CBS COPENHAGEN BUSINESS SCHOOL
HANDELSHØJSKOLEN

Inducing Novel Sound–taste Correspondences Via an Associative Learning Task

Escobar, Francisco Barbosa; Wang, Qian Janice

Document Version
Final published version

Published in:
Cognitive Science

DOI:
10.1111/cogs.13421

Publication date:
2024

License
CC BY

Citation for published version (APA):
Escobar, F. B., & Wang, Q. J. (2024). Inducing Novel Sound–taste Correspondences Via an Associative Learning Task. *Cognitive Science*, *48*(3), Article e13421. https://doi.org/10.1111/cogs.13421

Link to publication in CBS Research Portal

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025

COGNITIVE SCIENCE
A Multidisciplinary Journal

Check for updates

Cognitive Science 48 (2024) e13421
© 2024 The Authors. *Cognitive Science* published by Wiley Periodicals LLC on behalf of Cognitive Science Society (CSS).
ISSN: 1551-6709 online
DOI: 10.1111/cogs.13421

Inducing Novel Sound–Taste Correspondences via an Associative Learning Task

Francisco Barbosa Escobar,[a,b] Qian Janice Wang[a]

[a]*Department of Food Science, Faculty of Science, University of Copenhagen*
[b]*Department of Marketing, Copenhagen Business School*

Received 25 September 2023; received in revised form 8 January 2024; accepted 19 February 2024

Abstract

The interest in crossmodal correspondences, including those involving sounds and involving tastes, has experienced rapid growth in recent years. However, the mechanisms underlying these correspondences are not well understood. In the present study ($N = 302$), we used an associative learning paradigm, based on previous literature using simple sounds with no consensual taste associations (i.e., square and triangle wave sounds at 200 Hz) and taste words (i.e., sweet and bitter), to test the influence of two potential mechanisms in establishing sound–taste correspondences and investigate whether either learning mechanism could give rise to new and long-lasting associations. Specifically, we examined an emotional mediation account (i.e., using sad and happy emoji facial expressions) and a transitive path (i.e., sound-taste correspondence being mediated by color, using red and black colored squares). The results revealed that the associative learning paradigm mapping the triangle wave tone with a happy emoji facial expression induced a novel crossmodal correspondence between this sound and the word sweet. Importantly, we found that this novel association was still present two months after the experimental learning paradigm. None of the other mappings, emotional or transitive, gave rise to any significant associations between sound and taste. These findings provide evidence that new crossmodal correspondences between sounds and tastes can be created by leveraging the affective connection between both dimensions, helping elucidate the mechanisms underlying these associations.

This research was supported by a Carlsberg Young Researcher Fellowship received by Qian Janice Wang. We have no conflicts of interest to disclose. Data, script, and stimuli are openly available at the project's OSF page (https://osf.io/uxp29/).

Correspondence should be sent to Francisco Barbosa Escobar, Department of Marketing, Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark. E-mail: fjbe.marktg@cbs.dk

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Moreover, these findings reveal that these associations can last for several weeks after the experimental session through which they were induced.

Keywords: Crossmodal correspondences; Sound; Taste; Emotions; Color; Semantic; Associative learning

1. Introduction

Eating and drinking are some of the most multisensory experiences in individuals' everyday lives (Spence, 2017), in which most, if not all, sensory modalities come to play (Prescott, 2015). A key part of these experiences relates to flavor, which refers to the interaction of a myriad of sensory signals, mainly tastes and odors (Prescott, 2015) that form a unitary percept (Auvray & Spence, 2008). Here, the senses of taste and smell seem to be constitutive of flavor perception, while other senses such as audition seem to be modulatory (Spence, 2015b). Nevertheless, sound and flavor are tightly connected in food-related experiences. Indeed, although it has long been overlooked, audition plays a key role in flavor perception (Spence, 2015a). For instance, the sounds produced by the physical interaction with foodstuffs, whether it is with the hands or in the mouth (e.g., chewing, biting), can carry a great amount of information about their properties and set expectations about their eventual perception (Spence, 2012). In addition, sounds and flavor are linked through the environmental sounds present when eating or drinking that sometimes correlate to specific foods or drinks, such as the sounds commonly found in coffee shops (Spence, 2012).

Furthermore, previous research has evidenced crossmodal interactions and crossmodal correspondences between basic tastes and different dimensions of sounds (Guedes, Vaz Garrido, Lamy, Pereira Cavalheiro, & Prada, 2023; Knöferle & Spence, 2012; Velasco, Corradi, & Motoki, 2023). Crossmodal correspondences refer to, often unexpected, associations most people tend to have between dimensions or stimuli in two different sensory modalities (Spence, 2011, 2022). Correspondences between sound and tastes (as actual tastants and taste words) found in past literature have involved a variety of sound stimuli, including pure tones, musical tones, words, non-words, speech sounds, music, and soundtracks (see Guedes et al., 2023, for a review), and the most prominent psychoacoustical parameter studied relates to pitch. Overall, the most consistent (but by no means perfect) correspondences found in this literature relate to associations between sweetness (and in some cases also sourness) and high-pitch sounds and between bitterness and low-pitch sounds (Crisinel & Spence, 2009, 2010a, 2010b; Knoeferle, Woods, Käppler, & Spence, 2015; Qi, Huang, Li, & Wan, 2020; Wang, Wang, & Spence, 2016; Q. J. Watson & Gunther, 2017). For instance, using both short chord progressions and more complex soundtracks, Knoeferle et al. (2015) found that people encode the word sweet into high-pitch sounds, whereas they encode the word bitter as low-pitch sounds. Furthermore, using different concentrations of water-based basic taste solutions and a range of tones produced by a MIDI keyboard, Wang et al. (2016) found that the tones associated with the sour solutions had the highest frequency, followed by those associated with the sweet solutions, whereas the tones associated with bitterness had the lowest frequencies.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Even though previous research has uncovered crossmodal correspondences between sounds and basic tastes using different types of both sound and basic taste stimuli, the mechanisms behind these associations are not well understood, although some explanations have been proposed (see Guedes et al., 2023). In general, literature explicitly studying the underlying mechanisms of crossmodal associations is scarce. That being said, Barbosa Escobar, Velasco, Byrne, and Wang (2023) recently investigated the relative influence of two different mechanisms in the context of crossmodal associations between visual textures and temperature using an associative learning paradigm. In the present paper, we aimed to fill this gap in the literature related to sound–taste correspondences. More specifically, we build on the associative learning paradigm used by Barbosa Escobar et al. (2023), which can easily be used to investigate the underlying mechanisms of different crossmodal associations, to examine the relative influence of two potential mechanisms in the formation of novel sound–taste correspondences, namely, an emotional account and a transitive path (i.e., through color associations). In addition to the contributions of Barbosa Escobar et al. (2023), the present paper further investigates the role of learning in the formation of these novel correspondences. Furthermore, for the first time in the literature, we commenced investigating the lifespan of experimentally induced crossmodal correspondences and examined whether individuals still exhibited the correspondences created here several weeks after the experimental session. The present research contributes to a better understanding of how sound–taste correspondences may arise in individuals' minds, as it is still not well understood in the scientific literature and that novel correspondences involving these dimensions can be created. Importantly, from a broad perspective, this research further enhances the understanding of the important role (ease of) learning plays for crossmodal correspondences and multisensory integration in general. This research also contributes to the scientific literature by elucidating that the lifespan of crossmodally correspondences originating in the laboratory may last at least several weeks.

2. Theoretical background

As mentioned earlier, studies explicitly investigating the underlying mechanisms of sound–taste correspondences are limited, and there does not seem to be any statistical regularity in the environment connecting the two features, with the exception of perhaps sounds produced by orofacial gestures when consuming pleasant or unpleasant foods (Knöferle & Spence, 2012). Although the findings of previous literature hint at some possible mechanisms of sound–taste correspondences overall (see Guedes et al., 2023), which may include an emotional mediation account, based on the pleasantness of different timbres (Crisinel & Spence, 2010a, 2012) or different levels of pitch (Wang et al., 2016), as well as a structural account (i.e., louder sounds are matched with more intense basic tastes; Wang et al., 2016) and a semantic account (i.e., the same words, such as sweet, may be used to describe taste and musical attributes; Guedes, Prada, Garrido, & Lamy, 2022). In the present paper, we focus on an emotional mediation account and on intermediate associations with color, as per the transitive property of crossmodal correspondences.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.1. An emotional mediation account

Both sounds and tastes can be associated with different emotions and evoke different affective reactions. Regarding sound, research has shown that music, for instance, can evoke strong and vivid emotions (Juslin, Liljeström, Västfjäll, & Lundqvist, 2010; Koelsch, 2014; Zentner, Grandjean, & Scherer, 2008) to the extent that individuals often use music to self-regulate their mood and emotional state (Attie-Picker, Venkatesan, Newman, & Knobe, 2024; Saarikallio, 2011). While lyrics can carry a great deal of emotional information, they are not necessary for music to evoke emotions, although lyrics can modulate the intensity of these emotions (Barradas & Sakka, 2022; Brattico et al., 2011). Indeed, different psychoacoustic properties, such as loudness, pitch level, pitch contour, tempo, texture, and sharpness (Coutinho & Cangelosi, 2011; Jaquet, Danuser, & Gomez, 2014), as well as timbre (Le Groux & Verschure, 2012), in music have been shown to influence the valence and emotions evoked by music. Moreover, the timbre of the sounds produced by different musical instruments in isolation affects the valence and emotions generated by their sounds (Eerola, Ferrer, & Alluri, 2012). In addition to music, speech sounds (e.g., phonemes) can have specific inherent emotional qualities based on certain acoustic qualities such as the acoustic frequency components of phonemes (Auracher, Menninghaus, & Scharinger, 2020; Myers-Schulz, Pujara, Wolf, & Koenigs, 2013; Sabbatino, Troiano, Schweitzer, & Klinger, 2022). Furthermore, previous research has shown that, similar to music and speech sounds, changes in specific acoustic features (e.g., frequency spectrum, intensity, rate) of environmental sounds can influence the emotional evaluation of said sounds (Ma & Thompson, 2015).

Considering taste, specific basic tastes are associated with different levels of valence and arousal and can evoke different emotions. Past studies have investigated these effects using various autonomic nervous system parameters (e.g., skin conductance, blood flow, temperature, and heart rate; Robin, Rousmans, Dittmar, & Vernet-Maury, 2003; Rousmans, 2000), as well as facial skin blood flow (Kashima & Hayashi, 2011), and skin conductance (Spinelli et al., 2023). In addition, taste words can evoke affective responses with different levels of valence and arousal, as shown by participants' self-reported ratings (Wang, Woods, & Spence, 2015). Broadly, this research has shown that sweetness generates positive emotional reactions. On the other hand, bitterness triggers negative emotional responses.

Given that both sounds with specific properties and basic tastes have been found to be associated with given emotions and affective reactions, it is possible that these stimuli are associated via an emotional mechanism. The emotional-mediation account of crossmodal correspondences poses that two different dimensions or stimuli may be matched together based on congruent emotional associations or similar affective evocations (Spence, 2020; Whiteford, Schloss, Helwig, & Palmer, 2018). The emotional mediation account may help explain correspondences for which there are no physical regularities in the environment (Spence, 2020), as it seems to be the case here. Due to the apparent absence of any physical relationships between sounds and tastes (although see Knöferle & Spence, 2012, who proposed that these associations may arise from differences in the pitch of sounds produced by tongue protrusions when tasting pleasant tastes compared to unpleasant ones), the emotional mediation account represents a good candidate to explain these correspondences. Indeed, as Wang et al. (2016) found, the valence dimension of affect seems to mediate associations between pitch and basic tastes. More specifically, using notes from 19 keys on a MIDI keyboard and three different levels of intensities of the five basic tastes, the higher the valence (i.e., pleasantness) of the taste, the higher the valence (i.e., pleasantness) of the single musical note participants chose to match with the taste. However, these results are different from Crisinel and Spence (2012), who did not find any relationship between the chosen pitch matched with dark chocolate. Nevertheless, in the latter study, only one type of food stimuli was used, and the valence of pitch was not measured.

