Records excluded in
T1 (n = 1121)
T2 (n = 281)
Reports sought for retrieval at
T1 (n = 56)
T2 (n = 23)
Reports not retrieved
(n = 0)
Reports assessed for eligibility:
Based on searches (n = 79)
Based on identification in references and review articles (n = 6)
Reports excluded due to inadequate:
Scope (n = 9)
Outcome (n = 8)
Publication type (n = 7)
Method (n = 1)
Reports included in review
(n = 60 reports; n = 94 studies)
Identification
Screening
Included

*Note*. T1 = Time 1 (April 2020); T2 = Time 2 (July 2022).

### 2.4. Data extraction

One author led the data extraction process, and a second author reviewed the coded data. The extracted characteristics were:

1. Article information: Authors, title, year, journal, and country of authors' affiliations.
2. Sample: Type of sample (e.g., college students, restaurant patrons), sample size, number or percentage of women, age (mean, standard deviation, and/or range), and sample selection criteria.
3. Method: Scope of the study, setting (e.g., field, lab, online), general procedures, stimuli type and origin (auditory and gustatory), measures.
4. Design: Conditions, design, randomization, baseline or control conditions, follow-up, independent and dependent variables.
5. Analysis and results: Data analysis, number of participants per group, means, standard deviations or standard errors, inferential statistics (e.g., Student's *t*, *F*-statistic, Pearson's *r*), and summary of results.

## 3. Results

The full data regarding studies' characteristics (article information, sample, method, design, and results) is available at osf.io/t4r76.

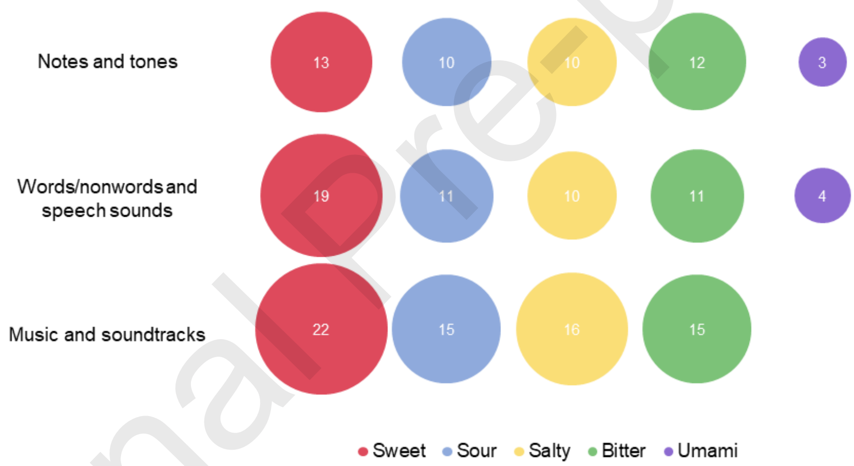
### 3.1. General characterization

The total number of manuscripts selected for extraction was 60, which reported 94 eligible studies. Most of these studies followed a within-subjects design ($n$ = 72), whereas 14 were mixed and eight were between-subjects. In terms of scope, 56 studies examined the crossmodal correspondences between audition and taste, and 38 tested the modulatory role of audition in taste perception.

Correspondence studies focused on the associations between basic tastes and (a) music and soundtracks[1] ($n$ = 22), (b) words, nonwords, and speech sounds ($n$ = 20), (c) tones, musical notes, and musical instruments ($n$ = 14). The most systematically examined taste category was sweetness ($n$ = 54), followed by bitterness ($n$ = 38), sourness ($n$ = 36), saltiness ($n$ = 36), and umami ($n$ = 7). The distribution of taste variables across categories of studies is presented in Figure 2.

**Figure 2**

*Number of Studies Examining the Correspondence Between Each Basic Taste and Various Categories of Sounds*



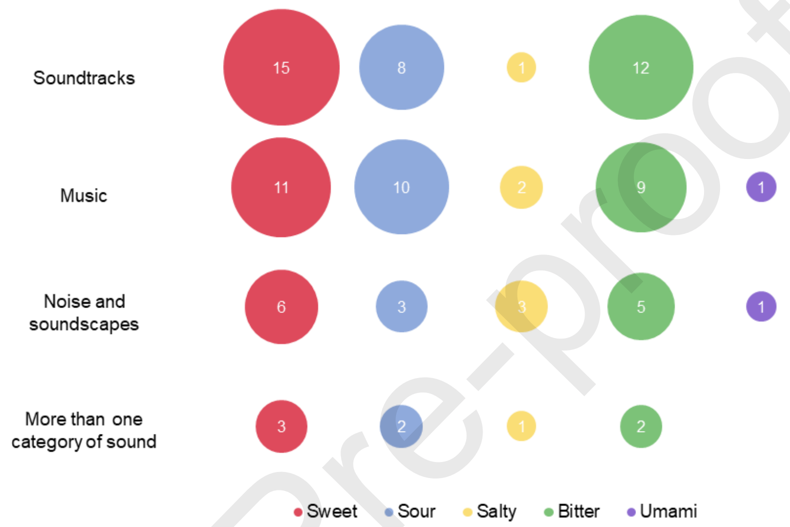

Studies testing the modulatory effects of audition in taste perception focused mainly on (a) music ($n$ = 12), (b) soundtracks ($n$ = 16), and (c) noise, tones, and soundscapes ($n$ = 7). Three studies simultaneously included stimuli from two categories, namely noise with soundtracks and music with soundscapes. Most studies examined the influence of sound on sweetness perception ($n$ = 35), followed by bitterness ($n$ = 28), sourness ($n$ = 23), saltiness ($n$ = 7), and umami ($n$ = 2). The

[1] Here, music refers to known pieces, usually in familiar genres (e.g., jazz, classical). Soundtracks are bespoke stimuli, commonly produced to match taste or flavor attributes.

number of studies testing the effects of different sounds on each of the basic tastes is presented in Figure 3.

**Figure 3**

*Number of Studies Examining the Influence of Different Sounds on the Perception of Each Basic Taste*




All the included articles were published after 2009, with about half (51.7%) released after 2017. International collaborations originated about 55% of the selected papers, with a majority of contributions from researchers affiliated with European institutions at the time of publication (in 66% of all papers).

**Table 1**

*Total Number of Studies Selected for Inclusion and Number of Studies per Category*

| | Number of Studies (*N* = 94) | % |
|---|---|---|
| Sound-taste correspondences studies | 56 | 60% |
| a. Music and soundtracks | 22 | 23% |
| b. Words, nonwords, and speech sounds | 20 | 21% |
| c. Tones, musical notes, and musical instruments | 14 | 15% |
| Taste modulation studies | 38 | 40% |

| | | |
|---|---|---|
| a. Music | 12 | 13% |
| b. Soundtracks | 16 | 17% |
| c. Noise, tones, and soundscapes | 7 | 7% |
| d. More than one category | 3 | 3% |

**3.2. Crossmodal correspondences between audition and taste**

**3.2.1. Tones, musical notes, and musical instruments**

The reviewed studies described several crossmodal correspondences, linking basic tastes with auditory attributes (for an overview, see Table 2). One of the most common psychoacoustical parameters associated with taste is sound frequency (or pitch). One of the earliest examples is Crisinel and Spence's (2009) study using the Implicit Association Test (IAT) to measure the association between tastes and sound frequency. Results indicated an association between higher-pitched sounds and sour-tasting foods and between lower-pitched sounds and bitter-tasting foods. Crisinel and Spence (2010a) presented further support in favor of the sour-high pitch association, as well as for a sweet taste-high pitch correspondence. Bitter and salty tastes, however, were not associated with lower-pitched sounds.

Since these studies relied on food names to represent taste categories (e.g., “dark chocolate” for bitter, “lemon juice” for sour), the same authors aimed to replicate these findings, using real tastants (e.g., sucrose) and flavors (e.g., vanilla) (Crisinel & Spence, 2010b). Here, participants had to select one musical note (played by different musical instruments) to match each gustatory stimulus. The results also pointed toward an association between sweet and sour tastes and high-pitched sounds, with participants choosing higher-pitched musical notes for the sucrose (sweet) and citric acid (sour) solutions, while caffeine (bitter) and monosodium glutamate (umami) led to the choice of lower-pitched sounds. Besides the association with musical notes, there were also differences in the choice of musical instruments (e.g., preference for piano sounds to match sucrose, while brass was preferred for caffeine).

A replication of Crisinel and Spence's (2009, 2010a, 2010b) studies presented further support for the association between bitter taste and low pitch and between sweet taste and high pitch (Watson & Gunther, 2017). However, the results differed according to the type of instrument (i.e., significant differences in the choice of pitch according to tastes were found for trombone tones but not for clarinet). Knöferle et al. (2015) also found support for the bitter-low pitch correspondence, as well as an association between high pitch and both sweetness and sourness. This experiment also found evidence for crossmodal mappings between basic taste words and other auditory parameters, such as roughness, discontinuity, tempo, and sharpness.

Wang et al. (2016) reported an effect of taste on pitch choice, with sweetness and sourness again leading to the choice of higher-pitched notes. In this study, three concentration levels were used for each taste solution to test the effect of taste intensity on pitch and volume choice. While the type of tastant was influential for pitch choice, taste intensity significantly impacted the choice of sound volume. A recent study with Chinese students also found a main effect of taste on pitch choice, with bitterness and saltiness being significantly associated with low-pitched sounds (Qi et al., 2020). An association between taste and choice of instruments was also observed, as participants showed a preference for specific instrument-taste pairings over others with significantly above-chance probability.

Three other studies reported pitch-taste correspondences in real foods and beverages. Reinoso-Carvalho, Wang, de Causmaecker, et al. (2016) found that individuals tended to match bitter beers to lower-pitched sounds and sweeter beers to higher-pitched sounds. In another study, chocolates that were rated as sweeter were matched to a higher pitch, whereas the bitter were matched to a lower pitch (Crisinel & Spence, 2012). Crisinel and Spence (2011) found that participants matched the taste of flavored milk solutions to different musical instruments. This study also reported an effect of the flavor (e.g., vanilla, lemon) of milk solutions on pitch choice, although no results were reported specifically for basic taste ratings.

One indirect source of evidence for taste-pitch correspondence comes from an experiment where participants saw images of different packages after a high- or low-pitched sound (Velasco et al., 2014). When asked if the package was more appropriate for a sweet or sour product, high-pitched sounds were more frequently associated with sourness, whereas low-pitched sounds were more associated with sweetness. These results align with previous findings observed for sourness but not sweetness, which is usually associated with a higher pitch.

Overall, across nine studies that specify the direction of association between sound pitch and sweet taste, six reported a link with higher-pitched sounds, one reported an association with a lower pitch, and two reported no association. In contrast, bitterness was associated with a lower pitch in five studies, whereas three reported no association. For sourness, five studies reported a link with a higher pitch, while two reported no association. The findings regarding saltiness are mixed, with five studies reporting null results, one reporting an association with a higher pitch and one with a lower pitch. Umami was associated with a lower pitch across two studies, and one reported no association. In the remaining studies where pitch-sound associations were reported (*n* = 3), only information regarding relative (pairwise) differences was available.

[Table 2]

#### 3.2.2. Words, nonwords, and speech sounds

The bouba-kiki effect is a popular example of sound symbolism, that is, the ability to associate a speech sound with no semantic meaning (e.g., “Bouba”) to a certain attribute in another sensory modality (e.g., rounded shapes) (Ramachandran &

Hubbard, 2001). In the case of taste, Gallace et al. (2011) found evidence of regularity in the association between different foods and different nonwords. For instance, salty potato crisps were significantly more “Takete” than brie cheese, which is more “Maluma”. However, the taste ratings of these foods (e.g., salty/sweet) were not significantly correlated with the choice of nonwords. Crisinel, Jones, and Spence (2012), on the other hand, found evidence of systematic preferences for certain nonwords in response to basic taste solutions. In this experiment, participants were instructed to rate each aqueous solution using scales anchored by nonword pairs, such as “Lula-Ruki”, “Maluma-Takete”, “Bobolo-Decter”, and “Bouba-Kiki”. Saltiness led to significant differences from the midpoint of all nonword scales (i.e., more “Ruki”, “Takete”, “Decter”, and “Kiki”), while sweetness was more strongly associated with “Maluma” and “Lula” than “Takete” and “Ruki”, respectively. The “Bobolo-Decter” scale was the only one to detect significant preferences in response to both the citric acid solution (i.e., sourness) and the caffeine solution (i.e., bitterness). Ngo et al. (2013) followed a similar approach in two experiments with British and Colombian participants. The authors concluded that more “rounded” nonwords (such as “Bouba”) and low-pitched sounds share some form of crossmodal correspondence with exotic juices that are rated as being sweet and low in sourness. However, there were also some differences between samples, such as an association between passion fruit and sharp sounds for the British but not for Colombian participants. Overall, research with nonwords suggests a tendency to associate sweetness with rounder sounds (e.g., “Maluma”) and sourness and saltiness with sharper sounds, although results seem to depend on the nonword being tested.

The effects of sound symbolism are particularly evident in cultures where certain words reflect underlying associations between speech sounds and sensory attributes. For instance, instead of employing adjectives, Japanese speakers may opt for a sound symbolic expression like “mofu-mofu” to describe the sensory experience of a warm, soft blanket (Sakamoto & Watanabe, 2013). Sakamoto and Watanabe (2016) tested the crossmodal correspondence between speech sounds and tastes by examining the spontaneous sound symbolic words (SSW) produced by native Japanese speakers when tasting drinks such as tea and coffee. The authors analyzed the phonemes in the first syllable of each SSW and found several correlations with taste ratings. For instance, participants resorted to phonemes like /sh/ and /zy/ in response to sweet tastes, /g/, /d/, /z/, and /e/ in response to salty and /d/, /z/, and /e/ in response to sour. Interestingly, these phonemes are unrelated to the Japanese adjectives for those tastes, except in the case of bitterness, in which SSW shared similar sounds with the word “nigai” (i.e., bitter), namely /n/, /z/, and /i/. In another set of studies with Japanese participants (Motoki et al., 2020), fictitious brand names were created by systematically manipulating the type of vowels (front/back) and consonants (fricative/stop and voiced/voiceless). The four experiments suggested regular patterns of association between the brand names and expectations towards the products’ taste. In the case of sweetness, three main effects were found, indicating that front (vs. back) vowels, fricative (vs. stop) consonants, and voiceless (vs. voiced) consonants increased the expected sweetness. In contrast, stop (vs. fricative) and voiced (vs. voiceless) consonants were more associated with bitterness. Saltiness was associated with voiced (vs. voiceless) and stop (vs. fricative) consonants in two studies, whereas sourness presented a less clear pattern of association. Pathak et al. (2020) and Pathak and Calvert (2021) later extended these findings to vowel length, showing that people expect words containing long vowels to be associated with sweetness. These findings were consistent across studies, either in evaluation tasks (e.g., asking participants to rate the “sweetness” of a hypothetical brand name) or in free-choice tasks (e.g., having participants invent their own brand names for a sweet product).

Simner et al. (2010) asked participants to match four tastants (sweet, sour, bitter, salty) with sounds varying in several speech-related attributes and spectral balance. Increasing concentration of tastants corresponded to higher levels of all sound attributes (vowel height, vowel front/backness, voice discontinuity, and spectral balance). Significant main effects were also observed for taste type. For instance, sweet tastes were associated with lower vowel height compared with the bitter, salty, and sour tastes. The sweet taste was also associated with more continuous vowels than the bitter and sour tastes and lower spectral balance compared with the sour taste. Finally, three experiments have shown that basic tastes may also be associated with voice qualities (Motoki, Pathak, & Spence, 2022). For instance, falsetto voices were matched more strongly with sweetness, whereas creaky voices were more associated with bitterness.

In terms of the associations between speech sounds and tastes, results were particularly consistent for the sweet taste. This taste category seems to be well characterized by long vowels (eight out of eight studies), voiceless consonants (four out of four studies), fricative and front vowels (three out of four studies), and falsetto voices (three out of three studies). The other tastes presented less consistent patterns of association across studies, although some tendencies may also be observed, such as voiced consonants for bitterness or modal voices for umami. These findings are summarized in Table 3.