Considering the literature presented earlier, the emotional mediation account may explain correspondences between sounds and basic tastes based on common emotional associations or evocations from both dimensions. Hence, we expected that an associative learning paradigm mapping a sound to an emotional stimulus with consensual taste associations would cause people to associate said sound with the corresponding taste more often than in the absence of such mapping (i.e., higher probability of association). More formally, we hypothesized that:

H_{1A}: Participants will categorize a sound as sweet more often after being exposed to an associative learning mapping matching the sound with an emotional stimulus related to sweetness than before being exposed to the mapping.

H_{1B}: Participants will categorize a sound as bitter more often after being exposed to an associative learning mapping matching the sound with an emotional stimulus related to bitterness than before being exposed to the mapping.

2.2. *A transitive path*

Given the lack of clear statistical links in the environment on which correspondences between sounds and tastes may emerge, besides an indirect emotional mediation account, it is worth considering that these correspondences may come from associations through a third type of stimuli or feature, a property of crossmodal correspondences termed transitivity (Deroy, Crisinel, & Spence, 2013; Deroy & Spence, 2013; Fields, Verhave, & Fath, 1984; Spence, 2019; Spence & Deroy, 2012). The transitivity property of crossmodal correspondences poses that if a dimension X in one sensory modality is associated with a dimension Y in another modality, and Y, in turn, is associated with a dimension Z in a third sensory modality, then the brain creates associations between X and Z (Deroy et al., 2013). For example, knowing that density and pitch are positively related and that there is a positive relationship between pitch and brightness, it is possible to arrive at a positive association between density and brightness (Boring & Stevens, 1936). The transitive property therefore allows two dimensions that do not statistically co-occur in the environment to be matched crossmodally. In the case of the sound–taste correspondences studied here, it is possible that they occur through transitive mapping with colors. In other words, they may occur via mediation by pitch–color and color–taste correspondences.

Multiple studies have uncovered crossmodal correspondences between different dimensions of sound and color (see Spence & Di Stefano, 2022, for a review). When it comes to pitch, a couple of studies have found associations between pitch and color hue and

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

brightness. For example, controlling for lightness and using both pure tones and complex sounds, Hamilton-Fletcher, Witzel, Reby, and Ward (2017) found a linear relationship between sounds with dominant frequency ranges and hue, so that low frequencies were associated with blue hues, whereas high frequencies (above 800 Hz) were associated with yellow hues. In addition, the latter authors found a positive relationship between chroma and both pitch and loudness. More recently, Sun et al. (2018) used an explicit test and an implicit association test (IAT) and found that people tend to associate high-pitch sounds (523 Hz) with red hues and low-pitch sounds (130 Hz) with blue hues. Despite the findings of these two studies, it is important to note that the results of other studies investigating correspondences between pitch and color are not consistent, questioning the existence of these associations, at least when luminance is controlled for (Spence & Di Stefano, 2022). Another dimension studied in relation to sound-color correspondences relates to timbre (Adeli, Rouat, & Molotchnikoff, 2014; Reuter et al., 2018). For example, in a conference paper, Reuter et al. (2018) reported having found associations between musical notes from different instruments and colors (e.g., trumpet–red, low register flute–blue).

Regarding color–taste associations, this is perhaps one of the oldest and most widely known sets of crossmodal correspondences, and a myriad of studies have uncovered correspondences between these two dimensions (e.g., Saluja & Stevenson, 2018; Tomasik-Krótki & Strojny, 2008; Velasco et al., 2016; Wan et al., 2014; see Spence et al., 2015 and Spence & Levitan, 2021, for reviews and discussions on the underlying mechanisms of these correspondences). For instance, Saluja and Stevenson (2018) found specific associations between colors and tastes using real tastants corresponding to the five basic tastes (i.e., sweet, salty, sour, bitter, and umami—called "meaty" in their study), each at three concentration levels, which were consistent with previous research using taste words. Overall, the literature so far has found that sweet is associated with red and pink hues, sour is associated with yellow and green hues, salty with blue and white hues, and bitter is associated with brown and black hues (Spence et al., 2015). As per umami, although Kikunae Ikeda suggested that if umami had a color, it would be yellow (Ikeda, 1909, 2002), the results of the few previous studies that have examined this taste have not yielded consistent associations (Tomasik-Krótki & Strojny, 2008; Wan et al., 2014; see Spence et al., 2015).

Informed by the transitive property of crossmodal correspondences, together with the literature on both sound–color and color–taste correspondences, which may help explain the associations between sounds and basic tastes, we expected that an associative learning paradigm that matched a sound with a color stimulus known to be associated with a specific taste would make people associate the sound with the corresponding taste more often (i.e., higher probability of association). More formally, we hypothesized that:

- H_{2A}: Participants will categorize a sound as sweet more often after being exposed to an associative learning mapping matching the sound with a color stimulus related to sweetness than before being exposed to the mapping.
- H_{2B}: Participants will categorize a sound as bitter more often after being exposed to an associative learning mapping matching the sound with a color stimulus related to bitterness than before being exposed to the mapping.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3. Main experiment

3.1. Methods

3.1.1. Participants

The required sample size was determined via a power analysis based on a goodness-of-fit test in G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) to obtain a statistical power of at least .80 using a medium effect size of Cohen's W $= 0.3$ with an alpha level of 0.05. The power calculation yielded a required sample size of 280 participants. Data collection took place on July 13, 2023. A total of 310 native English speakers from the United Kingdom (156 females, 153 males, 1 other), aged 18–40 years ($M_{age} = 30.97$ years, $SD_{age} = 5.94$), took part in the experiment.

Participants were recruited from Prolific (https://www.prolific.com/) and received GBP 1.35 in compensation. To increase the quality of the participants' pool in Prolific, we pre-screened for participants who had an approval rate greater than 98% and who were not on other crowdsourcing platforms, as determined by Prolific's pre-screening options. The data from eight participants were removed, as they achieved an overall level of correct responses in the learning blocks lower than 75%. The final sample consisted of 302 individuals (154 females, 147 males, 1 other), aged 18–40 years ($M_{age} = 31.00$ years, $SD_{age} = 5.93$). The experiment complied with the World Medical Association's Declaration of Helsinki, and it was approved by the University of Copenhagen's Research Ethics Committee, as an institutional review board, under the case #504-0404/23-5000.

3.1.2. Apparatus and materials

The sound stimuli consisted of two three-second simple tones of the same frequency but different timbre (i.e., waveforms) without known consensual basic taste associations, as found in a pre-test test to check for pre-existing associations (see the Supplementary Materials, S1 for details). One of the stimuli was a triangle-wave tone at 200 Hz, which showed no consensual taste associations in the main experiment. The other stimulus was a square-wave tone at 200 Hz, which contrary to our expectations and the pre-test presented an association with bitterness before the associative learning task in the main experiment. The sounds were generated in the open-source software Audacity v. 3.3.3 (https://www.audacityteam.org/). The sounds were normalized for loudness perception (with the triangle wave as a benchmark), by a sound researcher, the two authors, and a person without prior knowledge of the experiment. More specifically, the sound pressure of the square wave was reduced by 12 dB.

The stimuli corresponding to the emotional mechanism comprised two different emoji facial expressions (i.e., only eyebrows and smile) of two emojis extracted from the EmojiGrid (Toet et al., 2018), as originally used in Barbosa Escobar, Velasco, Byrne, and Wang (2022). The selected emojis were the lowest valence, lowest arousal (V1A1; henceforth called *sad*) and the highest valence, lowest arousal (V5A1; henceforth called *happy*). The *happy* emoji expression is considered to be related to sweetness given the shared positive valence of the expression and sweetness, whereas the sad emoji expression is considered to be related to bitterness due to their shared negative valence evocations (Spinelli et al., 2023).

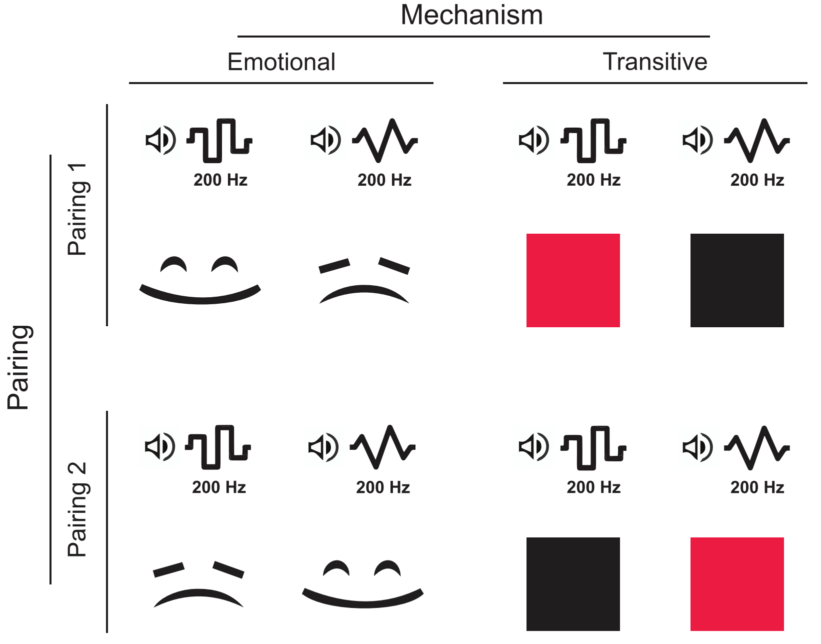

Fig. 1. Possible mappings in the associative learning task.
Note. The figure presents the four different mappings in the experiment, based on the interaction of mechanism and pairing. The images corresponding to the sound stimuli are for illustrative purposes only, as no visual cues were presented to represent sounds, and only a *Play* button was visible to play the sounds when necessary.

The stimuli relating to the transitive related to two colored squares, namely, a red one (R: 237, G:25, B:65) and a black one (R: 35, G:31, B:32). The former is associated with sweetness, whereas the latter is assumed to be associated with bitterness (Spence & Levitan, 2021).