[Table 3]

**3.2.3. Music and soundtracks**

The use of musical expressions like the Italian term “dolce” (i.e., to play in a slow, gentle manner) suggests that certain music attributes may be intuitively associated with taste. To test if such vocabulary is conceptually meaningful rather than accidental, Mesz et al. (2011) asked trained musicians to improvise in accordance with taste words (sweet, salty, sour, and bitter). The resulting improvisations revealed consistent patterns of articulation, duration, loudness, gradus, and dissonance. When these improvisations were later presented to non-musical experts, the underlying taste was guessed with significantly above change accuracy. Based on the patterns of association between taste and certain attributes of music composition, Mesz et al. (2012) developed a composition algorithm and found that participants were able to decode the taste associated with each algorithmically generated music with above-chance accuracy. More recently, Wang et al. (2021) contributed to understanding the auditory attributes of salty soundtracks. In this study, the authors found that emotional (negative valence, high arousal, minor mode) and other sonic attributes (long decay, high roughness, and regular rhythm) were systematically associated with the perception of “saltiness” in music.

While several taste-inspired musical compositions have been created in the past, they were usually tested independently. In an effort to compare soundtracks produced by different researchers and designers, Wang et al. (2015) compiled 24 soundtracks used in previous studies. The results showed that the different soundtracks elicited different taste associations. Of the 24 soundtracks, 21 led to a significant association with one particular taste, and in 14 of them, the most chosen taste word was the one intended by the composer. Sweet soundtracks were the most easily decoded by participants (56.9% “correct” associations), followed by salty (44.4%), sour (41.7%), and bitter soundtracks (31.4%). Following a different approach, the Taste and Affect Music Database (Guedes et al., 2022) provides subjective ratings and basic taste correspondences for a set of 100 music excerpts of different moods and genres. Overall, all four basic taste categories achieved basic taste correspondences above the chance level (25%), suggesting that music may communicate gustatory attributes even when not composed with that purpose in mind. Similar to Wang et al. (2015), sweetness was the most commonly identified taste, followed by bitterness, saltiness, and sourness.

To further explore the influence of sound attributes on taste-sound associations, Guetta and Loui (2017) tested the same violin melody, played in different styles, inspired by four basic tastes (sweet, sour, salty, bitter). Participants were able to match the sound clips to taste words with above-chance accuracy. In another experiment described in this same paper, participants correctly matched each melody style with a corresponding chocolate sample (e.g., matching the sweet melody with the sweeter chocolate). When separately comparing match rates for each taste group, only the sweet category resulted in significantly above-chance performance. In a field study, visitors of a science fair who were challenged to freely associate taste words with several musical pieces revealed a significant preference for words like “chocolate” and “tasty” when exposed to putatively sweet music, while sour music elicited associations with words like “fruits” or “sour” (Kontukoski et al., 2015). Moreover, when these participants were asked to mix different ingredients to create a drink to match the music, differences in the drinks' sugar and acid content were found between the “sour” and “sweet” music conditions.

The associations between soundtracks and taste may also have the potential to influence choice behavior. Individuals exposed to a salty soundtrack were more likely to select pictures of salty (vs. sweet) foods when asked to indicate what food they would rather eat. The opposite pattern was found for sweet soundtracks (Padulo et al., 2020). Sweet and salty soundtracks also led to higher visual fixation times in congruent compared to incongruent foods (i.e., fixation on sweet foods was longer when listening to sweet soundtracks; Peng-Li et al., 2020). Music genre may also contribute to shaping food choices, as evidenced in two experiments where jazz and classical music (vs. hip hop and rock/metal) were associated with a higher preference for healthy savory foods. Classical music also led to a higher preference for sweet foods (healthy or indulgent) compared to all other genres (Motoki, Takahashi, et al., 2022).

Among the articles included in the current review that empirically tested the effect of sound in shaping taste perception, six reported results of pilot studies validating the associations between music or soundtracks and basic tastes. For instance, Reinoso-Carvalho, Van Ee, Rychtarikova, Touhafi, Steenhaut, Persoone, Spence, et al. (2015) reported the process of composing three soundtracks to match chocolates with different bitter-to-sweet profiles. Wang et al. (2017) explored the concepts of sweetness and spiciness in two soundtracks composed to enhance both gustatory qualities in a new restaurant dish. In another pilot study with two pairs of contrasting (soft/hard and positive/negative mood) soundtracks, Reinoso-Carvalho, Gunn, ter Horst, et al. (2020) found that the positive soundtrack elicited the sweetest ratings and the negative soundtrack was more strongly associated with bitterness.

Wang and Spence (2016) created consonant and dissonant[2] versions of two short melodies and, as expected, observed that they were associated with sweetness and with sourness, respectively. Another experiment aimed at developing and testing soundtracks to elicit smoothness/creaminess and roughness, respectively (Reinoso-Carvalho, Wang, et al., 2017). Besides being rated as creamier, the smooth/creamy (vs. rough) soundtrack was also rated as sweeter. Finally, Hauck and Hecht (2019) tested how classical music pieces evoked taste and other sensory associations. This pilot test resulted in the selection of two pieces (Alban Berg's "Three pieces for orchestra" and Tchaikovsky's "Waltz of the Flowers") that differed significantly in sensory associations, including sweetness, sourness, and bitterness.

Table 4 summarizes the current studies of music-taste correspondences. Notwithstanding the diversity of approaches, the reviewed studies seem consistent in showing that sounds and tastes were matched by participants in a nonarbitrary fashion. The evidence seems stronger for the sweet taste, not only in terms of the studies reporting easier recognition of this sensation in comparison with others (e.g., Guedes et al., 2022; Knöferle et al., 2015; Wang et al., 2015) but also in terms of the relative

---

[2] Consonance is typically associated with stability and pleasantness, while dissonance communicates uneasiness and a need of resolution (Lahdelma & Eerola, 2020)

predominance of studies targeting this taste (*n* = 21), compared to saltiness (*n* = 15), sourness (*n* = 14), bitterness (*n* = 14) and umami (*n* = 0).

[Table 4]

**3.3. Modulating basic taste perception with auditory stimuli**

**3.3.1. Music**

There were 14 studies in our sample testing the effect of music on the perceived taste of foods and beverages, either alone (*n* = 12) or in comparison with other sonic stimuli (*n* = 2). Most of these focused on beverages with alcoholic content (*n* = 10), and the remaining studies focused on sweet foods, such as ice cream and chocolate.

In the case of alcoholic beverages, two studies in wine-tasting settings described a main effect of classical music on the sweetness rating of wines compared to tasting in silence (Spence et al., 2013), as well as differences in perceived acidity profiles when listening to a classical music piece that was selected to match the sensory attributes of the wine, compared to tasting in silence (Wang & Spence, 2015). An exploratory study found that a white chardonnay wine was perceived as sweeter when a piece of classical music was played compared to silence, but no differences in taste perception were found when compared to when a pop song was played (De Luca et al., 2019). Using the Temporal Dominance of Sensations (TDS) method, Wang et al. (2019) found that the presence of music led to significantly different dominance profiles for sensory attributes of a red wine, particularly for the bitter taste. In Hauck and Hecht's (2019) study, music (Berg vs. Tchaikovsky) influenced the perceived saltiness and sourness of wines and aqueous solutions, but not sweetness and bitterness.

The valence of music affected the sweetness and bitterness (but not sourness) ratings of beers (Reinoso-Carvalho et al., 2019). Pleasant (vs. unpleasant) music enhanced sweetness perception and decreased bitterness ratings. In another study, the mere presence of music seemed to enhance sweetness ratings of alcoholic beverages compared to control and other distracting conditions (Stafford et al., 2012). Finally, Reinoso-Carvalho, Velasco, et al. (2016) found that tasting a beer with a specific musical accompaniment was considered more enjoyable and led to higher sourness ratings than in silence.

Three studies in our sample report the effects of emotionally laden music on the taste of chocolate gelati using the TDS method. Two of these studies suggest that listening to liked music is associated with longer dominance of sweetness, while disliked music elicits a longer dominance of the bitter taste (Kantono et al., 2016, 2019). As in real-world settings, background music is often intertwined with ambient sounds, Lin et al. (2022) tested the effects of combinations of liked and disliked music with pleasant and unpleasant sounds. Once again, preferred music led to higher dominance of sweetness, whereas bitterness was more evident in unpleasant sound conditions (i.e., disliked music, unpleasant sounds, and a combination of both). Another study examined the influence of sonic stimuli in a more ecologically valid setting (Reinoso-Carvalho, Van Ee, Rychtarikova, Touhafi, Steenhaut, Persoone, & Spence, 2015). In this case, participants were customers of a chocolate shop who evaluated a chocolate praline while listening to a customized musical accompaniment or the shop's own soundscape. In this context, listening to music did not significantly impact taste ratings.

A summary of the findings is provided in Table 5. Given the large proportion of studies with alcoholic drinks (beer and wine), it is not surprising that there were more studies targeting sweet ($n$ = 13), sour ($n$ = 11), and bitter ($n$ = 11) tastes compared to saltiness ($n$ = 3), and umami ($n$ = 1). Overall, differences in bitterness profiles were more consistent, with reported differences in seven studies, whereas changes in sweetness perception were also reported in seven studies (although with a higher number of null results). In contrast, only three studies reported significant differences in sourness.

[Table 5]

### 3.3.2. Soundtracks

Sixteen studies in our sample tested the effect of soundtracks on sensory perception of foods and drinks. Generally, these auditory stimuli were produced by researchers and sound designers to match or enhance tastes or flavors. One of the first studies to empirically test this possibility showed that a bittersweet cinder toffee was perceived as tasting significantly more bitter when participants listened to a bitter soundtrack compared to a sweet soundtrack (Crisinel, Cosser, et al., 2012). Höchenberger and Ohla (2019) conducted two experiments to replicate these results. In the first experiment, the soundtracks influenced bitterness-sweetness ratings in the expected direction (i.e., the samples were perceived as sweeter under the sweet sound condition vs. the bitter sound condition). However, when a "no sound" condition was included in the statistical analysis, the effect of sound was no longer significant. Moreover, when independent taste rating scales were used in Experiment 2 (for sweet, bitter, salty, and sour) instead of a bipolar bittersweet scale (Experiment 1), the effect of sound was no longer significant.

Across three experiments, Reinoso-Carvalho, Wang, Van Ee, et al. (2016) found that sensory (i.e., sweet, bitter, and sour) properties of soundtracks may transfer to the sensory perception of beers. Interestingly, the modulatory effects of sounds for two of the beers were significant when comparing pairs of soundtracks differing in taste correspondences but not when comparing soundtracks and silence conditions. Conversely, for the remaining beer, the soundtracks led to different taste ratings only when comparing soundtrack and silence conditions.

One study testing the effect of a “creamy” (vs. “rough”) soundtrack on the taste of chocolate showed that sound could modulate perceived creaminess, as well as the perceived sweetness of the samples (Reinoso-Carvalho, Wang, et al., 2017). Another study with “sweet” and “bitter” soundtracks found differences in taste ratings for bitter chocolate but not for sweet chocolate (Reinoso-Carvalho, Van Ee, Rychtarikova, Touhafi, Steenhaut, Persoone, Spence, et al., 2015). Interestingly, the modulatory effects were more evident when comparisons were made using each participant’s individual correspondences between the soundtracks and tastes (i.e., pairing the sweet chocolate with the soundtrack that the participant selected as “sweet” in a pretest instead of the group average). Furthermore, the time of presentation of the soundtracks also leads to different modulatory effects. Specifically, while presenting the soundtracks before or during tasting significantly modulated taste ratings, this effect did not hold when a soundtrack was presented after tasting (Wang et al., 2020).

Although most soundtracks were created with taste associations in mind, two recent studies have highlighted the role of emotional associations in eliciting changes in taste perception. One study found stronger evidence for an effect of positive (vs. negative) soundtracks on the perceived sweetness of chocolates than for soft (vs. hard) soundtracks (Reinoso-Carvalho, Gunn, ter Horst, et al., 2020). Similar results were obtained when comparing different cultures, namely, between Japanese and Colombian participants (Reinoso-Carvalho, Gunn, Molina, et al., 2020).

Two additional studies in our sample tested the effect of consonant and dissonant soundtracks on the taste of juices (Wang & Spence, 2015; 2018). Both studies suggested that consonant soundtracks increased perceived sweetness, being also perceived as more pleasant than dissonant soundtracks. The study by Wang and Spence (2018) included both visual and auditory stimuli and found that valence (positive vs. negative) was more determinant than sensory modality (visual vs. auditory) in taste perception differences.

One study conducted in a naturalistic eating setting tested the effect of spicy (high pitch, fast tempo, and a distorted timbre) and sweet (high pitch, legato articulation, and consonant harmony) soundtracks on the perception of a restaurant dish (Wang et al., 2017). Here, although participants reported expecting the dish to be spicier when listening to a spicy soundtrack, actual taste (i.e., sweetness) ratings were unaffected by the soundtrack.

Table 6 summarizes the evidence in this category of studies. Once again, sweet ($n$ = 16) and bitter ($n$ = 12) sensations were examined more frequently, followed by sourness ($n$ = 8) and salty ($n$ = 1). Most studies reported significant differences in sweet ($n$ = 13) and bitter ($n$ = 9) attributes, whereas differences in sourness were reported in five studies.

[Table 6]

**3.3.3. Noise, tones, and soundscapes**

Participants in Burzynska's et al. (2019) study tasted two wines while in silence and while listening to a low (100 Hz) or high (1000 Hz) frequency sine wave tone. While the sound condition had a relative impact on some of the wine's attributes (e.g., higher aromatic intensity with the lower frequency tone), the only dependent measure related to basic tastes (acidity) was unaffected by the experimental manipulation.

Three studies examined the effect of noisy stimuli on taste perception. Woods et al. (2011) showed that food was rated significantly less sweet and salty under loud (vs. quiet) white noise. Yan and Dando (2015) tried to reproduce the auditory experience of a noisy airline cabin and exposed participants to the simulated noise for 30 minutes before and during the tasting of five aqueous solutions. Although the noise condition did not impact ratings for salty, bitter, and sour tastes, sweet taste perception was significantly suppressed by noise, regardless of concentration levels. In contrast, umami was perceived as more pronounced under the loud noise condition. Since it is common to conduct research on gustatory function with fMRI, Lorentzen et al. (2021) tested whether the noise of an fMRI scanner could represent a confound in these studies by impacting taste identification, intensity, or hedonic ratings. According to these results, no significant impact of noise (compared to silence) was found for any taste variables, regardless of the taste in question (sweet, salty, sour, or bitter).

Regarding soundscapes, Bravo-Moncayo et al. (2020) found that the taste of coffee was evaluated as more bitter and with a more intense aroma when listening to a filtered (less loud) version of a noisy food court soundscape, compared to an unfiltered (louder) background noise. Two studies in our review tested the effect of soundscapes on continuous dynamic measures of sensations using two different methods. In the Temporal Dominance of Sensations (TDS) method, participants select the most dominant sensory attribute at a given moment in time, while in the Temporal Check-All-That-Apply (TCATA) method, participants may choose more than one attribute simultaneously. One study employed the former method to assess differences in perception of ice-cream flavor while listening to a café soundscape or the same soundscape overlaid with bird, forest, and machine sounds (Xu, Hamid, Shepherd, Kantono, Reay, et al., 2019). These soundscapes elicited different taste and flavor trajectories and emotional experiences. For instance, when listening to the café-forest soundscape, sweetness was more dominant and for a longer duration at the start of consumption than the café-bird and café-machine soundscapes. The bitter taste was perceived as more dominant at the end of the consumption period when listening to the café control and café-machine soundscapes as compared to the café-birds and café-forest soundscapes. Additionally, a correlation was observed between evoked emotions and taste and flavor ratings, specifically between unpleasant emotions and bitterness dominance and between pleasant emotions and sweetness dominance. A similar result was obtained in a study using the TCATA method in which participants identified sweetness more frequently in positive valence conditions, such as a park or a café soundscape (Lin et al., 2019). In contrast, bitterness was more pronounced in high arousal, low valence conditions, such as bar, fast food, and food court soundscapes.

A summary of findings may be found in Table 7. Overall, the results seem to indicate some inconsistency regarding the influence of noise on taste perception, which could be attributable to methodological differences (e.g., stimuli characteristics). As far as soundscapes are concerned, differences in sweetness and bitterness perception were reported in accordance with the emotional tone of the stimuli.

[Table 7]

**4. Discussion**

The present work provides an overview of the current evidence on the relationship between audition and taste. Although not always seen as such, audition has been growingly acknowledged as a relevant sensory modality for taste (and flavor) perception (Spence, 2016). According to these findings, there are at least two ways by which audition and taste interweave.

First, it seems that there are regularities in how people match sounds to basic tastes. The scope of sound-taste associations seems to range from more elementary auditory stimuli, such as musical notes or speech sounds, to more complex stimuli, such as music. One important consequence of the multimodal nature of human perception is that stimulating one sensory modality may contribute to modulating the perception of information in another modality (Evans & Treisman, 2010). That seems to be the case with taste perception as well. Thus, a second complementary line of inquiry is now shedding light on how the crossmodal correspondences between sounds and tastes may result in different sensory experiences with foods and drinks. According to the reviewed literature, there are several ways whereby sound may influence taste perception. The most common strategies involve musical stimuli, either something one could expect to hear in a restaurant, such as pop, jazz, or classical music (e.g., Hauck

& Hecht, 2019) to less familiar, bespoke soundtracks that are intentionally composed to convey basic taste attributes (e.g., Wang et al., 2015).

In most eating settings, people are exposed to other auditory stimuli as well. Ambient noise in places like canteens or food courts is known to have a negative impact on customers' experience (Spence, 2014). However, it's still unclear how it may affect taste perception, as is evidenced by the mixed findings reviewed here. Indeed, the answer may depend on the type of noise as well as the basic taste in question. Research with soundscapes reinforces the need to pay attention to environmental sounds in eating settings, as these may shape how tastes are perceived. Notably, these studies (but also those with musical stimuli) seem to suggest that emotional/affective dimensions may be relevant to understanding the implications of sound for the taste experience. Particularly, pleasant emotions and affect may be significant factors for enhancing sweetness perception, whereas unpleasant emotions and affect seem to emphasize bitterness perception (Lin et al., 2019; Xu, Hamid, Shepherd, Kantono, Reay, et al., 2019).

4.1. Psychoacoustic attributes of taste-evoking stimuli

One quintessential challenge in crossmodal research is determining what attributes make up a sweet or a salty sound which can be approached by asking professional musicians to create musical compositions based on taste attributes (e.g., Mesz's et al. (2011). Another – more common – method is based on laypeople's intuitive matching of sounds to tastes (i.e., how sounds varying in sonic attributes such as pitch or loudness are differently matched to tastes).

The most commonly studied acoustic attribute in this regard is pitch. Previous research has shown that high-pitched sounds are crossmodally associated with attributes like brightness (Marks & Pierce, 1989), higher spatial position (Rusconi et al., 2006), or a small size (Bien et al., 2012; Fernández-Prieto et al., 2015). Regarding sound-taste associations, there seems to be a fair degree of consistency in how individuals match the sound frequency to some of the basic tastes. The majority of the reviewed studies suggest an association between higher-pitched sounds and sweet and sour tastes, whereas lower-pitched sounds are more frequently associated with bitterness. It should be noted, however, that this pattern may depend, for instance, on the type of instrument under analysis (Watson & Gunther, 2017). On the other hand, the taste-pitch patterns seemed to hold regardless of whether taste stimuli were presented as food words (Crisinel & Spence, 2009; 2010; Qi et al., 2020) or actual gustatory stimuli (Crisinel & Spence, 2011; Reinoso-Carvalho, Wang, de Causmaecker, et al., 2016).

Besides pitch, other acoustic attributes may be manipulated to mimic basic tastes: For instance, Mesz et al. (2011) found that musicians systematically manipulate attributes like duration, dissonance, articulation, and loudness. Knöferle et al. (2015) later showed that individuals with no musical training are also able to match tastes to sounds with different roughness, discontinuity, tempo, and sharpness. Wang et al. (2015) provided an overview of the taste-evoking soundtracks released to date and systematized some of the main acoustic attributes (pitch, instruments, harmony, and articulation) of bitter, salty, sour, and sweet soundtracks. Although some attributes were consistently associated with tastes (for example, sweetness and consonance), others were less stable across taste categories (e.g., bitter soundtracks may be consonant or dissonant, low-pitched or medium-pitched). For that reason, defining the sonic attributes of taste-evoking stimuli remains a challenge.

### 4.2. Understanding taste-sound correspondences

Understanding the acoustic mechanics of a sweet or salty sound may help understand how taste-sound correspondences unfold. However, it is also important to question why gustation and audition interweave in predictable ways. Most studies reviewed here do not explicitly seek to test conceptual hypotheses for sound-taste associations or sonic seasoning effects. Nonetheless, we should briefly mention some likely conceptual explanations for the findings reviewed here.

At least four possible mechanisms underlying taste-sound correspondences have been put forward, namely, statistical, intensity-matching, semantic, and hedonic hypotheses (Knöferle & Spence, 2012). These mechanisms are summarized in Table 8, alongside illustrative examples taken from this review. The statistical co-occurrence hypothesis posits that individuals learn to match attributes that tend to coexist frequently together in natural environments. For instance, color-taste associations (e.g., between red-colored products and sweetness) are thought to depend on previous knowledge that fruits tend to transition from colors at the green end of the spectrum to colors at the red end of the spectrum as they ripen (Foroni et al., 2016; Maga, 1974). In the case of sound-taste associations, such associative learning may not be as straightforward since it is not easy to imagine the natural sound of bitter food. Still, it has been proposed that innate orofacial gestures in response to pleasant and unpleasant tastes may have associated speech sounds. For instance, protruding the tongue outward and upward in response to a pleasant taste could result in higher frequency sounds when exhaling (Knöferle & Spence, 2012). Motoki et al. (2020) provide another illustration of a statistical mechanism. In their study of brand names and taste expectations, the authors argue that people may learn to associate certain food brand names with tastes based on real-world patterns (for example, based on the common occurrence of brand names containing stop and voiced consonants for salty products).

The intensity matching hypothesis posits that crossmodal correspondences may stem from mappings based on magnitude judgments. The so-called “prothetic” (or magnitude-related) hypothesis is in line with Wang’s et al. (2016) findings, where participants consistently matched higher sound loudness to higher concentrations of tastants. Although intensity-matching seemed to explain (or, at least, show consistency with) the results observed for sound volume, the same was not the case with pitch. While the expected pattern (higher tastants concentration to higher pitch) was observed for some of the tastants, the opposite pattern was observed for others. Moreover, the results also depended on the measure in question (i.e., objective concentration measures or subjective intensity ratings).

On a different note, the fact that pitch is verbally described as differing in height (i.e., “high or “low”) serves as an example of a semantic commonality in which the same language terms are applied to describing different perceptual phenomena. According to this view, an association between spatial position and sound frequency could be attributed to this commonality (Knöferle & Spence, 2012). In different languages, the word “sweet” is a common synesthetic metaphor for describing acoustic qualities, such as a soft voice or a delicate way of playing an instrument (Mesz et al., 2012). This association has been hypothesized to be involved in why some studies report a seemingly easier recognition of sweet taste attributes in music pieces compared with other taste categories (e.g., Guedes et al., 2022; Knöferle et al., 2015; Wang et al., 2015). However, semantic commonalities fail to explain why individuals are also able to match soundtracks to other basic tastes with above-chance accuracy.

The hedonic matching hypothesis is an alternative explanation of sound-taste correspondences (Knöferle & Spence, 2012). This proposition is based on the assumption that some tastes are globally deemed more pleasant than others. Indeed, the sweet taste is thought to be innately preferred due to its association with maternal milk. In contrast, the bitter taste is often rejected as a protective caution against possible toxicity (Ventura & Mennella, 2011). Although adults may show different taste preferences, it is possible that conceptually, such a heuristic could account for some taste-sound correspondences. It has been shown that the choice of instrument sounds is associated with pleasantness ratings (e.g., piano associated with pleasantness and sweetness; brass associated with unpleasantness and bitterness; Crisinel & Spence, 2010; Crisinel & Spence, 2012). While the choice of instrument was significantly affected by pleasantness, the same was not always true for the choice of pitch. As Crisinel and Spence (2012) argue, the lack of a consistent association with pitch may be due to participants' difficulty in agreeing on how pleasantness and pitch relate (i.e., some higher-pitched sounds may be more pleasant, whereas others may not). Guetta and Loui (2017) also found support for a hedonic matching hypothesis in music-taste matching, as individuals who judged a sweet chocolate ganache as more pleasant also rated the sweet soundtracks as more pleasant. In contrast, those who preferred a salty ganache also rated a salty soundtrack as more pleasant. Similarly, Guedes et al. (2022) found several correlations between music-taste correspondences and emotional and affective variables (e.g., soundtracks evaluated as sweet were more associated with pleasant emotions, whereas the opposite pattern was found for bitterness).

Although the aforementioned conceptual hypotheses may help explain why people associate stimuli pertaining to the auditory and gustatory modalities in some instances, there are also findings that do not fit perfectly in the existing conceptual frameworks. For example, Knöferle et al. (2015) argue that some crossmodal associations (e.g., sour taste and high tempo) are not easily framed in either of the existing working hypotheses and require further research and theoretical reflection.

### 4.3. Understanding sonic seasoning

Spence et al. (2019) have argued that the mere existence of a taste-sound association does not assure that a given auditory stimulus will result in changes in taste perception. Previously, five mechanisms underlying the modulatory role of audition in taste perception have been posited (Wang, 2017): response bias, sensory expectations, attention capture, physiological response, and emotion mediation (for a summary, see Table 8).

In one study reviewed here, participants tasted chocolate with a soundtrack delivered before, during, or after the tasting (Wang et al., 2020). Notably, taste ratings were affected when the soundtrack was delivered during the tasting but not after tasting (Experiment 1), suggesting that sonic seasoning corresponds to an actual perceptual effect rather than a mere response bias. No differences were observed when the soundtrack was played before (vs. during) tasting (Experiment 2), which seems consistent with the idea that auditory stimuli may alter taste expectations. These, in turn, may influence how food or drinks are perceived (see also Wang et al., 2017). These findings support a sensory expectations account but not a response bias explanation.

An attention-capture account has been used to explain findings in TDS studies. Wang et al. (2019) argue that crossmodal associations may direct individuals' attention toward congruent attributes of a gustatory stimulus, thus making it appear dominant in the flavor matrix. Alternatively, evidence emerging from other TDS studies shows that physiological responses may have an important role in modulating taste perception. Listening to music and soundscapes has been shown to elicit electrophysiological changes (e.g., skin conductance, heart rate, or respiratory rate), which seem to correlate with changes in flavor perception (Kantono et al., 2019; Xu, Hamid, Shepherd, Kantono, & Spence, 2019).

Based on the current evidence, physiological response mechanisms seem to depend on a second explanatory variable, namely, emotion. What the literature on TDS appears to suggest is that auditory stimuli change individuals' affective states, which, in turn, lead to a differentiated taste experience. Similar to what has been said about crossmodal correspondences, pleasant emotions are tendentially correlated with sweetness dominance, whereas unpleasant emotions highlight the bitter taste (Kantono et al., 2019; Lin et al., 2019; Xu, Hamid, Shepherd, Kantono, & Spence, 2019).

Similar to what has been described for TDS studies, research with "static" taste ratings also shows support for an effect of stimulus valence on sweetness ratings of juices and beers (Reinoso-Carvalho et al., 2019; Wang & Spence, 2018). Intriguingly enough, it has recently been argued that the emotional connotations of music may have a larger effect on taste ratings than crossmodal attributes alone. In two studies, participants tasted chocolate while listening to a soft (vs. hard) soundtrack or a positive (vs. negative) soundtrack (Reinoso-Carvalho, Gunn, Molina, et al., 2020; Reinoso-Carvalho, Gunn, ter Horst, et al., 2020). Although the pieces of music with crossmodal attributes were effective in modulating taste perception, the emotional music led to larger differences in taste ratings. One could argue that this result was perhaps due to the former pair of songs evoking textural rather than taste attributes more directly. Yet, the soft music had been previously shown to correspond to a sweet sensation, contrarily to the hard song, which was more associated with bitterness.

In a similar vein, the feelings evoked by auditory stimuli may also be transposed to the hedonic evaluation of foods and drinks. This process, usually known as "sensation transference", may help explain why liked music affects not only how participants evaluate the taste of foods but also how much they enjoy them (Guedes et al., 2023; Kantono et al., 2016; Reinoso-Carvalho, Wang, Van Ee, et al., 2016). Sensation transference has also been proposed as a potential mechanism to explain differences in taste ratings, as the pleasantness of a song may also be transferred to the evaluation of attributes such as sweetness or bitterness, which are positively and negatively associated with pleasantness, respectively (Reinoso-Carvalho, Touhafi, et al., 2016).

In future studies, it would be of interest to unravel the effects of emotional and crossmodal attributes, for instance, by testing pairs of sweet (vs. bitter) soundtracks with average valence ratings and positive (vs. negative) soundtracks controlling for crossmodal correspondences. Disentangling these effects would provide a valuable contribution to the field since knowing which attributes to look for in sound stimuli would be of major importance for informing future studies and interventions. If emotional connotations were solely responsible for leading sonic seasoning effects, one would perhaps expect emotional variables to fully mediate the effect of sound on taste ratings. However, that was not the case in Wang's et al. (2020) study, where soundtrack-taste associations fully mediated this relationship while valence did not. Further research would be needed to check whether the same findings would be obtained with different stimuli and/or other affective dimensions (e.g., arousal, dominance).

**Table 8**

*Summary of Putative Explanatory Mechanisms for Crossmodal Correspondences and Sonic Seasoning Effects*

| Proposed mechanism | General principle | Empirical example |
|---|---|---|
| *Crossmodal correspondences*[1] | | |
| Statistical co-occurrence | Individuals match features of stimuli from gustatory and auditory modalities based on natural co-occurrences | Association of brand names containing stop and voiced consonants with saltiness based on the frequency of real-world brands of salty products whose names contain those sounds (Motoki et al., 2020) |
| Intensity matching | Increases in the magnitude of one attribute of gustation are mapped onto the magnitude of one attribute of audition | Consistent matching between higher sound loudness and higher concentrations of tastants (e.g., sugar) (Wang et al., 2016) |
| Semantic matching | Correspondences stem from the shared use of a same term or concept to describe sensations from different sensory modalities | The easier association of sound attributes to the sweet taste could be due to the use of the word "sweet" to describe musical attributes (Guedes et al., 2022) |
| Hedonic matching | Individuals match sounds and tastes based on shared affective attributes, such as pleasantness | Matching of more pleasant instrument sounds (e.g., piano) with pleasant tastes (e.g., sweetness) (Crisinel & Spence, 2010; Crisinel & Spence, 2012) |
| *Sonic seasoning*[2] | | |
| Response bias | Differences in taste ratings could reflect influences in individuals' evaluation | Taste ratings are affected when a soundtrack is delivered during the |

| | | |
|---|---|---|
| | process rather than actual perceptual changes | tasting but not after (Wang et al., 2022 - Study 1) |
| Sensory expectations | Sounds induce sensory expectations which influence the subsequent tasting experience | No significant differences in taste ratings when a soundtrack is delivered before rather than during eating (Wang et al., 2022 - Study 2) |
| Attention capture | Auditory stimuli may direct attention to specific components of the taste experience | Sonic attributes commonly linked with sourness are associated with the perceived dominance of this taste in wine (Wang et al., 2019) |
| Physiological response | Sounds may induce physiological responses similar to those associated with the taste experience | Sounds elicit electrophysiological changes (e.g., heart rate), which correlate with changes in flavor perception (Kantono et al., 2019; Xu, Hamid, Shepherd, Kantono, & Spence, 2019) |
| Emotion mediation | Sounds influence participants' emotions, and these, in turn, shape the subsequent taste evaluation | Correlation between pleasant emotions induced by music/ soundscapes and sweetness perception (Lin et al., 2019; Reinoso-Carvalho et al., 2019). |

[1] See Knöferle & Spence (2012)

[2] See Wang (2017)

4.4. The nature-nurture of crossmodal perception and individual differences

There is a longstanding debate around the universal/innate or learned/culture-dependent nature of crossmodal associations. Support for an innatist outlook of crossmodal perception may come from comparative studies (e.g., Ludwig et al., 2011) as well as developmental research showing the early emergence of crossmodal mappings (Dolscheid et al., 2014; Walker et al., 2010). On the other hand, it has been shown that individuals are able to learn to integrate stimuli from different sensory modalities that were previously unrelated, such as stiffness and luminance (Ernst, 2007). Perhaps one of the most famous examples of multisensory links, the "Bouba-Kiki" effect, has been cited as a widely shared crossmodal association (Ćwiek et al., 2021). Interestingly, Bremner et al. (2013) found that the Himba, a remote population in Namibia with no written language, were able to match sounds and shapes just as

Western participants do. However, they matched sweeter chocolates with angular rather than rounded shapes, contrary to what is commonly found with Western participants.