3.1.3. Design

The experiment followed the same general associative learning design as Experiment 2 in Barbosa Escobar et al. (2023) while investigating two different mechanisms. More specifically, the experiment followed a 2 (Mechanism: emotional, transitive) × 2 (Pairing: pairing 1, pairing 2) × 2 (Time: before learning, after learning) mixed design, with mechanism and pairing as between-subjects factors and time as within-subject factor. Each participant was randomly assigned to one of four possible groups based on the interaction of mechanism and pairing. The pairings related to the specific emotional or transitive stimulus each sound was mapped to. Nevertheless, in the analysis, the pairing was analyzed based on the basic taste associated with the emotional or transitive stimuli with which each sound was paired. For instance, as stated earlier, the sweetness-related stimuli consisted of the *happy* emoji expression and the red square. The bitterness-related stimuli were the *sad* emoji expression and the black square. Fig. 1 presents all the possible mappings based on the experimental design.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

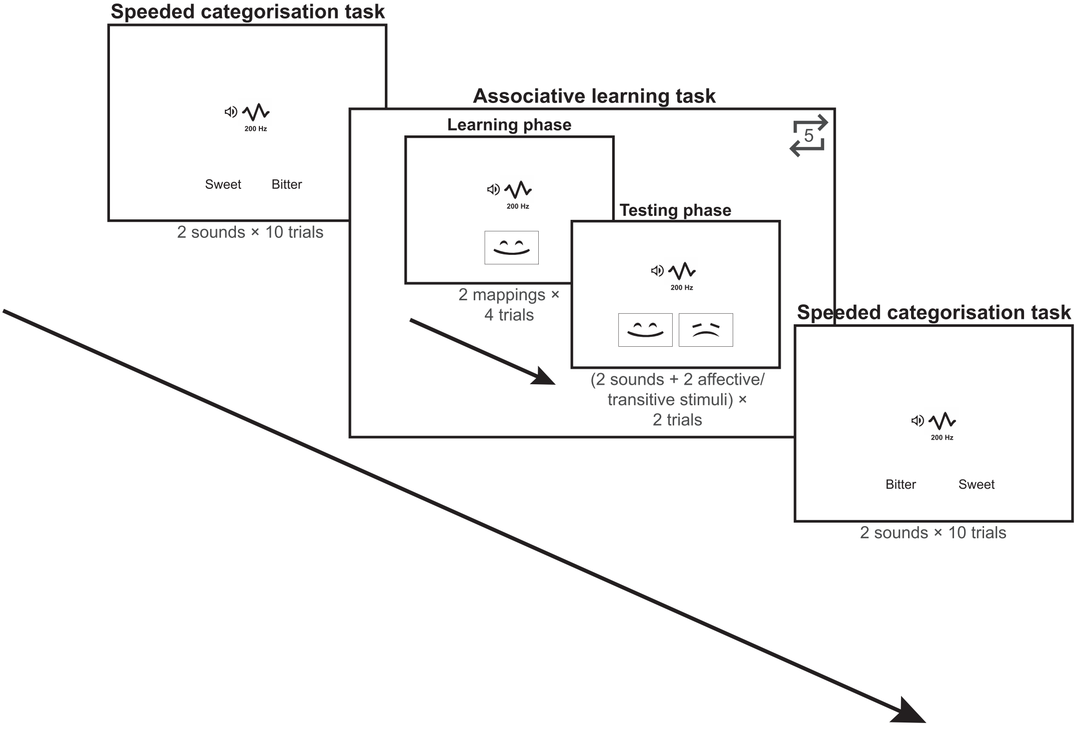

Fig. 2. Experimental procedure.
Note. The figure presents a schematic representation of the experimental procedure consisting of speeded categorization tasks before and after an associative learning task.

3.1.4. Procedure

The procedure was the same as Experiment 2 in Barbosa Escobar et al. (2023). The experiment was programmed on Gorilla (https://gorilla.sc/), and participants could only use either a desktop or a laptop to participate in the experiment. The experiment comprised two speeded categorization tasks, one before and the other after a self-paced associative learning task. Fig. 2 presents a schematic representation of the experimental design. Before starting the experiment, participants completed a sound calibration task. Here, a sine-wave pure tone at 150 Hz was played, and participants were asked to adjust the volume level of their system to a comfortable level. In addition, they heard the word carbonation (generated by a text-to-speech generator; https://freetools.textmagic.com/) and adjusted their volume level further.

Each speeded categorization task comprised 20 trials, consisting of 10 trials for each tone. In each trial, participants were tasked with categorizing, as promptly as possible, a tone (played automatically) as either sweet or bitter by pressing the “F” or “J” keys on their keyboard. The keys corresponded to the left-hand and right-hand sides of the screen, based on mappings provided to participants before starting each speeded categorization task. The mappings of the keys to basic tastes remained visible at the bottom of the screen during each trial.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The basic taste mapped to each key was counterbalanced between the initial and the last task for each participant.

The associative learning task consisted of two alternating phases, namely, a learning phase (composed of five learning blocks) and a testing phase (composed of five testing blocks), repeated five times. In the learning blocks, participants were exposed to two mappings presented in random order, one involving the triangle wave tone and the other involving the square wave tone. Hence, each tone was mapped to either an emotional (i.e., the *happy* or the *sad* emoji expression) or a transitive (i.e., the red or the black square) stimulus. The learning block consisted of eight trials (i.e., four trials per mapping). Each trial began with a fixation cross in the center of the screen visible for 500 ms, followed by an interstimulus pause of 100 ms. After each learning block, the testing block started, in which participants performed a matching task. They were presented with a target stimulus in the center of the screen, and they were tasked with selecting the other stimuli that corresponded to it based on the mappings experienced in the learning blocks. The target stimulus could be either a tone or an emotional/transitive stimulus, and participants had two response choices. Each testing block comprised eight trials, so each stimulus was presented twice. In the testing blocks, each trial began with a fixation cross followed by an interstimulus trial as in the learning trials. Once a response was given in each testing trial, response feedback appeared in the center of the screen for 200 ms, and it consisted of a green checkmark for a correct response or a red “X” for an incorrect response. The allocation of the stimuli to the response options was counterbalanced during each testing block.

3.1.5. Data analysis

The data analysis followed that of Experiment 2 in Barbosa Escobar et al. (2023), and both categorization responses and reaction times (RTs; presented in the Supplementary Materials, S1) were analyzed. In addition to the analyses conducted in Barbosa Escobar et al. (2023), to examine learning differences in the mappings corresponding to the two different mechanisms, we examined learning time and accuracy in the associative learning task by analyzing RTs and response accuracy in the testing phases of the associative learning task. Furthermore, we conducted a deeper analysis of the relative influence of the emotional and transitive mechanisms in the creation of novel crossmodal correspondences. All analyses were conducted using R statistical software (R Core Team, 2023). First, we visualized the proportion of categorization responses for each sound via mosaic plots. Then, we fit separate logistic generalized linear mixed models (GLMMs) for the emotional and the transitive mechanisms to analyze the probability that each tone was categorized as sweet. In the GLMMs, the dependent variable was specified as a binary variable indicating whether the tone was categorized as sweet. The interactions of pairing, tone, and time were specified as fixed effects, and participant ID was specified as a random effect. The GLMMs were conducted via the *glmer* function of the {lme4} R package (Bates, Mächler, Bolker, & Walker, 2015). Each GLMM was tested via likelihood ratio tests (LRT) against its corresponding null model consisting only of the random effect. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) were also computed. Differences in the estimated probabilities that the sounds were categorized as sweet were analyzed via Bonferroni-corrected pairwise comparisons of the estimated

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

marginal means resulting from the models via the *emmeans* functions of the {emmeans} R package (Lenth, 2023). Moreover, pseudo-R^2s were computed via the *rsq.glmm* function of the {rsq} R package (Zhang, 2022).

To further investigate the relative influence of the emotional and transitive mechanisms in the creation of novel correspondences for both types of tones, we fitted an additional comprehensive logistic GLMM using the full dataset. Similar to the earlier logistic GLMM on categorization responses, the binary variable indicating whether the tone was categorized as sweet was specified as the dependent variable. The interactions and main effects of mechanism, pairing, and time were specified as fixed effects, and tone and participant ID were specified as random effects. Given that we aimed to analyze the effect of both mechanisms for both tones, and following Judd, Westfall, and Kenny (2012), the tone was specified as a random effect. Subsequently, changes in the association probabilities from before to after the associative learning task under the two different pairings for the two mechanisms were computed. As an indication of the magnitude of the effects of the two mechanisms, the difference between the changes in association probability corresponding to the two mechanisms was computed separately for the sweet- and the bitter-related pairings.

To analyze learning time in the associative learning task, the RTs of the trials in the testing phases were first treated for outliers by winsorizing values more than three median absolute deviations (MADs) plus or minus the median of each participant. Subsequently, a GLMM with Gamma distribution and the identity link function on RT was conducted with mechanism (i.e., emotional vs. transitive) as the main fixed effect. The correct response, a binary variable indicating whether the participant responded correctly to the testing trial was added as a covariate. Moreover, stimulus type (i.e., sound, emoji, color patch) was added as a random effect, given that the dynamics of evaluating a sound clip compared to an emoji or a color patch were different. Participant ID was also added as a random effect. Pairwise comparisons were then computed. Furthermore, to further examine learning in the associative learning task, we conducted a similar Gamma GLMM with the specific mapping participants were exposed to, instead of mechanism, as the main fixed effect. Then, Bonferroni-corrected pairwise comparisons were computed. Learning accuracy in the associative learning task was analyzed by fitting a logistic GLMM on a binary variable indicating whether the participant responded to each testing trial correctly with mechanism as main fixed effect and stimulus type and participant ID as random effects.

3.2. Results

3.2.1. Categorization responses

The proportion of categorization responses based on the different factors for the two sounds is visualized via mosaic plots in Fig. 3. A visual inspection of the proportion of categorization responses showed that for the triangle wave tone, both the emotional and the transitive sweet-related pairings increased the number of times participants categorized the tone as sweet, although the effect of the emotional sweet-related pairings seemed to be larger. On the other hand, for the bitter-related pairings, neither emotional nor transitive seemed to change participants' categorization. As for the square wave tone, both the bitter-related and sweet-related

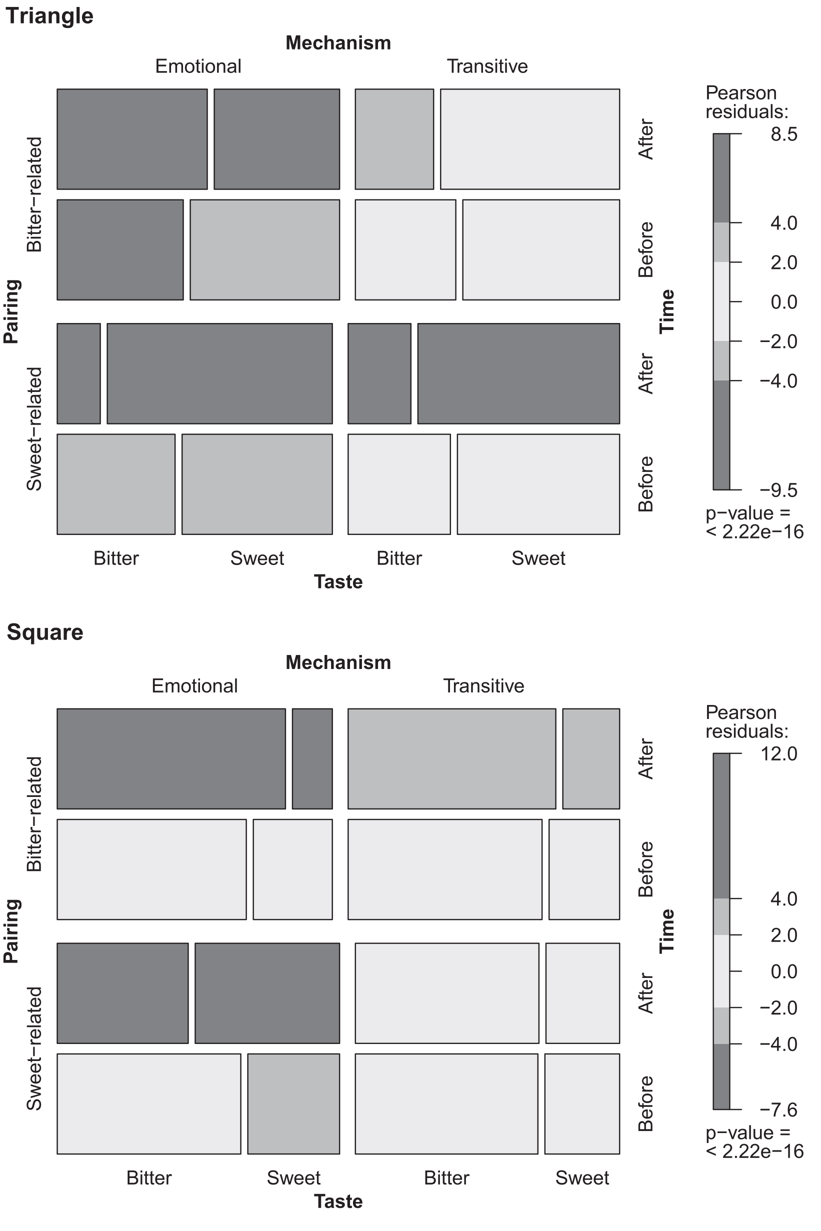

Fig. 3. Mosaic plots of categorization responses.
Note. The mosaic plots show the proportion in which each tone (*triangle* in the upper panel and *square* in the lower panel) was categorized as either sweet or bitter (indicated on the bottom side) before and after the associative learning paradigm (indicated on the right-hand side) comprising the different mechanisms (indicated on the top side) and pairings (indicated on the left-hand side). Pearson residuals indicate the strength of deviation from independence in standard deviations. A residual greater than 2 or less than −2 represents a significant departure at the 95% level.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 1
Results of the logistic GLMMs