Currently, cross-cultural comparisons of taste-sound associations are scarce. This is potentially problematic, considering that while some aspects of music perception seem cross-culturally shared, others are not (Knöferle et al., 2015; Qi et al., 2020). Similarly, if some crossmodal correspondences depend on linguistic features, such as shared concepts to describe stimuli in different modalities (e.g., the "sweet" word may be used to describe tastes, aromas, and sounds), the generalizability of findings could be constrained by language boundaries as well (Wang, 2017). In the case of the taste-pitch correspondence, Qi et al. (2020) found that Chinese participants behaved similarly to Western participants in matching higher pitch with sweet and sour tastes and lower pitch with bitter. Ngo et al. (2013) found that Colombian and British participants agreed on matching sweeter juices (e.g., pineapple, mango) with rounder sounds (e.g., "Bouba") and lower-pitched sounds (although no actual sound stimuli were used here, but rather a visual scale ranging from "low-pitched" to "high pitched"). Some specificities were also observed, such as British participants associating passion fruit juice with sharp sounds (e.g., "Takete"), while Colombian participants did not show a clear preference for rounded or sharp sounds.

Regarding the ability to decode taste attributes in more complex sound stimuli (e.g., soundtracks), Knöferle et al. (2015) found that participants from India and the USA performed with above-chance accuracy. Even so, US participants showed higher accuracy, particularly for the sweet taste. In another study, Chinese and Danish participants showed a similar ability to decode sweet and salty soundtracks (Peng-Li et al., 2020). When these soundtracks were played, there was a higher likelihood of selecting a congruent food (e.g., choosing a sweet food when listening to a sweet soundtrack), regardless of culture. However, some differences emerged when including a silence condition. Notably, Danish participants were more likely to choose salty foods when listening to the salty soundtrack (vs. silence), whereas no differences were observed for Chinese participants. When a sweet soundtrack was played (vs. silence), participants were more likely to choose a sweet food than a salty one. However, this difference was only marginal for Danish participants.

Recently, soundtracks previously associated with soft (vs. hard) sensations were shown to have a negligible effect on basic taste ratings of chocolates in a study with South Korean participants (Reinoso-Carvalho, Gunn, ter Horst, et al., 2020). Instead, emotional (i.e., positive vs. negative) music led to more pronounced sonic seasoning effects. Similar results were obtained in a direct cross-cultural comparison with Japanese and Colombian participants (Reinoso-Carvalho, Gunn, Molina, et al., 2020). This result was somewhat surprising, given that soft/hard soundtracks had been previously shown to affect the taste evaluation of chocolates in a study with Western participants (Reinoso-Carvalho, Wang, et al., 2017). However tempting as it may be to interpret these results in light of cultural differences, there are important methodological differences between these studies that should be considered. Notably, soundtrack condition varied within-participants in the latter study, whereas the former two followed a between-participants design. Moreover, Reinoso-Carvalho et al. (2017) contrasted only the two (soft and hard) soundtracks. Thus, it is yet unknown whether the sonic seasoning effects of crossmodally-corresponding music and emotional music differ between cultures.

Apart from cultural differences, a more empirical inquiry is needed regarding the role of individual differences in multisensory perception (for a recent discussion, see Spence, 2022). For instance, in one study, participants with no musical training were more likely to associate one music piece with sweetness, whereas musically trained individuals exhibited more frequent bitter taste associations (Wang et al., 2015). In the first large-scale norming study with sound-taste correspondences, some weak but significant correlations were observed between taste ratings and facets of musical sophistication (Guedes et al., 2022). In addition, associations between taste preferences and taste correspondences were also observed. For example, participants who reported a higher preference for sour foods provided more frequent sourness correspondences. In another study, taste sensitivity (namely, bitter sensitivity) was shown to influence taste-pitch mappings, as more sensitive individuals matched bitter solutions with lower-pitched sounds (although the small number of participants in this group advises caution in interpreting this result, Wang et al., 2016).

One important matter for debate is whether researchers should treat auditory stimuli nomothetically (i.e., assuming that sounds communicate crossmodal and/or emotional attributes similarly to all individuals) or ideographically (i.e., seeking to treat each individual differently according to their idiosyncratic evaluation of sounds). One of the reviewed studies suggests that this may be an important distinction since individual pairings (in this case of soundtracks and chocolates) may originate somewhat different sonic seasoning effects compared to pairings based on the sample's average (Reinoso-Carvalho, Van Ee, Rychtarikova, Touhafi, Steenhaut, Persoone, Spence, et al., 2015). Similarly, opting for self-selected sound stimuli may be a proper strategy for eliciting emotional effects. In effect, individual music preferences can modulate subjective reactions to musical stimuli (Fuentes-Sánchez et al., 2022). Hence, while some people may feel displeased by a heavy metal song, others may report an intensely enjoyable experience. Following this notion, previous studies have taken an idiographic approach to determine the most adequate stimuli based on individual selections of liked and disliked music genres (Kantono et al., 2016, 2019; Lin et al., 2022). From an applied perspective, the two approaches may serve complementary goals. While an idiographic strategy may best fit the purpose of individualized interventions, understanding shared mappings between sounds and tastes may be relevant for interventions in collective meal contexts, such as restaurants or canteens.

4.5. Limitations

The current paper is based on a systematic outlook on the growing body of literature regarding audition-gustation interactions. This approach aimed at ensuring a comprehensive understanding of the current state of knowledge while minimizing the risk of bias that hinders the reliability of traditional reviews. By defining the interactions between these two sensory modalities as the scope of the review, the resulting findings encompass a vast and diverse body of literature. While this may be advantageous in a relatively young field of inquiry, it also has limitations regarding the conclusions to be drawn from the reviewed data. The heterogeneity of data in a broad-scope review may challenge data synthesis and interpretation (Higgins et al., 2019) and can, perhaps, open more questions than it answers. In effect, future narrow-focused reviews are still needed to address more specific questions, such as those regarding the efficacy of specific sound-based interventions for changing taste-related outcomes. Although the current review protocol opened the spectrum of studies to be considered for analysis, there are still limitations, for instance, in terms of the lack of grey literature. The current protocol relied on central sources (i.e., databases) to obtain data, thus excluding findings that may have been disseminated outside the traditional scientific publishing channels (Mahood et al., 2014). In addition, the formulation of the search string may have resulted in a larger representation of crossmodal research based on auditory stimuli such as soundtracks or musical notes, compared to other stimuli such as phonemic sounds. Thus, it would be appropriate to complement the findings of this review with further readings of research in the field of sound symbolism (e.g., Athaide & Klink, 2012; Klink, 2000; Motoki et al., 2021; Pathak et al., 2021, 2022; Pathak & Calvert, 2020).

An additional limitation concerns the “file drawer problem” (otherwise known as publication bias), which refers to the null results that remain unpublished or unwritten (Franco et al., 2014). Thus, it is possible that some studies resulting in statistically insignificant results were not captured by the present review, which could result in an over-representation of “positive” results.

This search strategy allowed the identification of a significant body of peer-reviewed studies, which were summarized in a table of data extraction (provided as a supplement to this review). The data extraction strategy aimed at obtaining a synthesis of findings as well as information regarding methods, procedures, and other general characterization data. One less explored aspect of the data extraction protocol regards the theoretical underpinnings and conceptual hypothesis underlying the studies. Since many of the reviewed papers did not explicitly test explanatory mechanisms nor derive their hypothesis from known theoretical frameworks, the data extraction protocol did not include the assessment of theoretical/conceptual data. Future works may wish to bridge this gap, for instance, by integrating findings according to overarching theoretical frameworks (for an example, see Motoki et al., 2022).

4.6. Applicability and future directions

The current review aimed at summarizing the existing evidence on sound-taste associations as well as reflecting on the future directions of the field. As the interest in this area of research seems to be gaining momentum (as seen by the large slice of studies published in the past five years), it is perhaps the moment to seek to bring coherence to the field. For example, as a growing number of auditory stimuli becomes available, it is relevant to test them comprehensively, under similar circumstances, and based on comparable measurement scales (e.g., Guedes et al., 2022; Wang et al., 2015). Similarly, replication studies also contribute to the robustness of findings and the confidence in how to read them (Höchenberger & Ohla, 2019; Watson & Gunther, 2017). Currently, there is a great diversity in the use of measurement scales, as well as in research designs, which hinders the comparability of different studies. For example, in sonic seasoning studies, it is unclear whether comparisons are to be made between contrasting stimuli (e.g., a sweet vs. a bitter song) or between music/soundtracks and other sonic stimuli (e.g., white noise, silence). So far, the different options regarding what is to be compared with what have led to disparate findings. In some cases, including a silence condition instead of comparing pairs of contrasting soundtracks may be sufficient to overshadow the effects of music (Höchenberger & Ohla, 2019; Reinoso-Carvalho, Wang, van Ee, et al., 2016). It is also highly questionable whether two sweet soundtracks are as equivalent as to be used interchangeably. The fact that music differs in taste associations but also other conceptual dimensions should be taken into account when comparing findings, as two sweet soundtracks are hardly perceived alike. Validating stimulus sets and testing sonic stimuli across contexts will perhaps allow for greater comparability and, consequently, a more robust quantitative synthesis (such as meta-analyses). Similarly, the full reporting of results (including non-significant findings) and pre-registration of studies will undoubtedly contribute to a more coherent and robust body of evidence. Moreover, reporting effect sizes (or improving data availability to allow calculation) will be of most importance to future efforts toward providing quantitative syntheses of findings.

Another relevant challenge concerns the diversification of samples, given the widely discussed pitfalls of convenience sampling of undergraduate students that seems to plague the reproducibility of behavior research in general (Peterson & Merunka, 2014). Several studies reviewed here rely on samples of visitors or customers of shops, restaurants, or public events. Yet, it is unknown whether these provide an adequate approximation to the general population (e.g., in terms of sociodemographic characteristics, such as educational attainment, but also in engagement or openness to multisensory eating experiences). Likewise, while several of the reviewed studies were conducted in so-called real-world settings (e.g., stores, restaurants), experimental apparatus are still often artificial and lab-like in some respects (e.g., individual booths, use of headphones). Isolating sonic stimuli from the broader context is somewhat problematic since real eating environments are likely to involve a combination of music, speech sounds, and background noise (although, for an approximation, see Lin et al., 2022). Of particular importance is the prevalent use of headphones in most studies. It is likely that this feature of most sonic seasoning studies (not commonly shared in real-world situations) may be important to bring about the desired crossmodal effects, not only by directing attention to the auditory stimuli but also by evoking the notion that gustatory and auditory stimuli origin from the same location (i.e., inside the head, Spence et al., 2019). In sum, the question of ecological validity will be a likely challenge for future studies that will require an effort to bring research closer to real-life circumstances.

To date, there have been exciting applications of sonic stimuli to improve the multisensory experience with foods. One popular example, the dish "Sound of the Sea" by British chef Heston Blumenthal[3] illustrates how audition may be convened to design novel and memorable multisensory experiences in fine dining settings (Spence & Shankar, 2010). With growing digitalization and technological advancements, it also becomes possible to deliver accessible multisensory experiences at home. In the past years, a variety of applications have been released to enhance customers' experience of products like champagne, beer, or chocolate through sound (for a review, see Spence et al., 2019). Initiatives like these undoubtedly contribute to the enjoyment of the eating situation, which may be an important motivation for changing eating habits. As far as sonic seasoning is concerned, there is a sharp predominance of research on highly palatable products, including alcoholic beverages and sweet products. For other sensory modalities, multisensory strategies have been called upon as potential avenues for enhancing sensory acceptance and enjoyment of healthier products, such as sugar-reduced foods (Alcaire et al., 2017; Hidaka & Shimoda, 2014; Hutchings et al., 2019). It is feasible to hypothesize that sound stimuli may also contribute to

[3] The signature dish consists of sea products served in a shell containing an iPod. A seaside soundscape overlaying sounds of waves and seagulls is delivered through headphones.

mitigating the effects of sugar or salt reduction, both on sensory and hedonic grounds. Recently, Peng-li et al. (2021) demonstrated that music might be selectively engineered to communicate healthiness attributes and, consequently, shape healthier food choices. Similarly, background music and noise may systematically nudge participants toward healthier food choices (Biswas et al., 2019). It would be interesting to test whether crossmodally-corresponding background sounds in meal settings could also contribute to promoting healthier eating by means of their effect on taste perception. A recent experiment suggested that this may be a feasible endeavor for sugar consumption, as sweet music was shown to influence taste perception and the acceptance of products with lower sugar levels (Guedes et al., 2023).

## 5. Conclusions

The literature on the multisensory interactions between audition and taste perception is rapidly growing. In this work, we sought to provide an outlook of the current understanding of how sound and taste relate by systematically reviewing the existing body of literature. This review allowed the identification of a great variety of points of contact between the two sensory modalities. Two main lines of research were outlined, namely, showing that people systematically associate attributes from gustatory and auditory domains and that taste perception may be liable to the influence of auditory stimuli.

This review sought to synthesize the current evidence to inform future empirical inquiry. Practitioners interested in developing multisensory integration techniques to improve taste perception may also find this review useful for informing the scope and methodological approach of their work.

**Acknowledgments**

This research was financially supported by FCT - Fundação para a Ciência e a Tecnologia with a grant awarded to the first author (SFRH/BD/145929/2019). Part of this research was funded by Project LISBOA-01-0145-FEDER-028008, co-funded by the Lisboa 2020 Program, Portugal 2020, and European Union through FEDER funds and by national funds through the Fundação para a Ciência e Tecnologia.

The authors would like to thank Sofia Coelho for her support in the title and abstract screening stage.

**Declaration of interest**

None

**Data code and availability statements**

The extracted data is openly available at osf.io/t4r76.

**Author contributions**

David Guedes: Conceptualization, Investigation, Formal analysis, Methodology, Writing – original draft.

Margarida V. Garrido: Conceptualization, Methodology, Writing – review & editing.

Elsa Lamy: Conceptualization, Methodology, Writing – review & editing.

Bernardo P. Cavalheiro: Formal analysis.

Marília Prada: Conceptualization, Methodology, Writing – review & editing.