	Emotional			Transitive		
Fixed effects	OR	CI	p	OR	CI	p
Intercept	0.42	[0.36, 0.49]	< .001	0.35	[0.29, 0.42]	< .001
$\text{Pairing}_{\text{Sweet-related}}$	1.20	[0.96, 1.49]	.108	1.13	[0.86, 1.48]	.382
$\text{Tone}_{\text{Triangle 200 Hz}}$	2.85	[2.30, 3.53]	< .001	4.55	[3.49, 5.92]	< .001
$\text{Time}_{\text{After}}$	0.42	[0.32, 0.54]	< .001	0.75	[0.59, 1.37]	.017
$\text{Pairing}_{\text{Sweet-related}} \times \text{Tone}_{\text{Triangle 200 Hz}}$	0.90	[0.66, 1.22]	.497	0.90	[0.59, 1.37]	.617
$\text{Pairing}_{\text{Sweet-related}} \times \text{Time}_{\text{After}}$	5.31	[3.83, 7.35]	< .001	1.31	[0.94, 1.83]	.109
$\text{Tone}_{\text{Triangle 200 Hz}} \times \text{Time}_{\text{After}}$	1.69	[1.22, 2.33]	.001	2.00	[1.44, 2.77]	< .001
$\text{Pairing}_{\text{Sweet-related}} \times \text{Tone}_{\text{Triangle 200 Hz}} \times \text{Time}_{\text{After}}$	1.10	[0.70, 1.73]	.670	1.07	[0.67, 1.69]	.781
Random effects						
σ^2	3.29			3.29		
Participants	154			148		
Observations	6,160			5,920		

Note. The table presents the results of the logistic GLMMs corresponding to the emotional and the transitive mechanisms. OR = odds ratios; CI = 95% confidence interval.

emotional pairings seemed to slightly increase the number of times participants categorized the tone as bitter.

The results of all the logistic GLMMs for the speeded categorization tasks are presented in Table 1. The model fit results are presented in Table 2. The estimated marginal mean probabilities resulting from the logistic GLMMs are presented in Fig. 4.

The logistic GLMM involving the emotional mechanism revealed significant two-way interaction effects of pairing and time and of tone and time and significant main effects of tone and time. Regarding the two-way interaction effect, as expected, for the triangle wave tone, the probability that it was categorized as sweet with the sweet-related pairings was greater after the associative learning task (0.84, 95% CI = [0.81, 0.87]) than before (0.56, 95% CI = [0.51, 0.61]; $p < .001$). In addition, the probability that the triangle wave tone was categorized as sweet with the bitter-related pairings was marginally but significantly lower, and just below the 50% threshold, after the associative learning task (0.46, 95% CI = [0.41, 0.50]; $p < .001$) than before (0.54, 95% CI = [0.50, 0.59]). Regarding the square wave tone, the probability that it was categorized as sweet under the sweet-related pairings was higher, and just above the 50% threshold, after the associative learning task (0.52, 95% CI = [0.48, 0.57]) than before (0.33, 95% CI = [0.30, 0.37]; $p < .001$. As for the square-wave tone with the bitter-related pairings, the probability that it was categorized as sweet was even lower after the associative learning task (0.15, 95% CI = [0.12, 0.18]) than before (0.29, 95% CI = [0.26, 0.33]; $p < .001$). When it comes to the main effects, overall, the probability that

Table 2
Model comparison results of logistic GLMMs

Model	Effects	AIC	BIC	*D*	LRT X^2	*p*	Pseudo-R^2s R^2_{Model}	R^2_{Fixed}	R^2_{Random}
Null	Participant	8,509	8,522				< .01		
Emotional	Pairing × Tone × Time	7,513	7,573	7	1,010.1	< .001	.16	.15	.03
Null	Participant	8,147	8,160				.03		
Transitive	Pairing × Tone × Time	7,046	7,106	7	1,114.8	< .001	.22	.17	.04

Note. The table presents the model fit results of the logistic GLMMs for the emotional and transitive mechanisms against their corresponding null models. The null models only included participants' IDs. AIC = Akaike information criterion; BIC = Bayesian information criterion; LRT = likelihood ratio test.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

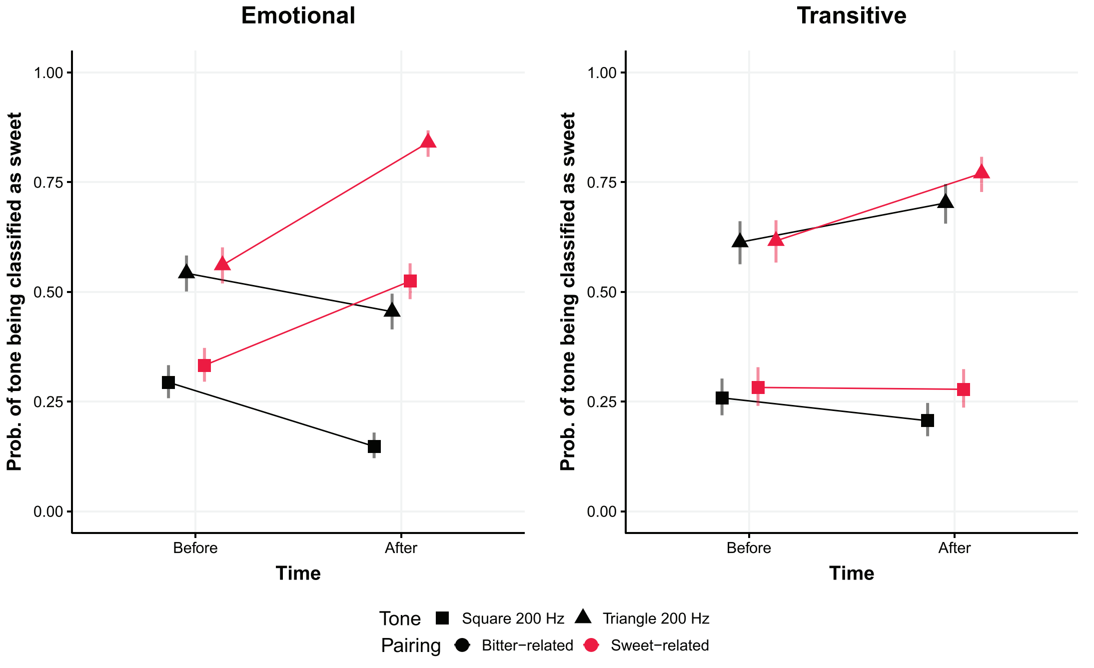

Fig. 4. Estimated marginal means from the logistic GLMM models.
Note. The figure presents the estimated marginal means deriving from the two logistic GLMMs corresponding to the main experiment. The *y*-axis represents the probability that the tones were categorized as sweet. The timing of the speeded categorization task, before or after the associative learning paradigm, is indicated in the *x*-axis. Error bars represent the 95% CIs resulting from the models.

the triangle wave tone (0.62, 95% CI = [0.60, 0.64]) was categorized as sweet was higher than that of the square wave tone (0.31, 95% CI = [0.29, 0.33]; $p < .001$). In addition, the probability that the tones, overall, were categorized as sweet was greater after the associative learning task (0.49, 95% CI = [0.46, 0.51]) than before (0.43, 95% CI = [0.41, 0.45]; $p < .001$).

The logistic GLMM involving the transitive path revealed significant main effects of tone and time and significant two-way interaction of tone and time (Table 2). In relation to the two-way interaction, as expected, with the triangle wave tone under the sweet-related pairings, the probability that it was categorized as sweet was higher after the associative learning task (0.77, 95% CI = [0.73, 0.81]) than before (0.62, 95% CI = [0.57, 0.66]; $p < .001$). However, contrary to our expectations, when it comes to the triangle wave tone and the bitter-related pairings, the probability that it was categorized as sweet was greater after the associative learning task (0.70, 95% CI = [0.66, 0.75]) than before (0.61, 95% CI = [0.56, 0.66]; $p < .001$). Regarding the square wave tone, with the bitter-related pairings, the probability that it was categorized as sweet was marginally lower after the associative learning task (0.21, 95% CI = [0.17, 0.25]) than before (0.26, 95% CI = [0.22, 0.30]; $p < .001$). There was not a significant difference in the probability that the square wave tone was categorized as sweet

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 3
Results of the comprehensive logistic GLMM

Fixed effects	OR	CI	p
Intercept	0.73	[0.25, 2.07]	.549
$\text{Mechanism}_{\text{Emotional}}$	0.94	[0.80, 1.11]	.488
$\text{Pairing}_{\text{Sweet-related}}$	1.07	[0.92, 1.25]	.400
$\text{Time}_{\text{After}}$	1.07	[0.92, 1.25]	.384
$\text{Mechanism}_{\text{Emotional}} \times \text{Pairing}_{\text{Sweet-related}}$	1.07	[0.86, 1.33]	.523
$\text{Mechanism}_{\text{Emotional}} \times \text{Time}_{\text{After}}$	0.54	[0.43, 0.67]	$< .001$
$\text{Pairing}_{\text{Sweet-related}} \times \text{Time}_{\text{After}}$	1.38	[1.11, 1.71]	.004
$\text{Mechanism}_{\text{Emotional}} \times \text{Pairing}_{\text{Sweet-related}} \times \text{Time}_{\text{After}}$	3.70	[2.72, 5.04]	$< .001$
Random effects			
σ^2	3.29		
Tones	2		
Participants	154		
Observations	6,160		

Note. The table presents the results of the comprehensive logistic GLMM aimed at examining the relative influence of the emotional and transitive mechanisms. OR = Odds Ratios; CI = 95% confidence interval.

before (0.28, 95% CI = [0.24, 0.33]) or after the associative learning task (0.28, 95% CI = [0.24, 0.32]; $p = .859$) with the sweet-related pairings. As per the main effects, overall, the triangle wave tone (0.68, 95% CI = [0.65, 0.71]) had a higher probability of being categorized as sweet than the square wave tone (0.26, 95% CI = [0.23, 0.28]; $p < .001$). Moreover, broadly speaking, the probability that the tones were categorized as sweet was higher after the associative learning task (0.49, 95% CI = [0.45, 0.52]) than before (0.44, 95% CI = [0.41, 0.46]; $p < .001$).