**References**

Alcaire, F., Antúnez, L., Vidal, L., Giménez, A., & Ares, G. (2017). Aroma-related cross-modal interactions for sugar reduction in milk desserts: Influence on consumer perception. *Food Research International*, *97*, 45–50. https://doi.org/10.1016/j.foodres.2017.02.019

Athaide, G. A., & Klink, R. R. (2012). Creating global brand names: The use of sound symbolism. *Journal of Global Marketing, 25*(4), 202–212. https://doi.org/10.1080/08911762.2012.744123

Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and Cognition*, *17*(3), 1016–1031. https://doi.org/10.1016/j.concog.2007.06.005

Bertelsen, A. S., Mielby, L. A., Alexi, N., Byrne, D. V, & Kidmose, U. (2020). Sweetness enhancement by aromas: Measured by descriptive sensory analysis and relative to reference scaling. *Chemical Senses*, *45*(4), 293–301. https://doi.org/10.1093/chemse/bjaa012

Bertelsen, A. S., Zeng, Y., Mielby, L. A., Sun, Y. X., Byrne, D. V., & Kidmose, U. (2021). Cross-modal effect of vanilla aroma on sweetness of different sweeteners among Chinese and Danish Consumers. *Food Quality and Preference*, *87*(April 2020), 104036. https://doi.org/10.1016/j.foodqual.2020.104036

Besnard, P., Passilly-Degrace, P., & Khan, N. A. (2015). Taste of fat: A sixth taste modality? *Physiological Reviews*, *96*(1), 151–176. https://doi.org/10.1152/physrev.00002.2015

Bien, N., ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size. Crossmodal binding in pitch-size synesthesia: A combined TMS, EEG and psychophysics study. *NeuroImage*, *59*(1), 663–672. https://doi.org/10.1016/j.neuroimage.2011.06.095

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. https://doi.org/10.1007/s11747-018-0583-8

Bravo-Moncayo, L., Reinoso-Carvalho, F., & Velasco, C. (2020). The effects of noise control in coffee tasting experiences. *Food Quality and Preference*, 104020. https://doi.org/10.1016/j.foodqual.2020.104020

Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J., & Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners. *Cognition*, *126*(2), 165–172. https://doi.org/10.1016/j.cognition.2012.09.007

Burzynska, J., Wang, Q. J., Spence, C., & Bastian, S. E. P. (2019). Taste the Bass: Low frequencies increase the perception of body and aromatic intensity in red wine. *Multisensory Research*, *32*(4–5), 429–454. https://doi.org/10.1163/22134808-20191406

Calvo, C., Salvador, A., & Fiszman, S. M. (2001). Influence of colour intensity on the perception of colour and sweetness in various fruit-flavoured yoghurts. *European Food Research and Technology*, *213*(2), 99–103. https://doi.org/10.1007/s002170100359

Chamoun, E., Liu, A. A., Duizer, L. M., Darlington, G., Duncan, A. M., Haines, J., & Ma, D. W. (2019). Taste sensitivity and taste preference measures are correlated in healthy young adults. *Chemical senses, 44*(2), 129-134. https://doi.org/10.1093/chemse/bjy082

Crisinel, A.-S., & Spence, C. (2009). Implicit association between basic tastes and pitch. *Neuroscience Letters*, *464*(1), 39–42. https://doi.org/10.1016/j.neulet.2009.08.016

Crisinel, A.-S., & Spence, C. (2011). Crossmodal associations between flavoured milk solutions and musical notes. *Acta Psychologica*, *138*, 155–161. https://doi.org/10.1016/j.actpsy.2011.05.018

Crisinel, A.-S., & Spence, C. (2012). The impact of pleasantness ratings on crossmodal associations between food samples and musical notes. *Food Quality and Preference*, *24*(1), 136–140. https://doi.org/10.1016/j.foodqual.2011.10.007

Crisinel, A.-S., Spence, C. (2010b). As bitter as a trombone: Synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Attention, Perception & Psychophysics*, *72*(7), 1994–2002. https://doi.org/10.3758/APP.72.7.1994

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. https://doi.org/10.1016/j.foodqual.2011.08.009

Crisinel, A.-S., Jones, S., & Spence, C. (2012). “The sweet taste of Maluma”: Crossmodal associations between tastes and words. *Chemosensory Perception*, *5*(3–4), 266–273. https://doi.org/10.1007/s12078-012-9133-9

Crisinel, A.-S., & Spence, C. (2010a). A sweet sound? Food names reveal implicit associations between taste and pitch. *Perception*, *39*(3), 417–425. https://doi.org/10.1068/p6574

Ćwiek, A., Fuchs, S., Draxler, C., Asu, E. L., Hiovain, K., Kawahara, S., Koutalidis, S., Lippus, P., Lupyan, G., Oh, G. E., & Paul, J. (2021). The bouba/kiki effect is robust across cultures and writing systems. *Philosophical Transactions of the Royal Society B*, *377*, 20200390. https://doi.org/10.1098/rstb.2020.0390

De Luca, M., Campo, R., & Lee, R. (2019). Mozart or pop music? Effects of background music on wine consumers. *International Journal of Wine Business Research*, *31*(3), 1751–1062. https://doi.org/10.1108/IJWBR-01-2018-0001

Dolscheid, S., Hunnius, S., Casasanto, D., & Majid, A. (2014). Prelinguistic infants are sensitive to space-pitch associations found across cultures. *Psychological Science*, *25*(6), 1256–1261. https://doi.org/10.1177/0956797614528521

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision*, *7*(5), 1–14. https://doi.org/10.1167/7.5.7

Evans, K. K., & Treisman, A. (2010). *Natural cross-modal mappings between visual and auditory features*. *Journal of vision, 10*, 1–12. https://doi.org/10.1167/10.1.6.Introduction

Fernández-Prieto, I., Navarra, J., & Pons, F. (2015). How big is this sound? Crossmodal association between pitch and size in infants. *Infant Behavior and Development*, *38*, 77–81. https://doi.org/10.1016/j.infbeh.2014.12.008

Foroni, F., Pergola, G., & Rumiati, R. I. (2016). Food color is in the eye of the beholder: The role of human trichromatic vision in food evaluation. *Scientific Reports*, *6*. https://doi.org/10.1038/srep37034

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science, 345*(6203), 1502-1505. https://doi.org/10.1126/science.1255484

Fuentes-Sánchez, N., Pastor, R., Eerola, T., Escrig, M. A., & Pastor, M. C. (2022). Musical preference but not familiarity influences subjective ratings and psychophysiological correlates of music-induced emotions. *Personality and Individual Differences*, *198*, 111828. https://doi.org/10.1016/j.paid.2022.111828

Gallace, A., Boschin, E., & Spence, C. (2011). On the taste of “Bouba” and “Kiki”: An exploration of word-food associations in neurologically normal participants. *Cognitive Neuroscience*, *2*(1), 34–46. https://doi.org/10.1080/17588928.2010.516820

Guedes, D., Prada, M., Garrido, M. V., & Lamy, E. (2022). The taste & affect music database: Subjective rating norms for a new set of musical stimuli. *Behavior Research Methods*. https://doi.org/10.3758/s13428-022-01862-z

Guedes, D., Prada, M., Lamy, E., & Garrido, M. V. (2023). Sweet music influences sensory and hedonic perception of food products with varying sugar levels. *Food Quality and Preference*, 104752. https://doi.org/10.1016/j.foodqual.2022.104752

Guéguen, N., Jacob, C., Le Guellec, H., Morineau, T., & Lourel, M. (2008). Sound level of environmental music and drinking behavior: A field experiment with beer drinkers. *Alcoholism: Clinical and Experimental Research*, *32*(10), 1795–1798. https://doi.org/10.1111/j.1530-0277.2008.00764.x

Guetta, R., & Loui, P. (2017). When music is salty: The crossmodal associations between sound and taste. *PLoS ONE*, *12*(3), e0173366. https://doi.org/10.1371/journal.pone.0173366

Hauck, P., & Hecht, H. (2019). Having a drink with Tchaikovsky: The crossmodal influence of background music on the taste of beverages. *Multisensory Research*, *32*(1), 1-24. https://doi.org/10.1163/22134808-20181321

Hidaka, S., & Shimoda, K. (2014). Investigation of the effects of color on judgments of sweetness using a taste adaptation method. *Multisensory Research*, *27*(3-4), 189-205. https://doi.org/10.1163/22134808-00002455

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). Cochrane handbook for systematic reviews of interventions. John Wiley & Sons.

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*(3), 267–275. https://doi.org/10.1002/jnr.24308

Holt-Hansen, K. (1976). Extraordinary experiences during cross modal perception. *Perceptual and Motor Skills*, *43*(3), 1023–1027. https://doi.org/10.2466/pms.1976.43.3f.1023

Holt-Hansen, K. (1968). Taste and pitch. *Perceptual and Motor Skills*, *27*(1), 59–68. https://doi.org/10.2466/pms.1968.27.1.59

Hutchings, S. C., Low, J. Y. Q., & Keast, R. S. J. (2019). Sugar reduction without compromising sensory perception. An impossible dream? *Critical Reviews in Food Science and Nutrition*, *59*(14), 2287–2307. https://doi.org/10.1080/10408398.2018.1450214

Kantono, K., Hamid, N., Shepherd, D., Lin, Y. H. T., Skiredj, S., & Carr, B. T. (2019). Emotional and electrophysiological measures correlate to flavour perception in the presence of music. *Physiology and Behavior*. https://doi.org/10.1016/j.physbeh.2018.11.012

Kantono, K., Hamid, N., Shepherd, D., Yoo, M. J. Y., Grazioli, G., & Carr, B. T. (2016). Listening to music can influence hedonic and sensory perceptions of gelati. *Appetite*, *100*, 244–255. https://doi.org/10.1016/j.appet.2016.02.143

Keast, R. S., & Costanzo, A. (2015). Is fat the sixth taste primary? Evidence and implications. *Flavour*, *4*(1), 1–7. https://doi.org/10.1186/2044-7248-4-5

Klink, R. R. (2000). Creating brand names with meaning: The use of sound symbolism. *Marketing Letters, 11*(1), 5–20.

Knöferle, K., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic Bulletin & Review*, *19*(6), 992–1006. https://doi.org/10.3758/s13423-012-0321-z

Knöferle, K. M., Woods, A., Kappler, F., & Spence, C. (2015). That sounds sweet: using cross-modal correspondences to communicate gustatory attributes. *Psychology & Marketing*, *32*(1), 107–120. https://doi.org/10.1002/mar.20766

Kontukoski, M., Luomala, H., Mesz, B., Sigman, M., Trevisan, M., Rotola-Pukkila, M., & Hopia, A. I. (2015). Sweet and sour: Music and taste associations. *Nutrition and Food Science*, *45*(3), 357–376. https://doi.org/10.1108/NFS-01-2015-0005

Lahdelma, I., & Eerola, T. (2020). Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-65615-8

Le Bon, S.-D., Payen, L., Prunier, L., Steffens, Y., Horoi, M., Vaira, L. A., Hopkins, C., Lechien, J. R., & Saussez, S. (2021). Making scents of loss of taste in COVID-19: Is self-reported loss of taste due to olfactory dysfunction? A prospective study using psychophysical testing. *International Forum of Allergy & Rhinology*, *11*(10), 1504–1507. https://doi.org/10.1002/alr.22815

Lin, Y. H. T., Hamid, N., Shepherd, D., Kantono, K., & Spence, C. (2019). Environmental sounds influence the multisensory perception of chocolate gelati. *Foods*, *8*(4). https://doi.org/10.3390/foods8040124

Lin, Y. H. T., Hamid, N., Shepherd, D., Kantono, K., & Spence, C. (2022). Musical and non-musical sounds influence the flavour perception of chocolate ice cream and emotional responses. *Foods*, *11*(12), 1784. https://doi.org/10.3390/foods11121784

Lorentzen, K. L., Nørgaard, H. J., Thrane, J. F., & Fjaeldstad, A. W. (2021). Effects of acoustic fMRI-noise on taste identification, liking, and intensity. *Current Research in Behavioral Sciences*, *2*, 100054. https://doi.org/10.1016/j.crbeha.2021.100054

Ludwig, V. U., Adachi, I., & Matsuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(51), 20661–20665. https://doi.org/10.1073/pnas.1112605108

Maga, J. A. (1974). Influence of color on taste thresholds. *Chemical Senses*, *1*(1), 115–119. https://doi.org/10.1093/chemse/1.1.115

Mahood, Q., Van Eerd, D., & Irvin, E. (2014). Searching for grey literature for systematic reviews: challenges and benefits. *Research synthesis methods, 5*(3), 221-234. https://doi.org/10.1002/jrsm.1106

Marks, L. E., & Pierce, J. B. (1989). *On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness*. *15*(3), 586–602. https://doi.org/10.1037//0096-1523.15.3.586

Mathiesen, S. L., Mielby, L. A., Byrne, D. V., & Wang, Q. J. (2020). Music to eat by: A systematic investigation of the relative importance of tempo and articulation on eating time. *Appetite*, *155*, 1–10. https://doi.org/10.1016/j.appet.2020.104801

McCarron, A., & Tierney, K. J. (1989). The effect of auditory stimulation on the consumption of soft drinks. *Appetite*, *13*(2), 155–159. https://doi.org/10.1016/0195-6663(89)90112-8

McElrea, H., & Standing, L. (1992). Fast music causes fast drinking. *Perceptual and Motor Skills*, *75*(2), 362. https://doi.org/10.2466/pms.1992.75.2.362

Mennella, J. A., & Bobowski, N. K. (2015). The sweetness and bitterness of childhood: Insights from basic research on taste preferences. *Physiology & Behavior*, *152*, 502–507. https://doi.org/10.1016/j.physbeh.2015.05.015

Mesz, B., Sigman, M., & Trevisan, M. A. (2012). A composition algorithm based on crossmodal taste-music correspondences. *Frontiers in Human Neuroscience*, *6*, 71. https://doi.org/10.3389/fnhum.2012.00071

Mesz, B., Trevisan, M. A., & Sigman, M. (2011). The taste of music. *Perception*, *40*(2), 209–219. https://doi.org/10.1068/p6801

Motoki, K., Park, J., Pathak, A., & Spence, C. (2021). Constructing healthy food names: On the sound symbolism of healthy food. *Food Quality and Preference, 90,* 104157. https://doi.org/10.1016/j.foodqual.2020.104157

Motoki, K., Park, J., Pathak, A., & Spence, C. (2022). The connotative meanings of sound symbolism in brand names: A conceptual framework. *Journal of Business Research, 150*, 365-373. https://doi.org/10.1016/j.jbusres.2022.06.013

Motoki, K., Pathak, A., & Spence, C. (2022). Tasting prosody: Crossmodal correspondences between voice quality and basic tastes. *Food Quality and Preference*, *100*, 104621. https://doi.org/10.1016/j.foodqual.2022.104621

Motoki, K., Saito, T., Park, J., Velasco, C., Spence, C., & Sugiura, M. (2020). Tasting names: Systematic investigations of taste-speech sounds associations. *Food Quality and Preference*, *80*. https://doi.org/10.1016/j.foodqual.2019.103801

Motoki, K., Takahashi, N., Velasco, C., & Spence, C. (2022). Is classical music sweeter than jazz? Crossmodal influences of background music and taste/flavour on healthy and indulgent food preferences. *Food Quality and Preference*, *96*, 104380. https://doi.org/10.1016/j.foodqual.2021.104380

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan — a web and mobile app for systematic reviews. *Systematic Reviews, 5*(210). https://doi.org/10.1186/s13643-016-0384-4

Ngo, M. K., Velasco, C., Salgado, A., Boehm, E., O'Neill, D., & Spence, C. (2013). Assessing crossmodal correspondences in exotic fruit juices: The case of shape and sound symbolism. *Food Quality and Preference*, *28*(1), 361–369. https://doi.org/10.1016/j.foodqual.2012.10.004

Padulo, C., Mangone, M., Brancucci, A., Balsamo, M., & Fairfield, B. (2020). Crossmodal congruency effects between sound and food pictures in a forced-choice task. *Psychological Research*. https://doi.org/10.1007/s00426-020-01406-0

Pathak, A., & Calvert, G. A. (2020). Sounds sweet, sounds bitter: How the presence of certain sounds in a brand name can alter expectations about the product's taste. *Food Quality and Preference, 83*, 103918. https://doi.org/10.1016/j.foodqual.2020.103918

Pathak, A., & Calvert, G. A. (2021). Sooo sweeet! presence of long vowels in brand names lead to expectations of sweetness. *Behavioral Sciences*, *11*, 12. https://doi.org/10.3390/bs11020012

Pathak, A., Calvert, G. A., & Motoki, K. (2020). Long vowel sounds induce expectations of sweet tastes. *Food Quality and Preference*, *86*, 104033. https://doi.org/10.1016/j.foodqual.2020.104033

Pathak, A., Calvert, G. A., & Motoki, K. (2021). Sound symbolism overrides articulation dynamics in the taste continuum. *Food Quality and Preference, 91*, 104186. https://doi.org/10.1016/j.foodqual.2021.104186

Pathak, A., Calvert, G. A., Motoki, K., & Park, J. (2022). How early acquired phonemes present in words (or brand names) can evoke the expectations of sweet tastes. *Food Quality and Preference, 96*, 104392. https://doi.org/10.1016/j.foodqual.2021.104392

Peng-Li, D., Byrne, D. V., Chan, R. C. K., & Wang, Q. J. (2020). 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 and Preference*, *85*(264). https://doi.org/10.1016/j.foodqual.2020.103962