3.2.2. *Relative influence of mechanisms*

The analysis revealed a significant three-way interaction of mechanism, pairing, and time, with an odds ratio greater than one, indicating that the emotional mechanism had a greater influence on the probability of association compared to the transitive mechanism (Table 3). The results also revealed significant two-way interactions of mechanism and time and of pairing and Time. The post-hoc tests revealed significant differences between the emotional and the transitive mappings under the sweet- and the bitter-related pairings after the associative learning task (the estimated marginal means resulting from the model are presented in Table 4). Under the sweet-related pairings and after the associative learning task, the probability that the sounds were associated with sweetness was greater with the emotional mapping ($p < .001$). In addition, under the bitter-related pairings and after the associative learning task, the probability that the sounds were associated with sweetness was lower with the emotional mapping than with the transitive one ($p < .001$). There were no significant differences between the emotional and the transitive mechanisms before the associative learning task with either the sweet-related pairings or with the bitter-related pairings ($p > .050$).

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 4
Estimated marginal mean probabilities from the comprehensive logistic GLMM

					95% CI	
Pairing	Time	Mechanism	Probability	SE	LL	UL
Sweet-related	Before	Transitive	.437	0.132	0.189	0.720
		Emotional	.439	0.132	0.191	0.722
	After	Transitive	.533	0.133	0.256	0.791
		Emotional	.698	0.113	0.411	0.885
Bitter-related	Before	Transitive	.421	0.130	0.180	0.706
		Emotional	.406	0.129	0.171	0.694
	After	Transitive	.437	0.132	0.190	0.720
		Emotional	.284	0.109	0.107	0.568

Note. The table presents the estimated marginal mean probabilities resulting from the logistic GLMM examining the relative effect of the emotional and transitive mechanisms. The probabilities are broken down by pairing, time, and mechanism. CI = confidence interval; LL = lower level; UL = upper level.

To further examine the relative influence of the different mechanisms, we analyzed differences between the emotional and the transitive mappings in the change of association probability (with sweetness) from before the associative learning task to after separately under the sweet- and the bitter-related pairings. Under the sweet-related pairings, the emotional mechanism triggered a positive change in association probability of 25.9 percentage points from before to after the associative learning task, whereas the transitive mechanism yielded a positive change of 9.7 percentage points. Furthermore, under the bitter-related pairings, the emotional mechanism generated a negative change in association probability of 12.3 percentage points from before to after the associative learning task, whereas the transitive mechanism resulted in a positive change of 1.7 percentage points.

3.2.3. Learning time and accuracy

The analysis of the testing phases of the associative learning task revealed a significant effect of the learning mechanism of the mappings on RTs (Table 5). Participants were significantly faster in responding to the emotional mappings (3,662 ms, 95% CI = [3,656, 3,668]) than to the transitive ones (3,729 ms, 95% CI = [3,721, 3,737]; $p < .001$). Furthermore, looking deeper into the associative learning time, the analysis revealed a significant effect of the specific mappings on RTs (Table 6) and significant differences in RTs between each of the associative mappings (ps $< .001$), where the *Triangle tone–Happy emoji, Square tone–Sad emoji* mapping presented the lowest RTs (3,420 ms, 95% CI = [3,409, 3,432]), followed by the *Triangle tone–Red color, Square tone–Black color* mapping (3,693 ms, 95% CI = [3,678, 3,708]), and by the *Triangle tone–Black color, Square tone–Red emoji* mapping (3,766 ms, 95% CI = [3,747, 3,785]). Finally, the *Triangle tone–Sad emoji, Square tone–Happy emoji* mapping (3,889 ms, 95% CI = [3,875, 3,903]) had the highest RT.

In relation to learning accuracy, the analysis failed to reveal an effect of learning mechanism on the percentage of correct responses in the testing phase of the associative learning task. There was no difference in accuracy in responding to the emotional associative learning

Table 5
Results of the Gamma and logistic GLMMs per mechanism in the associative learning task

	RT			Accuracy		
Fixed effects	Estimate	CI	p	OR	CI	p
Intercept$_{\text{Mechanism: Emotional}}$	3,958.60	[3,953.48, 3,963.71	< .001	48.37	[35.47, 65.97]	< .001
Mechanism$_{\text{Transitive}}$	67.11	[62.53, 71.69]	< .001	1.26	[0.83, 1.91]	.277
Correct response	−593.50	[−597.18, −589.81	< .001			
Random effects						
σ^2	0.15			3.29		
Participants	310			310		
Stimulus type	3			3		
Observations	12,400			5,920		

Note. The table presents the results of the logistic and Gamma GLMMs on the responses and RTs in the testing phases of the associative learning task with mechanism of the mapping as fixed factor. OR = odds ratios; CI = 95% confidence interval.

Table 6
Results of the Gamma GLMM per specific mapping in the associative learning task

	RT		
Fixed effects	Estimate	CI	p
Intercept$_{\text{Mapping: Triangle-Happy emoji}}$	3,717.23	[3,708.14, 3,726.31]	<.001
Mapping$_{\text{Triangle-Sad emoji}}$	468.16	[463.02, 473.29]	<.001
Mapping$_{\text{Triangle-Red emoji}}$	272.54	[259.09, 286.00]	<.001
Mapping$_{\text{Triangle-Black colour}}$	345.83	[337.41, 354.25]	<.001
Correct response	−593.52	[−603.5, −583.53]	<.001
Random effects			
σ^2	0.16		
Participants	310		
Observations	12,400		

Note. The table presents the results of the logistic and Gamma GLMMs on the responses and RTs in the testing phases of the associative learning task with specific mapping as fixed factor. OR = odds ratios; CI = 95% confidence interval.

mechanism (0.98, 95% CI = [0.97, 0.99]) compared to the transitive ones (0.98, 95% CI = [0.98, 0.99]; $p = .277$).

4. Follow-up study

4.1. Methods

To investigate how the effects of the associative learning paradigm had changed over time, and test whether the novel association between the triangle wave tone and sweetness was still

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

present several weeks after the main experimental session, we conducted a follow-up study and tested these associations once again. The follow-up study consisted of a single speeded categorization task identical to the ones in the main experiment, consisting of 20 trials (i.e., 10 trials per tone) and with the same stimuli, conducted on Gorilla (https://gorilla.sc/). The 302 participants that took part in the main experiment and remained after the data cleaning process received a request, through Prolific (https://www.prolific.com/), to participate in the follow-up study. Participants completed the follow-up two months after the experiment, and they received GBP 0.40 in compensation. Out of the 302 invited, 229 (76%) participants completed the follow-up study (154 females, 147 males, 1 other), aged 19–40 years ($M_{age} = 31.38$ years, $SD_{age} = 5.81$). Of the participants in the follow-up, 117 had been exposed to the emotional learning paradigm, and 112 had been exposed to the transitive one in the main experiment. The analysis consisted of two separate logistic GLMMs for the categorization data, one for the emotional and one for the transitive mechanisms, identical to the main experiment, fitted to the data of the main experiment together with the follow-up. Furthermore, we added a third level to the time fixed effect corresponding to the follow-up.

4.2. Results

4.2.1. Categorization responses

The results of the logistic GLMMs including the follow-up data for both the emotional and transitive mechanisms are presented in Table 7. The model fit results are presented in Table 8. In the emotional mechanism, the results revealed significant main effects of tone, time (after), time (follow-up), as well as significant two-way interactions of pairing and time (after), pairing and time (follow-up), tone and time (after), and tone and time (follow-up). With the transitive path, the results showed significant main effects of tone and time (after), as well as significant two-way interactions of tone and time (after) and tone and time (follow-up).

With the emotional mechanism, the post-hoc tests revealed that the only significant difference in the probability of association between the follow-up and before participants were exposed to the associative learning task was when the triangle wave tone was paired with the sweet-related stimulus and when the square wave tone was paired with the bitter-related stimulus (Fig. 5). With the triangle wave tone and the sweet-related pairing, although the probability that it was associated with sweetness in the follow-up (0.69, 95% CI = [0.64, 0.74]; $p < .001$) was lower than right after the associative learning task (0.84, 95% CI = [0.81, 0.87]; $p < .001$), it was significantly higher than before participants completed the associative learning task (0.56, 95% CI = [0.52, 0.61]; $p < .001$), and above the 50% threshold. With the square wave tone and the bitter-related pairing, the probability that it was associated with sweetness in the follow-up (0.20, 95% CI = [0.16, 0.24]; $p < .001$) was lower than before the associative learning task (0.29, 95% CI = [0.26, 0.34]; $p < .001$), but it was not significantly different from than right after it (0.15, 95% CI = [0.12, 0.18]; $p = .052$).

Regarding the transitive path, the only significant difference in probability associations with respect to before the associative learning task was with the triangle wave tone under the bitter-related stimulus. The probability that the latter was associated with sweetness in the

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 7
Results of the logistic GLMMs including the follow-up

	Emotional			Transitive		
Fixed effects	OR	CI	*p*	OR	CI	*p*
Intercept	0.42	[0.35, 0.49]	<.001	0.36	[0.30, 0.42]	<.001
$\text{Pairing}_{\text{Sweet-related}}$	1.20	[0.96, 1.49]	.109	1.13	[0.88, 1.45]	.339
$\text{Tone}_{\text{Triangle 200 Hz}}$	2.85	[2.30, 3.54]	<.001	4.43	[3.48, 5.64]	<.001
$\text{Time}_{\text{After}}$	0.42	[0.32, 0.54]	<.001	0.75	[0.59, 0.95]	.018
$\text{Time}_{\text{Follow-up}}$	0.59	[0.46, 0.76]	<.001	1.11	[0.87, 1.43]	.402
$\text{Pairing}_{\text{Sweet-related}} \times \text{Tone}_{\text{Triangle 200 Hz}}$	0.90	[0.66, 1.23]	.501	0.90	[0.62, 1.30]	.571
$\text{Pairing}_{\text{Sweet-related}} \times \text{Time}_{\text{After}}$	5.32	[3.84, 7.37]	<.001	1.31	[0.94, 1.82]	.110
$\text{Pairing}_{\text{Sweet-related}} \times \text{Time}_{\text{Follow-up}}$	1.52	[1.08, 2.15]	.017	0.71	[0.50, 1.01]	.056
$\text{Tone}_{\text{Triangle 200 Hz}} \times \text{Time}_{\text{After}}$	1.69	[1.22, 2.33]	.001	1.97	[1.43, 2.73]	<.001
$\text{Tone}_{\text{Triangle 200 Hz}} \times \text{Time}_{\text{Follow-up}}$	2.12	[1.51, 2.96]	<.001	1.38	[0.98, 1.95]	.066
$\text{Pairing}_{\text{Sweet-related}} \times \text{Tone}_{\text{Triangle 200 Hz}} \times \text{Time}_{\text{After}}$	1.10	[0.70, 1.73]	.669	1.06	[0.67, 1.68]	.790
$\text{Pairing}_{\text{Sweet-related}} \times \text{Tone}_{\text{Triangle 200 Hz}} \times \text{Time}_{\text{Follow-up}}$	0.92	[0.58, 1.48]	.743	0.97	[0.59, 1.58]	.894
Random effects						
σ^2	3.29			3.29		
Participants	154			148		
Observations	8,500			8,160		

Note. The table presents the results of the logistic GLMMs corresponding to the emotional and the transitive mechanisms including the data of the main experiment and the follow-up. OR = odds ratios; CI = 95% confidence interval.

follow-up (0.71, 95% CI = [0.66, 0.76]; $p < .001$) was higher than before the associative learning task (0.61, 95% CI = [0.56, 0.66]; $p = .001$) and right after it (0.70, 95% CI = [0.65, 0.74]; $p = .001$).

5. General discussion

The present research aimed to investigate whether novel crossmodal correspondences between sounds and basic taste words could be established from the ground up, through an intermediate associative learning paradigm based on an emotional-mediation account and a transitive path (i.e., through colors), used in a past study (Barbosa Escobar et al., 2023). Beyond the findings of the latter study, the present paper also examined how RTs and accuracy in the associative learning paradigm differed across the mechanisms of the different

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Table 8
Model comparison results of logistic GLMMs including the follow-up

				LRT			Pseudo-R^2s		
Model	Effects	AIC	BIC	*d*	X^2	*p*	R^2_{Model}	R^2_{Fixed}	R^2_{Random}
Null	Participant	11,729	11,743				<.01		
Emotional	Pairing × Tone × Time	10,331	10,423	11	1,420	<.001	.16	0.15	0.01
Null	Participant	11,261	11,275				.03		
Transitive	Pairing × Tone × Time	9,798	9,888	11	1,485	<.001	.19	0.17	0.02

Note. The table presents the model fit results of the logistic GLMMs for the emotional and transitive mechanisms against their corresponding null models including the data of the main experiment and the follow-up. The null models only included participants' IDs. AIC = Akaike information criterion; BIC = Bayesian information criterion; LRT = likelihood ratio test.

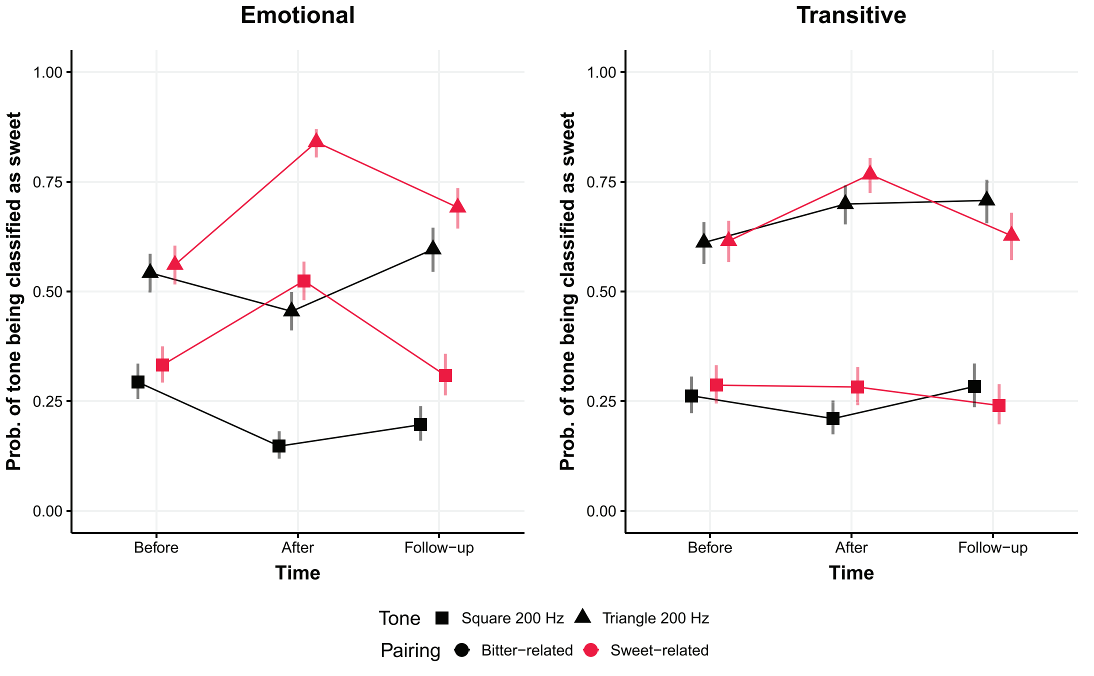

Fig. 5. Estimated marginal means from the logistic GLMM models including the follow-up.
Note. The figure presents the estimated marginal means deriving from the two logistic GLMMs including the data from the main experiment and the follow-up. The *y*-axis represents the probability that the tones were categorized as sweet. The timing of the speeded categorization task, before or after the associative learning paradigm and in the follow-up, is indicated in the *x*-axis. Error bars represent the 95% CIs resulting from the models.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

mappings, as well as the relative influence of both mechanisms. Moreover, we tested whether the novel correspondences formed were still present several weeks (i.e., two months) after the experimental session. To this end, two simple tones (i.e., a triangle wave tone at 200 Hz and a square wave tone at 200 Hz) were used as sound stimuli. Two emoji facial expressions (i.e., *happy* and *sad*) were used as stimuli for the emotional mediation account, and two colored squares (i.e., red and black) were used for the transitive path. Two months later after the experiment, we evaluated whether the associations were still present in the same participants who initially took part in the experiment. Broadly speaking, the results revealed that mapping a sound without previous consensual taste associations (i.e., the triangle wave tone) with a positive emotional stimulus can induce a crossmodal correspondence with sweetness. Furthermore, the results revealed a significant positive effect of RTs in the testing phases of the learning task across the different mappings on the change of probability association after the associative learning paradigm compared to before it. Importantly, albeit (unsurprisingly) less strongly, participants still exhibited this association two months after the main experiment.

The results of the present study revealed that the associative learning paradigm mapping the triangle wave sound, which was the sound that presented no consensual taste associations from the start, with the *happy* emoji facial expression significantly increased the probability (i.e., above 50%) that this sound was associated with sweetness. These results are in line with Wang et al. (2016) findings that sounds and tastes are matched based on the valence they evoke. However, the *sad* emoji expression did not have any effect on the association probability of the triangle wave sound with bitterness. Importantly, the findings from the present work start providing an answer to the long-held question regarding the lifespan of experimentally induced crossmodal associations (Spence, 2022). Past studies have demonstrated that novel crossmodal correspondences between dimensions that are not naturally correlated in the environment can be experimentally induced through short training sessions of less than an hour (Baier, Kleinschmidt, & Müller, 2006; Ernst, 2007; Flanagan, Bittner, & Johansson, 2008; Zangenehpour & Zatorre, 2010). These correspondences may arise from the fusion of sensory signals as coupling priors (Ernst, 2007; Körding et al., 2007; Parise & Spence, 2013). In addition, these crossmodal correspondences created in the laboratory can be regarded as a kind of statistical correspondence, as a correlation between stimuli or dimensions has been synthetically established (Parise & Spence, 2013). As Spence (2022) argued, these crossmodal correspondences induced experimentally are short-lasting or temporary and can be considered weak, and until this point, it was unknown how long they last in people's mind. It is worth noting that here, Spence (2022) referred to associations created by training participants to learn the specific mappings between the stimuli or dimensions involved in the correspondences per se. On the contrary, the novel association induced in the present work was created by establishing an underlying link between the stimulus on one side of the correspondence (i.e., a sound) to an emotional stimulus (i.e., a happy facial expression) related to the other side of the correspondence (i.e., sweetness). It is unclear which type of experimentally induced correspondence would have a longer lifespan. On the one hand, it could be argued that given that the correspondence created in the present study was created indirectly by establishing an intermediate link to a third dimension or stimuli, the resulting correspondence would be more short-lived than those created by training individuals to

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

directly associate the dimensions part of the correspondence. Nevertheless, it could also be argued that by establishing an intermediate link, the association could be more robust and last longer, as a correlational/statistical path has already been set for the correspondence to develop via a top-down approach. As past literature has shown, both bottom-up and top-down approaches play a role in the formation of crossmodal correspondences (Bolam, Boyle, Ince, & Delis, 2022; Getz & Kubovy, 2018; Peiffer-Smadja & Cohen, 2019). That being said, to the best of our knowledge, the present study is the first one to investigate the lifespan of experimentally induced crossmodal correspondences. We showed that the sound–taste correspondence created here, through an associative learning paradigm with a duration of less than 10 minutes, can last at least two months. Furthermore, it is possible that intermediate mappings that are learned more easily or effectively would yield longer-lasting novel crossmodal correspondences, as overall shorter learning times resulted in larger effects sizes of the associative learning paradigm in the present study, and hence could make the novel correspondences last longer.

Even though the results on learning accuracy did not reveal any difference between the mechanism of the different mappings, the results of the analyses on the RTs in the testing phases of the associative learning task revealed that participants responded the fastest to the mappings corresponding to the emotional mechanism. Furthermore, the specific mapping that participants took the shortest amount of time to respond related to the *Triangle tone–Happy emoji, Square tone–Sad emoji*. In turn, this latter mechanism was shown to be the one with the strongest relative effect in inducing a new correspondence. Together, these findings seem to suggest that easier learning may more effectively induce (stronger) novel crossmodal correspondences. In the present study, the emotional cues (i.e., emoji facial expressions) may have been learned more easily, as they are captured by the brain with higher priority, relative to more neutral stimuli such as color patches, given the large relevance in survival and well-being of the former (Brosch, Pourtois, & Sander, 2010).

When it comes to the square wave, contrary to our expectations, it was associated with bitterness fairly strongly from the start, which may have been caused by relative compatibility effects (Parise & Spence, 2013; Spence, 2019). Crossmodal correspondences tend to be relative judgements, and most times correspondences arise from explicit comparisons (Brunetti, Indraccolo, Del Gatto, Spence, & Santangelo, 2018; Deroy & Spence, 2013). Hence, given that participants were exposed to the two sounds (i.e., triangle wave and square wave) in the same experimental session, it is possible that participants judged the square wave sound as mbitter relative to the triangle wave sound. This association may be due to the different waveforms used (Knöferle & Spence, 2012), which at the same time may cause discrepancies in their perceived pitch (Hermes, 2023) and in their evoked valence (Baird, Parada-Cabaleiro, Fraser, Hantke, & Schuller, 2018). In addition, previous research has found that specific timbres, in the form of musical notes produced by different musical instruments, can give rise to different taste associations (Crisinel & Spence, 2010b). In this case, the *sad* emoji facial expression strengthened the square wave sound's association with bitterness, but the *happy* emoji expression did not significantly increase the square wave's association with sweetness. Regarding the transitive path, none of the mappings significantly influenced the associations of either of the sounds.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Put together, these results provided support to H_{1A}, but failed to support H_{1B}, thus demonstrating that an emotional mediation mechanism, as operationalized here, can induce crossmodal correspondences between sounds and sweetness by associating sounds with positively valenced emotional stimuli. Importantly, as the results revealed, the emotional mechanism had a higher relative influence in generating a novel crossmodal correspondence relative to the transitive mechanism. The change in association probability from before to after the associative learning task with the emotional mechanism was 16.2 basis points larger than the change with the transitive mechanism for the sweet-related pairings and 13.9 basis points larger for the bitter-related pairings. Notably, the results presented here revealed that the correspondence induced here (i.e., a triangle wave tone and sweetness created through a mapping of the tone to a positively valenced emoji facial expression) can last several weeks after the main experimental session. On the other hand, the results failed to support H_{2A} and H_{2B}, suggesting that a transitive path through colors may not be responsible for sound–taste correspondences, or at least, not the specific sound–taste correspondences investigated in the present study. That being said, these results do not discard the possibility that there may be other mechanisms underlying these correspondences. Here, it may be worth pointing out a potential similarity between the present findings and synesthesia, as even though these are different phenomena, past research has argued that an associative learning mechanism may underlie synesthetic experiences (M. R. Watson, Akins, Spiker, Crawford, & Enns, 2014; Yon & Press, 2014).