Peng-li, D., Mathiesen, S. L., Chan, R. C. K., Byrne, D. V, & Wang, J. (2021). Sounds Healthy: Modelling sound-evoked consumer food choice through visual attention. *Appetite*, *164*, 105264. https://doi.org/10.1016/j.appet.2021.105264

Peterson, R. A., & Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research*, *67*(5), 1035–1041. https://doi.org/10.1016/j.jbusres.2013.08.010

Piqueras-Fiszman, B., Alcaide, J., Roura, E., & Spence, C. (2012). Is it the plate or is it the food? Assessing the influence of the color (black or white) and shape of the plate on the perception of the food placed on it. *Food Quality and Preference*, *24*(1), 205–208. https://doi.org/10.1016/j.foodqual.2011.08.011

Proserpio, C., Laureati, M., Bertoli, S., Battezzati, A., & Pagliarini, E. (2016). Determinants of obesity in Italian adults: the role of taste sensitivity, food liking, and food neophobia. *Chemical senses, 41*(2), 169-176. https://doi.org/10.1093/chemse/bjv072

Qi, Y., Huang, F., Li, Z., & Wan, X. (2020). Crossmodal correspondences in the sounds of chinese instruments. *Perception*, *49*(1), 81–97. https://doi.org/10.1177/0301006619888992

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia - A Window into perception, thought and language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Reinoso-Carvalho, F., Dakduk, S., Wagemans, J., & Spence, C. (2019). Not just another pint! The role of emotion induced by music on the consumer's tasting experience. *Multisensory Research*, *32*(4/5), 367–400. https://doi.org/10.1163/22134808-20191374

Reinoso-Carvalho, F., Gunn, L. H., ter Horst, E., & Spence, C. (2020). Blending emotions and cross-modality in sonic seasoning: Towards greater applicability in the design of multisensory food experiences. *Foods*, *9*(12), 1876. https://doi.org/10.3390/foods9121876

Reinoso-Carvalho, F., Gunn, L., Molina, G., Narumi, T., Spence, C., Suzuki, Y., ter Horst, E., & Wagemans, J. (2020). A sprinkle of emotions vs a pinch of crossmodality: Towards globally meaningful sonic seasoning strategies for enhanced multisensory tasting experiences. *Journal of Business Research*, *117*, 389–399. https://doi.org/10.1016/j.jbusres.2020.04.055

Reinoso-Carvalho, F., Touhafi, A., Steenhaut, K., van Ee, R., Velasco, C. (2017). Using Sound to Enhance Taste Experiences: An Overview. In M. Aramaki, R. Kronland-Martinet, & S. Ystad (eds.) *Bridging People and Sound. CMMR 2016. Lecture Notes in Computer Science*, 10525. Springer, Cham. https://doi.org/10.1007/978-3-319-67738-5_19

Reinoso-Carvalho, F., Van Ee, R., Rychtarikova, M., Touhafi, A., Steenhaut, K., Persoone, D., & Spence, C. (2015). Using sound-taste correspondences to enhance the subjective value of tasting experiences. *Frontiers in Psychology*, *6*, 1309. https://doi.org/10.3389/fpsyg.2015.01309

Reinoso-Carvalho, F., Van Ee, R., Rychtarikova, M., Touhafi, A., Steenhaut, K., Persoone, D., Spence, C., & Leman, M. (2015). Does music influence the multisensory tasting experience? *Journal of Sensory Studies*, *30*(5), 404–412. https://doi.org/10.1111/joss.12168

Reinoso-Carvalho, F., Velasco, C., van Ee, R., Leboeuf, Y., & Spence, C. (2016). Music influences hedonic and taste ratings in beer. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00636

Reinoso-Carvalho, F., Wang, Q. J., de Causmaecker, B., Steenhaut, K., van Ee, R., & Spence, C. (2016). Tune that beer! Listening for the pitch of Beer. *Beverages*, *2*(4), 31. https://doi.org/10.3390/beverages2040031

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. https://doi.org/10.1016/j.foodqual.2016.03.009

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. https://doi.org/10.1016/j.appet.2016.10.026

Roballey, T. C., McGreevy, C., Rongo, R. R., Schwantes, M. L., Steger, P. J., Wininger, M. A., & Gardner, E. B. (1985). The effect of music on eating behavior. *Bulletin of the Psychonomic Society*, *23*(3), 221–222. https://doi.org/10.3758/BF03329832

Rusconi, E., Kwan, B., Giordano, B. L., Umiltà, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, *99*(2), 113–129. https://doi.org/10.1016/j.cognition.2005.01.004

Sakamoto, M., & Watanabe, J. (2016). Cross-modal associations between sounds and drink tastes/textures: A study with spontaneous production of sound-symbolic words. *Chemical Senses*, *41*(3), 197–203. https://doi.org/10.1093/chemse/bjv078

Sakamoto, M., & Watanabe, J. (2013). Effectiveness of onomatopoeia representing quality of tactile texture: a comparative study with adjectives. In *13th National Conference of the Japanese cognitive linguistics association* (pp. 473-485).

Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. *Perception*, *39*(4), 553–569. https://doi.org/10.1068/p6591

Skrzypulec, B. (2021). Constitutivity in Flavour Perception. *Erkenntnis*, *0123456789*. https://doi.org/10.1007/s10670-021-00503-9

Spence, C. (2012). Auditory contributions to flavour perception and feeding behaviour. *Physiology and Behavior*, *107*(4), 505–515. https://doi.org/10.1016/j.physbeh.2012.04.022

Spence, C. (2014). *Noise and its impact on the perception of food and drink*. *Flavour, 3*(1), 1-17. https://doi.org/10.1186/2044-7248-3-9

Spence, C. (2015a). Multisensory flavour perception. *Cell*, *161*(1), 24–35. https://doi.org/10.1002/9781118929384.ch16

Spence, C. (2015b). Eating with our ears: assessing the importance of the sounds of consumption on our perception and enjoyment of multisensory flavour experiences. *Flavour*, *4*(1), 1–14. https://doi.org/10.1186/2044-7248-4-3

Spence, C. (2016). Sound: The forgotten flavor sense. In B. Piqueras-Fiszman & C. Spence (Eds.), *Multisensory flavor perception: From fundamental neuroscience through to the marketplace* (pp. 81-105). Woodhead Publishing. https://doi.org/10.1016/B978-0-08-100350-3.00005-5

Spence, C. (2019). On the relationship(s) between color and taste/flavor. *Experimental Psychology*, *66*(2), 99–111. https://doi.org/https://doi.org/10.1027/1618-3169/a000439

Spence, C. (2022). Exploring group differences in the crossmodal correspondences. *Multisensory Research*, *35*, 495–536. https://doi.org/10.1163/22134808-bja10079

Spence, C., Reinoso-Carvalho, F., Velasco, C., & Wang, Q. J. (2019). Extrinsic auditory contributions to food perception & consumer behaviour: An interdisciplinary review. *Multisensory Research*, *32*(4/5), 275–318. https://doi.org/10.1163/22134808-20191403

Spence, C., Richards, L., Kjellin, E., Huhnt, A.-M., Daskal, V., Scheybeler, A., Velasco, C., & Deroy, O. (2013). Looking for crossmodal correspondences between classical music and fine wine. *Flavour*, *2*(1), 1–13. https://doi.org/10.1186/2044-7248-2-29

Spence, C., & Shankar, M. U. (2010). The influence of auditory cues on the perception of, and responses to, food and drink. Journal of Sensory Studies, 25(3), 406-430. https://doi.org/10.1111/j.1745-459X.2009.00267.x

Stafford, L. D., Fernandes, M., & Agobiani, E. (2012). Effects of noise and distraction on alcohol perception. *Food Quality and Preference*, *24*(1), 218–224. https://doi.org/10.1016/j.foodqual.2011.10.012

Stevenson, R. J. (2014). Flavor binding: Its nature and cause. *Psychological Bulletin*, *140*(2), 487–510. https://doi.org/10.1037/a0033473

Stevenson, R. J., Prescott, J., & Boakes, R. A. (1999). Confusing tastes and smells: How odours can influence the perception of sweet and sour tastes. *Chemical Senses*, *24*(6), 627–635. https://doi.org/10.1093/chemse/24.6.627

Stroebele, N., & de Castro, J. M. (2006). Listening to music while eating is related to increases in people's food intake and meal duration. *Appetite*, *47*(3), 285–289. https://doi.org/10.1016/j.appet.2006.04.001

Velasco, C., Salgado-Montejo, A., Marmolejo-Ramos, F., & Spence, C. (2014). Predictive packaging design: Tasting shapes, typefaces, names, and sounds. *Food Quality & Preference*, *34*, 88–95. https://doi.org/10.1016/j.foodqual.2013.12.005

Ventura, A. K., & Mennella, J. A. (2011). Innate and learned preferences for sweet taste during childhood. *Current Opinion in Clinical Nutrition and Metabolic Care*, *14*(4), 379–384. https://doi.org/10.1097/MCO.0b013e328346df65

Walker, P., Gavin Bremner, J., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, *21*(1), 21–25. https://doi.org/10.1177/0956797609354734

Wang, Q. J. (2017). *Assessing the mechanisms behind sound-taste correspondences and their impact on multisensory flavour perception and evaluation* [Doctoral thesis]. University of Oxford. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.736057

Wang, Q. J., & Spence, C. (2015). Assessing the effect of musical congruency on wine tasting in a live performance setting. *i-Perception*. https://doi.org/10.1177/2041669515593027

Wang, Q. J., & Spence, C. (2016). "Striking a Sour Note": Assessing the influence of consonant and dissonant music on taste perception. *Multisensory Research*, *29*(1–3), 195–208. https://doi.org/10.1163/22134808-00002505

Wang, Q. J., & Spence, C. (2018). "A sweet smile": the modulatory role of emotion in how extrinsic factors influence taste evaluation. *Cognition and Emotion*, *32*(5), 1052–1061. https://doi.org/10.1080/02699931.2017.1386623

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 and Preference*, *58*, 1–9. https://doi.org/10.1016/j.foodqual.2016.12.014

Wang, Q. J., Keller, S., & Spence, C. (2021). Metacognition and crossmodal correspondences between auditory attributes and saltiness in a large sample study. *Multisensory Research*, *34*(8), 785–805. https://doi.org/10.1163/22134808-bja10055

Wang, Q. J., Mesz, B., Riera, P., Trevisan, M., Sigman, M., Guha, A., & Spence, C. (2019). Analysing the impact of music on the perception of red wine via temporal dominance of sensations. *Multisensory Research*, *32*(4–5), 455–472. https://doi.org/10.1163/22134808-20191401

Wang, Q. J., Spence, C., & Knöferle, K. (2020). Timing is everything: Onset timing moderates the crossmodal influence of background sound on taste perception. *Journal of Experimental Psychology: Human Perception and Performance*. https://doi.org/10.1037/xhp0000820

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*(4), 345–356. https://doi.org/10.1093/chemse/bjw007

Wang, Q. J., Woods, A. T., & 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. https://doi.org/10.1177/2041669515622001

Watson, Q. J., & Gunther, K. L. (2017). Trombones elicit bitter more strongly than do clarinets: A partial replication of three studies of Crisinel and Spence. *Multisensory Research*, *30*(3–5), 321–335. https://doi.org/10.1163/22134808-00002573

Woods, A. T., Poliakoff, E., Lloyd, D. M., Kuenzel, J., Hodson, R., Gonda, H., Batchelor, J., Dijksterhuis, G. B., & Thomas, A. (2011). Effect of background noise on food perception. *Food Quality and Preference*, *22*, 42–47. https://doi.org/10.1016/j.foodqual.2010.07.003

World Health Organization (2012). *Guideline: Sodium intake for adults and children*. World Health Organization. https://www.who.int/publications/i/item/9789241504836

World Health Organization (2015). Guideline: *Sugars intake for adults and children*. World Health Organization. https://www.who.int/publications/i/item/9789241549028

Xu, Y., Hamid, N., Shepherd, D., Kantono, K., Reay, S., Martinez, G., & Spence, C. (2019). Background soundscapes influence the perception of ice-cream as indexed by electrophysiological measures. *Food Research International*, *125*, 108564. https://doi.org/10.1016/j.foodres.2019.108564

Xu, Y., Hamid, N., Shepherd, D., Kantono, K., & Spence, C. (2019). Changes in flavour, emotion, and electrophysiological measurements when consuming chocolate ice cream in different eating environments. *Food Quality and Preference*, *77*, 191–205. https://doi.org/10.1016/j.foodqual.2019.05.002

Yan, K. S., & Dando, R. (2015). A crossmodal role for audition in taste perception. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 590–596. https://doi.org/10.1037/xhp0000044

Zampini, M., & Spence, C. (2004). The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *Journal of Sensory Studies*, *19*(5), 347–363. https://doi.org/10.1111/j.1745-459x.2004.080403.x

**Table 2**

*Synthesis of Studies Testing the Association of Tones, Notes, and Musical Instruments with Basic Tastes*

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | | Results |
|---|---|---|---|---|---|---|---|---|
| | | | Sweet | Sour | Salty | Bitter | Umami | |
| Crisinel and Spence (2009) | High-pitched and low-pitched sounds | BT-tasting and SO-tasting food names | | x | | x | | Stronger association between lower-pitched sounds and visually presented BT-tasting foodstuffs and between higher-pitched sounds and SO-tasting foodstuffs. |
| Crisinel and Spence (2010a) – Study 1 | High-pitched and low-pitched sounds | SW and SA food words | x | | x | | | The association between SW tastes and high-pitched sounds and between SA tastes and low-pitched sounds was stronger than the association between SW tastes and low-pitched sounds and between SA tastes and high-pitched sounds. |
| Crisinel and Spence (2010a) – Study 2 | Same as in Study 1 | SA, SW, BT, and SO food words and neutral words | x | x | x | x | | SW and SO tastes were associated with high-pitched sounds. No pitch-taste associations for BT and SA. |
| Crisinel and Spence (2010b) | Musical notes | Gustatory stimuli representing the five basic tastes | x | x | x | x | x | The association of SW and SO tastes to high-pitched notes was confirmed. By contrast, UM and BT tastes were preferentially matched to low-pitched notes. All tastes gave rise to significant preferences in the choice of instrument (medium effect sizes for SW and BT, small effect sizes for SA and SO). |
| Crisinel and Spence (2011) | Musical notes | Milk with different flavors and fat contents | x | x | x | x | | Basic taste ratings were not independent of the choice of instrument (small effect sizes for SA, medium for SW and SO, and large for BT). Flavor significantly affected pitch choice, but no data was available for basic taste ratings. |
| Crisinel and Spence (2012) - Study 1 | Musical notes | Chocolate with different cocoa content | x | x | x | x | | SW ratings were positively correlated with pitch, whereas BT ratings were negatively correlated with pitch. Taste ratings were not associated with the choice of instrument. While there was an overall correlation between pleasantness ratings and the chosen pitch, it disappeared when a single stimulus was considered in isolation. |
| Crisinel and Spence (2012) - Study 2 | NA | NA | x | x | x | x | | The proportion of participants choosing a taste-related adjective was not significantly higher than chance for either of the two pairs (SO-BT or SW-SA). |
| Knöferle et al. (2015) | Short chord progression | NA | x | x | x | x | | SW was associated with a higher pitch, low roughness, and low discontinuity. BT was associated with a lower pitch, high roughness, high discontinuity, and low tempo. SO was associated with high pitch, high roughness, and high tempo. A main effect was found for sharpness, although most pairwise comparisons were nonsignificant. The attack/onset time of the sounds was not reliably linked to any of the basic tastes. |
| Qi et al. (2020) | Musical notes | Chinese words to describe SO, SW, BT, SA, and UM tastes | x | x | x | x | x | Lower pitch was more preferentially matched to a BT or SA taste than SO, SW, or UM. Certain types of Chinese instruments were preferentially matched to taste terms, except for SO (yunluo and guzheng for SW, dizi and erhu for BT, dizi for SA, yunluo for UM). |

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | | Results |
|---|---|---|---|---|---|---|---|---|
| Reinoso-Carvalho, Wang, de Causmaecker, et al. (2016) | Sounds with different pitch levels | Beers | x | | | x | | People tended to match beers with BT-range profiles at significantly lower pitch ranges when compared to the average pitch of a much sweeter beer. |
| Velasco et al. (2014) | High-pitched sound and low-pitched sound | NA | x | x | | | | SW tastes were better expressed through rounded shapes, typefaces, and names (soft, rounded), while SO tastes were better conveyed through more angular shapes, typefaces, and names (sharp, angular). In addition, sounds having a low pitch enhanced the perception of SW whereas high-pitched sounds enhanced the perception of SO. |
| Wang et al. (2016) - Study 1 | Musical notes | Samples of BT, SW, SO, SA, and UM solutions | x | x | x | x | x | Taste quality significantly affected participants' choice of pitch and volume ratings. SW and SO solutions were matched to a significantly higher pitch than the BT, SA, and UM solutions. The SO solution was matched to a significantly higher volume than the SA, SW, and UM solutions. Lower concentration solutions were matched to lower volume and higher concentration solutions were matched with higher volume. The perceived intensity of the samples was correlated with pitch choice, except for SO and UM. |
| Wang et al. (2016) - Study 2 | Same as in Study 1 | Same as in Study 1 | x | x | x | x | x | Taste quality significantly affected participants' choice of pitch and volume ratings. SW and SO solutions were matched to a significantly higher pitch than the BT, SA, and UM solutions. Pitch was positively correlated with SO taste intensity and negatively with UM intensity. |
| Watson and Gunther (2017) | Musical notes | SW, SA, and BT samples | x | | x | x | | This study replicated previous findings of low-pitch/BT and high-pitch/SW crossmodal correspondences. Results differed according to instrument type (trombone or clarinet). |

*Note*. NA = Non-applicable; SW = Sweet; SO = Sour; SA = Salty; BT = Bitter; UM = Umami.

**Table 3**

*Synthesis of Studies Testing the Association of Words, Nonwords, and Speech Sounds with Basic Tastes*

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | | Results |
|---|---|---|---|---|---|---|---|---|
| | | | Sweet | Sour | Salty | Bitter | Umami | |
| Crisinel, Jones, et al. (2012) | Soft-sounding nonwords (Maluma, Lula, Bouba, Bobolo) and hard-sounding nonwords (Kiki, Ruki, Takete, Decter) presented as scale anchors | 12 gustatory stimuli to represent the five basic tastes and more complex flavors | x | x | x | x | x | Salt was consistently associated with sharper sounds (Ruki, Takete, Decter, Kiki). Sugar was significantly more associated with the soft-sounding words Lula (vs. Ruki) and Maluma (vs. Takete), but not Bobolo or Bouba. SO and BT were more associated with Decter (vs. Bobolo), but no differences were observed in the other scales. UM did not give rise to any significant preference in any of the scales. |
| Gallace et al. (2011) | Nonwords (Kiki–Bouba, Maluma–Takete, Lula–Ruki, Decter–Bobolo) | 12 different foods | x | | x | x | | There were crossmodal associations between complex foods/flavors and words, but no significant correlations were found between basic taste ratings (e.g., SA-SW) and word scales (e.g., Bouba-Kiki). |
| Motoki et al. (2020) - Study 1 | Fictitious brand names manipulated for vowel (front/back) and consonant sounds (fricative/stop, voiced/voiceless) | Taste words | x | x | x | x | | Front (vs. back) vowels, fricative (vs. stop) consonants, and voiceless (vs. voiced) consonants respectively increased the expected SW. Fricative (vs. stop) consonants and voiceless (vs. voiced) consonants respectively increased the expected SO. Back (vs. front) vowels and voiced (vs. voiceless) consonants respectively increased expected SA. Stop (vs. fricative) consonants and voiced (vs. voiceless) consonants respectively increased the expected BT. |
| Motoki et al. (2020) - Study 2 | Fictitious brand names manipulated for vowel and consonant sounds (different sets of words) | Taste words | x | x | x | x | | The main effects indicated that front (vs. back) vowels, fricative (vs. stop) consonants, and voiceless (vs. voiced) consonants respectively increased the expected SW. Stop (vs. fricative) consonants increased the expected SO. Stop consonants increased the expected SA more than fricative consonants. Stop (vs. fricative) vowels and voiced (vs. voiceless) consonants respectively increased expected BT. |
| Motoki et al. (2020) - Study 3a, 3b | Fictitious brand names manipulated for vowel and consonant sounds (different sets of words) | Taste words | x | x | x | x | | Experiment 3a found that fricative (vs. stop) consonants and voiceless (vs. voiced) consonants respectively increased the expected SW. Stop consonants and voiced consonants respectively increased the expected SA more than fricative or voiceless consonants. Voiced (vs. voiceless) consonants increased the expected BT. No main effects were observed for expected SO.<br>Experiment 3b found that front (vs. back) vowels and voiceless (vs. voiced) consonants respectively increased expected SW. Front (vs. back) vowels increased the expected SO. Voiced (vs. voiceless) consonants increased expected SA and BT. |
| Motoki, Pathak, and Spence (2022) - Study 1 | 24 vocal stimuli differing in the types of phonation (modal, whispery, creaky, and falsetto) | NA | x | x | x | x | x | Falsetto voices were matched more strongly with SW than the other voices; creaky voices were matched more strongly with SA and BT than the other voices; creaky and falsetto voices were matched more strongly with SO than were the modal voices; modal voices |

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | | Results |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | were matched more strongly with UM than were the falsetto and creaky voices. |
| Motoki, Pathak, and Spence (2022) - Study 2 | Eight vocal stimuli (modal, whispery, creaky, and falsetto) selected from Study 1 | NA | x | x | x | x | x | Falsetto voices were matched more with SW than the other voice types; creaky sounds were matched more with SO than the modal and whispery voices; modal and creaky voices were matched more with SA than the falsetto voices; creaky voices were matched more with BT than the other voices; modal, falsetto and whispery voices were matched more with UM than creaky voices. |
| Motoki, Pathak, and Spence (2022) - Study 3 | Same as in study 2 | Pictures of SW, SO, SA, BT, and UM foods | x | x | x | x | x | Falsetto voices were matched more with SW and SO foods than the other voice types; modal voices were matched more with SA, BT, and UM foods than the other voice types. |
| Ngo et al. (2013) - Study 1 | Line scales corresponding to speech sounds (e.g., Bouba-Kiki) and pitch (e.g., low pitch vs. high pitch) | Six sweetened juices | x | x | | | | For Colombian participants, words containing more ‘rounded’ speech sounds (e.g., Bouba and Maluma) and low-pitched sounds all share some form of crossmodal correspondence with tastes/flavors that are rated as being low in SO. Words containing more ‘sharp’ speech sounds (Kiki and Takete) and high-pitched sounds share some form of crossmodal correspondences with SO tastes/flavors instead. |
| Ngo et al. (2013) - Study 2 | Same as in Study 1 | Same as in Study 1 | x | x | | | | For the participants tested in Oxford, words containing more ‘rounded’ speech sounds (e.g., Bouba and Maluma) and low-pitched sounds all share some form of crossmodal correspondence with tastes/flavors that are rated as being SW and low in SO. Words containing more ‘sharp’ speech sounds (Kiki and Takete) and high-pitched sounds share some form of crossmodal correspondences with SO tastes/flavors instead. |
| Pathak and Calvert (2021) - Study 1 | Ten hypothetical brand name (HBN) pairs differing in vowel type (short vs. long) | NA | x | | | | | Participants rated HBNs with long (vs. short) vowels as more appropriate for a very (vs. less) SW chocolate |
| Pathak and Calvert (2021) - Study 2 | NA | NA | x | | | | | Participants chose a significantly higher number of long (vs. short) vowels for creating BNs of very SW chocolates and used a significantly higher number of long vowels for creating BNs for very (vs. less) SW chocolates. |
| Pathak and Calvert (2021) - Study 3 | 20 hypothetical brand names (HBN) used in Study 1 | Chocolate images (floral/very SW and circular/SW shapes) | x | | | | | Participants associated HBNs with long vowels with the floral (very SW) shape significantly more than with the circular (SW) shape and associated the floral shape more with HBNs with long vowels than with the HBNs with short vowels. |
| Pathak and Calvert (2021) - Study 4 | 10 HBN with long vowels | SW vs. non-SW products | x | | | | | D values significantly different from zero suggested a stronger association of HBN with long vowels with SW products than with non-SW products. Participants were faster in associating SW products with the HBN with long vowels (than with the non-SW products). Differences in error rates were not significantly different. |
| Pathak et al. (2020) - Study 1 | Ten bi-syllabic word pairs differing | SW and non-SW products | x | | | | | The study found a stronger association of long vowels with expectations of SW compared to short vowels. |

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | Results |
|---|---|---|---|---|---|---|---|
| | in vowel type (short vs. long) | | | | | | |
| Pathak et al. (2020) - Study 2 | NA | Images of six SW food items (natural and man-made) and six non-SW items (natural and man-made) | x | | | | Participants used a significantly higher number of long vowels to create novel words for SW (vs. non-SW) food products. |
| Pathak et al. (2020) - Study 3a | Nine bi-syllabic word pairs from Study 1 | SW and very SW products | x | | | | Long (vs. short) had a stronger association with very SW (vs. SW) products. |
| Pathak et al. (2020) - Study 3b | Nine bi-syllabic word pairs from Study 1 | BT and very BT products | | | | x | Vowel length was not significantly associated with BT (vs. very BT) products. |
| Sakamoto and Watanabe (2016) | NA | Six drinks presented in their original form or with the addition of soy sauce, water, or carbonated water | x | x | x | x | This study explored the perceptual structure of gustation categories by analyzing a variety of phonemes of Japanese sound symbolic words spontaneously produced to express taste/textures. The results showed that each sound was associated with a few specific taste evaluation scales. Most phonemes were associated with taste scales as well as texture scales. |
| Simner et al. (2019) | Four different sound continua selected to examine different acoustic qualities: F1, F2, voice discontinuity, and spectral balance | Four tastants (SW, SA, BT, SO) in two concentrations (medium, high) | x | x | x | x | Increasing concentrations of taste corresponded to increasing values in F1, F2, and spectral balance, and more staccato vowel sounds. SW tastes were judged to be low in frequency in F1, F2, and spectral balance. In the latter two sliders, they were judged lower than SO tastes, and in the first slider, they were judged lower than BT, SA, and SO tastes. An additional tendency in F2 was for BT to be rated lower than SO. In other words, these sliders revealed a sequence from SW to BT to SO (with SA numerically between BT and SO). Finally, the SW taste was judged to match smoother, more continuous vowel sounds than the BT and SO tastes, which were judged to match more staccato sounds. |

*Note*. NA = Non-applicable; SW = Sweet; SO = Sour; SA = Salty; BT = Bitter; UM = Umami.

**Table 4**

*Synthesis of Studies Testing the Association of Soundtracks, Music, and Sound Clips with Basic Tastes*

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | Results |
|---|---|---|---|---|---|---|---|
| | | | Sweet | Sour | Salty | Bitter | |
| Guedes et al. (2022) | 100 instrumental musical excerpts | NA | x | x | x | x | All four tastes presented choice rates above what would be expected by chance (i.e., 25%). Significant correlations were observed between taste correspondences and emotional/affective dimensions (e.g., between SW ratings and pleasant emotions). Sweet soundtracks were more easily identified, followed by BT, SA, and SO. |
| Guetta and Loui (2017) - Study 1 | 8 violin music clips for SW, SA, BT, and SO tastes | NA | x | x | x | x | Participants were able to match the clips to taste words with above-chance accuracy. |
| Guetta and Loui (2017) - Study 2 | Same as in Study 1 | NA | x | x | x | x | Participants represented complex auditory stimuli consistent with taste dimensions. The style of musical playing (articulation, nuances, accents, richness) contributed to distinguishing the stimuli from one another. |
| Guetta and Loui (2017) - Study 3 | Same as in Study 1 | 4 types of (SW, SO, SA, BT) chocolate ganache | x | x | x | x | Participants matched chocolates with sounds with above-chance accuracy. Match rates were significantly above-chance for the SW category only. |
| Hauck and Hecht (2019) - Pilot study | 5 classical music pieces | NA | x | x | x | x | Two soundtracks differed significantly in 14 out of 16 dimensions, including SW, SO, and BT. |
| Knöferle et al. (2015) - Study 2 | 4 soundtracks varying the low-level properties of a 30-second piece of music | NA | x | x | x | x | The participants decoded the correct taste word for each piece of music at a level that was significantly above chance. The observed matching performance was most accurate for the SW taste, slightly less accurate for the SA taste, and least accurate for the BT and SO tastes. |
| Knöferle et al. (2015) - Study 3 | Same as in Study 2 | NA | x | x | x | x | Both the U.S. and the Indian sample exceeded chance level in the number of correct mappings. US participants, controlling for music experience, were significantly better at matching the sounds to the correct taste words. Compared to BT sounds, SW sounds were significantly easier to match, while SO and SA sounds were significantly harder to match. |
| Kontukoski et al. (2015) | 4 musical pieces (2 SW and 2 SO) | SW and SO liquid ingredients | x | x | | | Exposure to the SW or SO musical pieces elicited some congruent taste associations in the free association task (e.g., chocolate for SW music) and in the food-pairs task (e.g., stronger choice preference for lemon for SO music). SW or SO elements in the music were reflected in sugar content and acid content of foods developed in association with music. |
| Mesz et al. (2012) | Four musical pieces produced by an algorithm (SO, BT, SW, SA) | NA | x | x | x | x | Results showed that participants could decode well above chance the taste word of the composition. |
| Mesz et al. (2011) - Study 1 | Trained musicians improvised on a MIDI keyboard with piano timbre | NA | x | x | x | x | In free improvisation, taste words elicited consistent musical patterns: BT improvisations were low-pitched and legato (without interruption between |

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | Results |
|---|---|---|---|---|---|---|---|
| | | | | | | | notes), SA were staccato, SO were high-pitched and dissonant, and SW were consonant, slow, and soft. |
| Mesz et al. (2011) - Study 2 | 3 melody improvisations corresponding to each taste word (SO, BT, SW, SA) | NA | x | x | x | x | Non-musical expert listeners classified with high performance the taste word which had elicited the improvisation. |
| Motoki, Takahashi et al. (2022) - Study 1 | 20 soundtracks of four genres (Jazz, Classical, Hip-hop, and Rock/Metal) | 16 (healthy savory, indulgent savory, healthy SW, and indulgent SW foods) food names | x | | | | Listening to Classical music increased people’s preferences for both healthier and indulgent SW foods as compared with the other musical genres. Positive valence mediated the relationship between music genre and indulgent and healthy SW foods. |
| Motoki, Takahashi et al. (2022) - Study 2 | Same as Study 1 | Same as Study 1 | x | | | | Listening to Classical music increased people’s preferences for both healthier and indulgent SW foods compared to the other musical genres. Positive valence mediated the relationship between music genre and indulgent and healthy SW foods. |
| Padulo et al. (2020) | 4 SA, 4 SW, 4 neutral (environmental soundtracks), and 4 silent soundtracks | 128 (SW/high-calorie, SW/low-calorie, SA/high-calorie, SA/low-calorie) food images | x | | x | | Results indicate that low- and high-calorie SA food selection was greater while listening to SA soundtracks, while low- and high-calorie SW food selection was greater while listening to SW soundtracks compared to neutral soundtracks. |
| Peng-Li et al. (2020) | A SW and a SA soundtrack | 16 SW and SA food images | x | | x | | Across both cultures, participants spent more time fixating on SW food while listening to SW music and SA food when listening to SA music, while no differences were observed in the no music condition. Participants’ choices in each sound condition were consistent with fixation time spent. |
| Reinoso-Carvalho, Gunn, ter Horst, and Spence (2020) - Pre-test | 2 songs produced to correspond with softness and hardness and 2 songs to prompt positive and negative emotional effects | NA | x | | | x | A main effect of music was found for the cross-modal ratings. The positive song evoked the sweetest ratings, followed by the soft song. Negative and hard songs evoked the most BT ratings. |
| Reinoso-Carvalho, Van Ee, Rychtarikova, Touhafi, Steenhaut, Persoone, Spence, and Leman (2015) - Control study | 3 (SW, BT, and Medium) soundtracks | NA | x | | | x | Participants’ evaluations of the musical selections differed significantly, with the SW soundtrack rated as the sweetest, the BT soundtrack as the most BT, and the medium soundtrack falling in between. |
| Reinoso-Carvalho et al. (2017) - Pilot study | 2 soundtracks corresponding to smoothness/creaminess and roughness | NA | x | | | x | The creamy soundtrack was rated as significantly sweeter than the rough soundtrack. |
| Wang et al. (2017) - Pilot study | A SW and a spicy soundtrack | NA | x | x | x | | The spicy and SW soundtracks were matched significantly more frequently to spicy and SW tastes, respectively, than all other options. |
| Wang et al. (2015) | 5 BT, 5 SA, 7 SO, and 7 SW soundtracks previously produced by various researchers and designers | NA | x | x | x | x | Out of the 24 soundtracks, only three had nonsignificant preferences in the choice of taste matches. Overall, the SW soundtracks most effectively evoked the taste intended by the composer, whereas |

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | Results |
|---|---|---|---|---|---|---|---|
| | | | | | | | the BT soundtracks were the least effective. |
| Wang and Spence (2016) - Control experiment | Consonant and dissonant versions of two short melodies | NA | x | x | | | Participants reliably associated the consonant soundtracks with SW and the dissonant soundtracks with SO. |
| Wang et al. (2021) | 36 short sound clips varying in 13 musical (e.g., articulation, tempo, consonance) and emotional (valence and arousal) attributes | NA | | | x | | The results revealed that SA was associated most strongly with a long decay time, high auditory roughness, and a regular rhythm. Regarding emotional associations, SA was matched with negative valence, high arousal, and minor mode. |

*Note*. NA = Non-applicable; SW = Sweet; SO = Sour; SA = Salty; BT = Bitter.