In the case of the *sad* emoji facial expression, although it strengthened the square wave tone's association with bitterness, it only brought the triangle wave tone's association with bitterness marginally above the 50% threshold right after the associative learning task, and later in the follow-up, it increased the tone's association with sweetness. It is possible that these effects were related to the affective loadings of the sad emoji expression. Bitterness triggers negative affective reactions, but these tend to be high in arousal (Barbosa Escobar & Wang, 2023; Lim & Green, 2007; Spinelli et al., 2023). In the present study, the *sad* emoji facial expression used is negatively valenced but low in arousal. Hence, it may not have evoked a robust enough association with bitterness in the associative learning mapping to trigger the taste correspondence in the neutral triangle wave tone, whereas it may have been sufficient to strengthen a sound (e.g., square wave) already associated with bitterness.

One reason behind the failure of the transitive associative learning paradigm could be that the black and red color stimuli may not have had strong enough associations with the target bitter and sweet tastes, respectively. Even though previous literature has found associations between bitterness and black hues (Spence et al., 2015; Spence & Levitan, 2021), recent research using a more fine-grained gamut of colors and a battery of methods revealed that the color associations with bitterness are less defined than previously thought, and people do not seem to strongly associate bitterness with the color black; although sweetness is associated with red and pink hues fairly consistently (Velasco, Barbosa Escobar, et al., 2023). Thus, in the present study, the black hue stimuli may not have truly evoked bitterness in the associative learning paradigm and hence may not have induced a crossmodal association between the sounds and bitterness.

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

An alternative explanation for the generation of the novel association between the triangle-wave tone and sweetness, but not bitterness, may lie in differences in the metabolic value of the tastants associated with the two basic tastes (Green, Nachtigal, Hammond, & Lim, 2012) that gives rise to the integration of odors and tastes (Prescott, 2015). Previous research has found that odor enhancement by taste is driven by the nutritive value of tastants (Linscott & Lim, 2016; Wang et al., 2019). More specifically, odor enhancement can be generated through sweet, salty, and umami, as these are found in nutritive tastants that contain critical compounds for the functioning of the human body (e.g., carbohydrates or proteins). However, the use of sour and bitter does not result in odor enhancement, given that these are present in non-nutritive or harmful tastants (e.g., toxins). In this way, sweetness (found in nutritive tastants), as opposed to bitterness (found in non-nutritive tastants), is more conducive to the generation of novel correspondences, especially when induced through an associative learning paradigm. Bridging the potential explanations posed thus far, based on the fact that sweetness is more important for survival compared to bitterness, as well as the larger relevance of emotional stimuli compared to more neutral ones, it is possible to argue that appraisal may be at play in the formation of crossmodal correspondences, whereby stimuli with higher-value outcomes for individuals are more easily or more strongly associated (Heyes, 2012; Morand-Ferron, 2017).

5.1. Limitations and future directions

Several limitations of the present study are worth noting. First, only two sounds were used, and they both consisted of simple tones, whereas most of auditory stimuli in real life are complex tones with different ranges of frequencies and differences in other psychoacoustic properties, which could yield different results (e.g., Hamilton-Fletcher et al., 2017). Nevertheless, using simple tones (e.g., triangle and square waves) permits the disentanglement of underlying mechanisms with fundamental psychoacoustic properties of sound. Future research should include different tones and complex sounds to test the robustness and generalizability of the findings presented here from both a fundamental and a more ecologically valid perspective. Another limitation comes from the lack of precise control of the volume at which participants heard the sound stimuli, which could have influenced the associations and the results. Although we mitigated the issue with a sound calibration step at the beginning of the experiment, this is a persistent issue in online studies.

Relating to the taste side of the correspondence, taste words and not actual tastants were used here. Using taste words allows for an easier operationalization of the experiment, lower costs, and the capability to have a larger sample size, compared to using real tastants, which should ideally be administered through precisely timed devices (e.g., gustometers) given the timed nature of the study. It cannot be discarded that using real tastants could influence the results found here. Although previous studies have found that associations using either taste words or actual tastants closely match each other, as found in the context of crossmodal correspondences between colors and tastes (Saluja & Stevenson, 2018) and between sounds and tastes (Wang et al., 2016). Concerning to the stimuli used for the associative learning paradigm, we only used two colors and two emoji expressions. Future research could further

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

investigate the effects found here using a wider array of both colors and emotional stimuli, as their effect could vary.

Regarding the follow-up study, while we examined participants' associations after two months of the experimental paradigm, it is not clear how long the correspondence between the triangle wave tone and sweetness will eventually last. Future studies can investigate the lifespan of experimentally induced crossmodal correspondences using even longer in time. Furthermore, our associative learning paradigm lasted less than 10 minutes. Further research could examine the effects of longer learning paradigms on the lifespan of these associations.

5.2. *Conclusions*

The findings presented here provide evidence that novel crossmodal correspondences between sounds and tastes can be created from the ground up through an associative learning paradigm mapping a sound without previous consensual associations with emotional stimuli. Notably, the present findings reveal that this novel correspondence can last at least several weeks. Nevertheless, the effect of the associative learning paradigm seems to work only with positively valenced emotional stimuli and with associations with sweetness. Furthermore, the findings here suggest that a transitive path through colors does not give rise to correspondences between tastes and sounds. Our work adds to the literature on crossmodal correspondences, and especially the highly inconsistent evidence in sound–taste correspondences, by deepening the understanding of how these correspondences may arise. Our work provides further evidence that sound–taste correspondences are mediated by affect. From a holistic perspective, the present research shows that ease of learning plays a critical role in the formation of crossmodal correspondences. Furthermore, our work provides insights, for the first time, regarding the lifespan of experimentally induced crossmodal correspondences. Finally, the present paper suggests that appraisal (i.e., in the sense of the importance of stimuli for survival) is an aspect that deserves more attention in the literature on crossmodal correspondences and multisensory integration. Together with Barbosa Escobar et al. (2023), the present study present the first step of a larger project aimed at making sense of general themes by testing a myriad of correspondences and novel aspects with this paradigm and combining their results. Such a project could further illuminate extant questions in the scientific literature on crossmodal correspondences, such as what really are crossmodal correspondences, how they differ from regular associations, and whether there are dimensions insusceptible by crossmodal correspondences (Parise, 2016). In addition, this project could generate entirely novel questions and topics that guide future research.

References

Adeli, M., Rouat, J., & Molotchnikoff, S. (2014). Audiovisual correspondence between musical timbre and visual shapes. *Frontiers in Human Neuroscience*, *8*, 352. https://doi.org/10.3389/fnhum.2014.00352

Attie-Picker, M., Venkatesan, T., Newman, G. E., & Knobe, J. (2024). On the value of sad music. *The Journal of Aesthetic Education*, *58*(1), 46–65. https://doi.org/10.5406/15437809.58.1.03

Auracher, J., Menninghaus, W., & Scharinger, M. (2020). Sound predicts meaning: Cross-modal associations between formant frequency and emotional tone in stanzas. *Cognitive Science*, *44*(10), e12906. https://doi.org/10.1111/cogs.12906

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

Baier, B., Kleinschmidt, A., & Müller, N. G. (2006). Cross-modal processing in early visual and auditory cortices depends on expected statistical relationship of multisensory information. *The Journal of Neuroscience*, *26*(47), 12260–12265. https://doi.org/10.1523/JNEUROSCI.1457-06.2006

Baird, A., Parada-Cabaleiro, E., Fraser, C., Hantke, S., & Schuller, B. (2018). The perceived emotion of isolated synthetic audio: The EmoSynth dataset and results. *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion*, *7*. https://doi.org/10.1145/3243274.3243277

Barbosa Escobar, F., Velasco, C., Byrne, D. V., & Wang, Q. J. (2022). The influence of emotional cues and anthropomorphism on product temperature expectations. *PsyArXiv*. https://doi.org/10.31234/osf.io/z6utv

Barbosa Escobar, F., Velasco, C., Byrne, D. V., & Wang, Q. J. (2023). Assessing mechanisms behind crossmodal associations between visual textures and temperature concepts. *Journal of Experimental Psychology: Human Perception and Performance*, *49*(6), 923–947. https://doi.org/10.1037/xhp0001131

Barbosa Escobar, F., & Wang, Q. J. (2023). Tasty vibes: Uncovering crossmodal correspondences between tactile vibrations and basic tastes. *Food Research International*, *174*, 113613. https://doi.org/10.1016/j.foodres.2023.113613

Barradas, G. T., & Sakka, L. S. (2022). When words matter: A cross-cultural perspective on lyrics and their relationship to musical emotions. *Psychology of Music*, *50*(2), 650–669. https://doi.org/10.1177/03057356211013390

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*. https://doi.org/10.18637/jss.v067.i01

Bolam, J., Boyle, S. C., Ince, R. A. A., & Delis, I. (2022). Neurocomputational mechanisms underlying crossmodal associations and their influence on perceptual decisions. *NeuroImage*, *247*, 118841. https://doi.org/10.1016/j.neuroimage.2021.118841

Boring, E. G., & Stevens, S. S. (1936). The nature of tonal brightness. *Proceedings of the National Academy of Sciences*, *22*(8), 514–521. https://doi.org/10.1073/pnas.22.8.514

Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., & Tervaniemi, M. (2011). A functional MRI study of happy and sad emotions in music with and without lyrics. *Frontiers in Psychology*, *2*, 308. https://doi.org/10.3389/fpsyg.2011.00308

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition & Emotion*, *24*(3), 377–400. https://doi.org/10.1080/02699930902975754

Brunetti, R., Indraccolo, A., Del Gatto, C., Spence, C., & Santangelo, V. (2018). Are crossmodal correspondences relative or absolute? Sequential effects on speeded classification. *Attention, Perception, & Psychophysics*, *80*(2), 527–534. https://doi.org/10.3758/s13414-017-1445-z

Coutinho, E., & Cangelosi, A. (2011). Musical emotions: Predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. *Emotion*, *11*(4), 921–937. https://doi.org/10.1037/a0024700

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

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., & 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

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Deroy, O., Crisinel, A.-S., & Spence, C. (2013). Crossmodal correspondences between odors and contingent features: Odors, musical notes, and geometrical shapes. *Psychonomic Bulletin & Review*, *20*(5), 878–896. https://doi.org/10.3758/s13423-013-0397-0

Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic Bulletin & Review*, *20*(4), 643–664. https://doi.org/10.3758/s13423-013-0387-2

Eerola, T., Ferrer, R., & Alluri, V. (2012). Timbre and affect dimensions: Evidence from affect and similarity ratings and acoustic correlates of isolated instrument sounds. *Music Perception*, *30*(1), 49–70. https://doi.org/10.1525/mp.2012.30.1.49

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

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of Experimental Behavior*, *42*(1), 143–157. https://doi.org/10.1901/jeab.1984.42-143

Flanagan, J. R., Bittner, J. P., & Johansson, R. S. (2008). Experience can change distinct size-weight priors engaged in lifting objects and judging their weights. *Current Biology*, *18*(22), 1742–1747. https://doi.org/10.1016/j.cub.2008.09.042

Getz, L. M., & Kubovy, M. (2018). Questioning the automaticity of audiovisual correspondences. *Cognition*, *175*, 101–108. https://doi.org/10.1016/j.cognition.2018.02.015

Green, B. G., Nachtigal, D., Hammond, S., & Lim, J. (2012). Enhancement of retronasal odors by taste. *Chemical Senses*, *37*(1), 77–86. https://doi.org/10.1093/chemse/bjr068

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*, *55*(3), 1121–1140. https://doi.org/10.3758/s13428-022-01862-z

Guedes, D., Vaz Garrido, M., Lamy, E., Pereira Cavalheiro, B., & Prada, M. (2023). Crossmodal interactions between audition and taste: A systematic review and narrative synthesis. *Food Quality and Preference*, *107*, 104856. https://doi.org/10.1016/j.foodqual.2023.104856

Hamilton-Fletcher, G., Witzel, C., Reby, D., & Ward, J. (2017). Sound properties associated with equiluminant colours. *Multisensory Research*, *30*(3–5), 337–362. https://doi.org/10.1163/22134808-00002567

Hermes, D. J. (2023). Pitch perception. In D. J. Hermes (Ed.), *The perceptual structure of sound: Current research in systematic musicology* (Vol. 11, pp. 381–448). Cham: Springer. https://doi.org/10.1007/978-3-031-25566-3_8

Heyes, C. (2012). Simple minds: A qualified defence of associative learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1603), 2695–2703. https://doi.org/10.1098/rstb.2012.0217

Ikeda, K. (1909). New seasonings. *Journal of the Chemical Society of Tokyo*, *30*(8), 820–836.

Ikeda, K. (2002). New seasonings. *Chemical Senses*, *27*(9), 847–849. https://doi.org/10.1093/chemse/27.9.847

Jaquet, L., Danuser, B., & Gomez, P. (2014). Music and felt emotions: How systematic pitch level variations affect the experience of pleasantness and arousal. *Psychology of Music*, *42*(1), 51–70. https://doi.org/10.1177/0305735612456583

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Juslin, P. N., Liljeström, S., Västfjäll, D., & Lundqvist, L.-O. (2010). How does music evoke emotions? Exploring the underlying mechanisms. In P. N. Juslin (Ed.), *Handbook of music and emotion: Theory, research, applications* (pp. 605–642). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oSo/9780199230143.003.0022

Kashima, H., & Hayashi, N. (2011). Basic taste stimuli elicit unique responses in facial skin blood flow. *PLoS One*, *6*(12), e28236. https://doi.org/10.1371/journal.pone.0028236

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Knoeferle, K. M., Woods, A., Käppler, 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

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

Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience*, *15*(3), 170–180. https://doi.org/10.1038/nrn3666

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B., & Shams, L. (2007). Causal inference in multisensory perception. *PLoS One*, *2*(9), e943. https://doi.org/10.1371/journal.pone.0000943

Le Groux, S., & Verschure, P. F. M. J. (2012). Subjective emotional responses to musical structure, expression and timbre features: A synthetic approach. In *9th International Symposium on Computer Music Modelling and Retrieval (CMMR 2012)*. https://ccrma.stanford.edu/~slegroux/publications/pubs/CMMR2012.pdf

Lenth, R. (2023). *emmeans: Estimated marginal means, aka least-squares means (version 1.8.5) [R software package]* [Computer software]. https://CRAN.R-project.org/package=emmeans

Lim, J., & Green, B. G. (2007). The psychophysical relationship between bitter taste and burning sensation: Evidence of qualitative similarity. *Chemical Senses*, *32*(1), 31–39. https://doi.org/10.1093/chemse/bjl033

Linscott, T. D., & Lim, J. (2016). Retronasal odor enhancement by salty and umami tastes. *Food Quality and Preference*, *48*, 1–10. https://doi.org/10.1016/j.foodqual.2015.08.004

Ma, W., & Thompson, W. F. (2015). Human emotions track changes in the acoustic environment. *Proceedings of the National Academy of Sciences*, *112*(47), 14563–14568. https://doi.org/10.1073/pnas.1515087112

Morand-Ferron, J. (2017). Why learn? The adaptive value of associative learning in wild populations. *Current Opinion in Behavioral Sciences*, *16*, 73–79. https://doi.org/10.1016/j.cobeha.2017.03.008

Myers-Schulz, B., Pujara, M., Wolf, R. C., & Koenigs, M. (2013). Inherent emotional quality of human speech sounds. *Cognition & Emotion*, *27*(6), 1105–1113. https://doi.org/10.1080/02699931.2012.754739

Parise, C., & Spence, C. (2013). Audiovisual cross-modal correspondences in the general population. In J. Simner & E. Hubbard (Eds.), *Oxford handbook of synesthesia* (pp. 790–815). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199603329.013.0039

Parise, C. V. (2016). Crossmodal correspondences: Standing issues and experimental guidelines. *Multisensory Research*, *29*(1–3), 7–28. https://doi.org/10.1163/22134808-00002502

Peiffer-Smadja, N., & Cohen, L. (2019). The cerebral bases of the bouba-kiki effect. *NeuroImage*, *186*, 679–689. https://doi.org/10.1016/j.neuroimage.2018.11.033

Prescott, J. (2015). Multisensory processes in flavour perception and their influence on food choice. *Current Opinion in Food Science*, *3*, 47–52. https://doi.org/10.1016/j.cofs.2015.02.007

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

R Core Team. (2023). *R: A language and environment for statistical computing* [Computer software]. https://www.r-project.org/

Reuter, C., Jewanski, J., Saitis, C., Czedik-Eysenberg, I., Siddiq, S., Kruchten, S., & Oehler, M. (2018). Colors and timbres—Consistency and tendencies of color-timbre mappings in non-synesthetic Individuals. In *Proceedings of the 34th Jahrestagung Der Deutschen Gesellschaft Für Musikpsychologie: Musik Im Audiovisuellen Kontext*. https://doi.org/10.13140/RG.2.2.31766.19526

Robin, O., Rousmans, S., Dittmar, A., & Vernet-Maury, E. (2003). Gender influence on emotional responses to primary tastes. *Physiology & Behavior*, *78*(3), 385–393. https://doi.org/10.1016/S0031-9384(02)00981-2

Rousmans, S. (2000). Autonomic nervous system responses associated with primary tastes. *Chemical Senses*, *25*(6), 709–718. https://doi.org/10.1093/chemse/25.6.709

Saarikallio, S. (2011). Music as emotional self-regulation throughout adulthood. *Psychology of Music*, *39*(3), 307–327. https://doi.org/10.1177/0305735610374894

Sabbatino, V., Troiano, E., Schweitzer, A., & Klinger, R. (2022). “Splink” is happy and “phrouth” is scary: Emotion intensity analysis for nonsense words. In *Proceedings of the 12th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis*, (pp. 37–50). https://doi.org/10.48550/arXiv.2202.12132

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Saluja, S., & Stevenson, R. J. (2018). Cross-modal associations between real tastes and colors. *Chemical Senses*, *43*(7), 475–480. https://doi.org/10.1093/chemse/bjy033

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971–995. https://doi.org/10.3758/s13414-010-0073-7

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

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

Spence, C. (2015b). Multisensory flavor perception. *Cell*, *161*(1), 24–35. https://doi.org/10.1016/j.cell.2015.03.007

Spence, C. (2017). *Gastrophysics: The new science of eating*. New York: Penguin.

Spence, C. (2019). On the relative nature of (pitch-based) crossmodal correspondences. *Multisensory Research*, *32*(3), 235–265. https://doi.org/10.1163/22134808-20191407

Spence, C. (2020). Temperature-based crossmodal correspondences: Causes and consequences. *Multisensory Research*, *33*(6), 645–682. https://doi.org/10.1163/22134808-20191494

Spence, C. (2022). Exploring group differences in the crossmodal correspondences. *Multisensory Research*, *35*(6), 495–536. https://doi.org/10.1163/22134808-bja10079

Spence, C., & Deroy, O. (2012). Crossmodal correspondences: Innate or learned? *I-Perception*, *3*(5), 316–318. https://doi.org/10.1068/i0526ic

Spence, C., & Di Stefano, N. (2022). Coloured hearing, colour music, colour organs, and the search for perceptually meaningful correspondences between colour and sound. *I-Perception*, *13*(3), 204166952210928. https://doi.org/10.1177/20416695221092802

Spence, C., & Levitan, C. A. (2021). Explaining crossmodal correspondences between colours and tastes. *I-Perception*, *12*(3), 204166952110182. https://doi.org/10.1177/20416695211018223

Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J., & Deroy, O. (2015). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour*, *4*(1), 23. https://doi.org/10.1186/s13411-015-0033-1

Spinelli, S., Pierguidi, L., Gavazzi, G., Dinnella, C., De Toffoli, A., Prescott, J., & Monteleone, E. (2023). Skin conductance responses to oral stimuli: The role of taste quality and intensity, and personality traits. *Food Quality and Preference*, *109*, 104917. https://doi.org/10.1016/j.foodqual.2023.104917

Sun, X., Li, X., Ji, L., Han, F., Wang, H., Liu, Y., Chen, Y., Lou, Z., & Li, Z. (2018). An extended research of crossmodal correspondence between color and sound in psychology and cognitive ergonomics. *PeerJ*, *6*, e4443. https://doi.org/10.7717/peerj.4443

Toet, A., Kaneko, D., Ushiama, S., Hoving, S., de Kruijf, I., Brouwer, A.-M., Kallen, V., & van Erp, J. B. F. (2018). EmojiGrid: A 2D pictorial scale for the assessment of food elicited emotions. *Frontiers in Psychology*, *9*, 2396. https://doi.org/10.3389/fpsyg.2018.02396

Tomasik-Krótki, J., & Strojny, J. (2008). Scaling of sensory impressions. *Journal of Sensory Studies*, *23*, 251–266. https://doi.org/10.1111/j.1745-459X.2008.00153.x

Velasco, C., Barbosa Escobar, F., Spence, C., & Olier, J. S. (2023). The taste of colours. *Food Quality and Preference*, *112*, 105009. https://doi.org/10.1016/j.foodqual.2023.105009

Velasco, C., Corradi, G., & Motoki, K. (2023). Harmony of senses: Exploring the impact of sound aesthetic features' on taste imagery. *Food Quality and Preference*, *111*, 104992. https://doi.org/10.1016/j.foodqual.2023.104992

Velasco, C., Michel, C., Youssef, J., Gamez, X., Cheok, A. D., & Spence, C. (2016). Colour–taste correspondences: Designing food experiences to meet expectations or to surprise. *International Journal of Food Design*, *1*(2), 83–102. https://doi.org/10.1386/ijfd.1.2.83_1

Wan, X., Woods, A. T., Van Den Bosch, J. J. F., McKenzie, K. J., Velasco, C., & Spence, C. (2014). Cross-cultural differences in crossmodal correspondences between basic tastes and visual features. *Frontiers in Psychology*, *5*, 1365. https://doi.org/10.3389/fpsyg.2014.01365

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Wang, Q. J., Mielby, L. A., Junge, J. Y., Bertelsen, A. S., Kidmose, U., Spence, C., & Byrne, D. V. (2019). The role of intrinsic and extrinsic sensory factors in sweetness perception of food and beverages: A review. *Foods*, *8*(6), 211. https://doi.org/10.3390/foods8060211

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), 2041669515622001. https://doi.org/10.1177/2041669515622001

Watson, M. R., Akins, K. A., Spiker, C., Crawford, L., & Enns, J. T. (2014). Synesthesia and learning: A critical review and novel theory. *Frontiers in Human Neuroscience*, *8*, 98. https://doi.org/10.3389/fnhum.2014.00098

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

Whiteford, K. L., Schloss, K. B., Helwig, N. E., & Palmer, S. E. (2018). Color, music, and emotion: Bach to the blues. *I-Perception*, *9*(6), 204166951880853. https://doi.org/10.1177/2041669518808535

Yon, D., & Press, C. (2014). Back to the future: Synaesthesia could be due to associative learning. *Frontiers in Psychology*, *5*, 702. https://doi.org/10.3389/fpsyg.2014.00702

Zangenehpour, S., & Zatorre, R. J. (2010). Crossmodal recruitment of