**Table 5**

*Synthesis of Studies Testing the Effect of Music on Basic Taste Perception*

| Study | Auditory stimuli | Gustatory stimuli | Basic Taste | | | | | Results |
|---|---|---|---|---|---|---|---|---|
| | | | Sweet | Sour | Salty | Bitter | Umami | |
| De Luca et al. (2018) | Peppy and lively vs. Fine music | White and red wine | x | x | | | | The white wine was perceived to be more delicate and sweeter if accompanied by a classical (vs. pop) music background. The red was perceived as less alcoholic when pop music (vs. silence) was transmitted, but no differences in taste ratings were observed. |
| Hauck and Hecht (2019) - Main study | Two classical music pieces selected in a pilot study | Red wine, white wine, sugar water, and citric acid solution | x | x | x | x | | There was a significant main effect of music on rating dimensions. The univariate effect of the factor ‘music’ was significant in SO and SA dimensions but not BT and SW. The two music pieces were previously shown to differ significantly in BT, SW, and SO but not SA. |
| Kantono et al. (2019) | 14 sound segments in different music genres (with the least liked, more liked, and closest to the mid-range liking score selected for each individual) | Bittersweet gelato samples | x | | | x | | Electrophysiological measures can covary sensory changes while listening to music. Ratings of positive emotions were associated with SW perception, while ratings of negative emotions were associated with BT perception. |
| Kantono et al. (2016) | Music samples representing 14 musical genres | Three (dark/BT, bittersweet, milk/SW) chocolate gelati | x | x | x | x | x | The TDS difference curves showed significant differences between gelati samples and music conditions. SW was perceived more dominant when neutral and liked music were played, while BT was more dominant for disliked music. |
| Lin et al. (2022) | 14 segments of music of different genres, 14 sounds, mixtures of liked music and pleasant sound (LMPS), and mixtures of disliked music and unpleasant sound (DMUS) unique to the individual panelists | Bittersweet chocolate ice cream | x | | | x | | Consuming ice cream during the liked music condition resulted in the longest duration of perceived SW, whereas BT was dominant under the disliked music and unpleasant sound (DMUS) and disliked music conditions, respectively. Positive emotions correlated with SW and cocoa in the positively valenced auditory conditions. Negative emotions were associated with BT and roasted tastes/flavors under the negatively valenced auditory conditions. |
| Reinoso-Carvalho, Velasco, et al. (2016) | A fragment of the song “Oceans of Light" by The Editors | Dark ale beer | x | x | | x | | The presence of the song seemed to have a modulatory effect on the perceived SO of the beer, but not SW and BT. |
| Reinoso-Carvalho, Van Ee, Rychtarikova, Touhafi, Steenhaut, Persoone, and Spence (2015) | A fragment of “Vem Morena, Vem” by Jorge Bem | Chocolate sample | x | x | x | x | | Customers reported a better tasting experience when the sounds were presented as part of the food’s identity and were willing to pay significantly more for the experience. The participants who heard the kitchen soundscape (condition B) instead of the song (conditions A, C, and D) rated the chocolate as tasting less SW, although this comparison failed to reach statistical significance. The comparisons between the ratings on BT, SO, and SA levels among conditions were inconclusive. |

| Study | Auditory stimuli | Gustatory stimuli | Basic Taste | | | | Results |
|---|---|---|---|---|---|---|---|
| Reinoso-Carvalho et al. (2019) - Study 1 | Positive and negative music | Beer (BT-dry pale lager) | x | x | | x | No differences were found between sensory ratings of the beer (SW, BT, and SO) when comparing positive versus negative or sound versus silence conditions. |
| Reinoso-Carvalho et al. (2019) - Study 2 | Two positive and one negative music track | Beer (pale-lager) | x | x | | x | Participants rated the beer as tasting sweeter while listening to the positive (vs. negative) music and more BT while listening to the negative (vs. positive) music. No differences were observed for SO ratings. |
| Reinoso-Carvalho et al. (2019) - Study 3 | Two positive and one negative music | Beer (strong dark ale) | x | x | | x | Participants rated the beer as tasting sweeter while listening to the positive (vs. negative) music and more BT while listening to the negative (vs. positive) music. No differences were observed for SO ratings. |
| Spence et al. (2013) - Study 2 | Four classical music pieces | Four wines with distinctive sensory characteristics (e.g., acidity, SW) | x | x | | | Participants perceived the wine as tasting sweeter and enjoyed the experience more while listening to the matching music than while tasting the wine in silence. No significant difference in acidity. |
| Stafford et al. (2012) | A piece of music and news articles recorded by a male voice | Five freshly prepared drinks of cranberry juice and vodka in different proportions | x | | | x | SW perception was significantly higher in the music than in control and other distracting conditions. BT ratings were lower in the music group compared to the control. |
| Wang and Spence (2015) | Two pieces of classical music (chosen to match each wine) | White and red wine | | x | | | The wine tasted while listening to Debussy was rated as significantly more acidic than while listening to Rachmaninoff, regardless of the type of wine. |
| Wang et al. (2019) | Two pieces of music that varied in tempo, mode, and instrumentation | Red wine | x | x | | x | There was a higher dominance duration for BT in both music conditions (vs. silence) but no differences between the two soundtrack conditions. No significant differences were observed in dominance durations of SW and acidity. |

*Note*. NA = Non-applicable; SW = Sweet; SO = Sour; SA = Salty; BT = Bitter.

**Table 6**

*Synthesis of Studies Testing the Effect of Soundtracks on Basic Taste Perception*

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | Results |
|---|---|---|---|---|---|---|---|
| | | | Sweet | Sour | Salty | Bitter | |
| Crisinel, Cosser, et al. (2012) | BT (trombone) and SW (piano) soundtracks | Cinder toffee | x | | | x | The samples were rated as significantly more BT when listening to the BT (vs. SW) soundtrack. |
| Höchenberger and Ohla (2019) - Study 1 | BT (low-pitch trombone) and SW (high-pitch piano) soundtracks | Cinder toffee | x | | | x | Samples tasted sweeter with a SW sound and more BT with a BT sound, but this effect disappeared when a “no-sound” control was included in the statistical model. |
| Höchenberger and Ohla (2019) - Study 2 | Same as in study 1 | Same as in study 1 | x | x | x | x | Taste quality-specific intensity was not influenced by sound. |
| Reinoso-Carvalho et al. (2017) | Two soundtracks corresponding to smoothness/creaminess and roughness | Four chocolates with 2 shapes (angular vs. round) and two formulas (71% vs. 81% cocoa content) | x | | | x | The participants reported that the chocolates tasted sweeter while listening to the creamy soundtrack and more BT while listening to the rough soundtrack. |
| Reinoso-Carvalho, Gunn, ter Horst and Spence (2020) | 2 soft/hard songs and 2 (positive/negative) songs | Two types of chocolates (milk and dark) | x | x | | x | The chocolates were rated as having a more intense flavor for those listening to softer (vs. harder) music. There was no evidence that participants rated the chocolate’s SW, BT, or SO differently while listening to soft (vs. hard) songs. The chocolate was rated as tasting sweeter when evaluated with positive (vs. negative) music, but no differences were observed for SO and BT. |
| Reinoso-Carvalho, Gunn, Molina et al. (2020) | 2 soft/hard songs and 2 (positive/negative) songs | Two types of chocolates (milk and dark) | x | | | x | Chocolate was rated as sweeter with positive music (regardless of chocolate and culture) and more BT with the negative music (except for dark chocolate in the Japanese sample). No differences concerning the sensory evaluation of the chocolates were significant with soft/hard music. |
| Reinoso-Carvalho, Van Ee, Rychtarikova, Touhafi, Steenhaut, Persoone, Spence, and Leman (2015) | 3 soundtracks congruent with SW, BT, and medium chocolates | BT, medium, and SW chocolates | x | | | x | The BT chocolate was rated as sweeter when listening to the subject SW soundtrack [vs. silence] and as more BT with the BT song [vs. medium sound], with the subject BT song [vs. subject sweet sound], and with the subject medium sound [vs. subject SW sound]. Differences were found only for the BT chocolate and more strongly for the subject-matched soundtracks (i.e., participants' individual music-chocolate matches). |
| Reinoso-Carvalho, Wang, van Ee, et al. (2016) - Study 1 | SW and BT soundtracks | Beer | x | | | x | The participants rated the beer as significantly sweeter when listening to the SW soundtrack than when listening to the BT soundtrack. |
| Reinoso-Carvalho, Wang, van Ee et al. (2016) - Study 2 | SW and SO soundtracks | Beer | x | x | | | The participants rated the beer as significantly sweeter when listening to the SW soundtrack than when listening to the SO soundtrack. |
| Reinoso-Carvalho, Wang, van Ee | SO and BT soundtracks | Beer | | x | | x | No significant differences were found in any of the taste ratings. |

| Study | Auditory stimuli | Gustatory stimuli | Basic taste | | | | Results |
|---|---|---|---|---|---|---|---|
| et al. (2016) - Study 3 | | | | | | | |
| Reinoso-Carvalho, Wang, van Ee et al. (2016) - Control | NA | Same as in Studies 1-3 | x | x | | x | The ratings differed significantly only for the beers in experiment 3, compared to silence. The beer was rated as significantly more BT and SO in single taste scales while listening to the BT soundtrack, compared to silence. The beer was also rated as more SO on the BT-SO scale while listening to the BT soundtrack. |
| Wang and Spence (2018) | Consonant and dissonant versions of a short melody | Juice mixture | x | x | | | Participants rated the juice as sweeter and less SO while exposed to more positive (vs. negative) visual and musical stimuli. |
| Wang et al. (2017) - Study 2 | Spicy and SW soundtracks and white noise | A dish with SW and spicy components | x | | | | Soundtracks modified people's evaluation of the expected and actual spiciness of foods but not of expected or actual ratings of SW taste. |
| Wang and Spence (2016) - Study 1 | Consonant and dissonant versions of two short piano melodies | Juice mixtures | x | x | | | Participants rated the fruit juice as tasting sourer while listening to the dissonant music than while listening to the consonant music. The correlations between music pleasantness ratings and SO-SW ratings were significant. |
| Wang and Spence (2016) - Study 2 | Consonant and dissonant versions of one melody with piano and trumpet sounds | Two blends of fruit juice | x | x | | | Participants rated the juice as tasting significantly sweeter while listening to consonant (vs. dissonant) music. The correlation between ratings of music pleasantness and ratings on the SO-SW scale was significant. |
| Wang et al. (2020) - Study 1 | A SW and BT soundtrack | 70% chocolate | x | | | x | The chocolates experienced with the sweet soundtrack were rated as sweeter than the chocolates experienced with the BT soundtrack. The taste-congruent soundtracks had no such effect on taste ratings for those participants who heard the soundtrack only after tasting. |
| Wang et al. (2020) - Study 2 | Same as in study 1 | Same as in study 1 | x | | | x | Participants rated the chocolates as tasting sweeter when they had listened to the SW soundtrack – either before or during tasting – compared to the BT soundtrack. |

*Note*. NA = Non-applicable; SW = Sweet; SO = Sour; SA = Salty; BT = Bitter.

**Table 7**

*Synthesis of Studies Testing the Effect of Notes, Noise, and Soundscapes on Basic Taste Perception*

| Study | Category | Auditory stimuli | Gustatory stimuli | Basic Taste | | | | | Results |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Sweet | Sour | Salty | Bitter | Umami | |
| Bravo-Moncayo et al. (2020) | Noise | Three versions of a food court noise (baseline unfiltered noise - UBN; passive-controlled noise - PBN, and active-controlled noise - ABN) | Coffee | x | x | | x | | Consumers rated the coffee as more BT while listening to the filtered (vs. unfiltered) noises. The differences in acidity and SW were not significant. |
| Burzynska et al. (2019) | Notes and tones | Sine waves with a bass frequency of 100 Hz and 1000 Hz | Two red wines | | x | | | | The acidity ratings of both wines did not differ significantly across sound conditions. |
| Lin et al. (2019) | Soundscape | Environmental sounds: café, fast food restaurant, bar, food court, and park | Chocolate gelato | x | | | x | | BT, roasted, and cocoa notes were more evident when the bar, fast food, and food court sounds were played. SW was cited more in the early mastication period when listening to park and café sounds. The valence evoked by the pleasant park sound was positively correlated with SW, whereas the arousal associated with bar sounds was correlated with BT, roasted, and cocoa attributes. |
| Lorentzen et al. (2021) | Noise | One "quiet" sound setting and one recording of fMRI noise | Five concentrations of four basic tastants (SW, SA, SO, and BT) from the Taste-Drop-Test | x | x | x | x | | Loud acoustic fMRI noise did not affect gustatory perception for any of the tastants. |
| Woods et al. (2011) | Noise | White noise (Quiet, loud) or no noise (Baseline) | Savory/crunchy (salted crisps), savory/soft (mini cheese), SW/crunchy (Biscuits), and SW/soft (flapjack) stimuli | x | | x | | | Background noise reduced the reported intensity of SW and SA, regardless of hardness (crunchy or soft). |
| Xu et al. (2019) | Soundscape | Café soundscape presented alone (control) or overlaid with bird (café-bird), forest (café-forest), and machine sounds (café-machine) | Chocolate ice cream | x | | | x | | SW and creaminess were dominant at the start of the consumption episode while listening to the café-forest soundscape. BT was perceived at the end of the consumption period in café/control and café-machine soundscapes. Listening to the café-machine soundscape evoked negative emotions associated with BT and creaminess, which were also associated with increased heart rate (HR) and respiration rate (RESP). When listening to the café-forest soundscape, ice cream was associated with SW and positive emotions. |
| Yan et al. (2015) | Noise | Cabin noise | Aqueous solutions for the five basic tastes at three concentration levels | x | x | x | x | x | Loud noise suppressed SW taste perception across all concentrations and enhanced UM in medium and high concentrations. SO, BT, and SA tastes were unaffected. |

*Note*. NA = Non-applicable; SW = Sweet; SO = Sour; SA = Salty; BT = Bitter; UM = Umami.

**Highlights**

- This review found 60 articles examining the interplay of audition and taste
- Basic tastes are crossmodally associated with sounds (e.g., musical notes, music)
- Auditory stimuli (e.g., soundscapes, music) influence taste perception

**Author Statement**

David Guedes: Conceptualization, Investigation, Formal analysis, Methodology, Writing – original draft.

Margarida V. Garrido: Conceptualization, Methodology, Writing – review & editing.

Elsa Lamy: Conceptualization, Methodology, Writing – review & editing.

Bernardo P. Cavalheiro: Formal analysis.

Marília Prada: Conceptualization, Methodology, Writing – review & editing.

**Declaration of interests**

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

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: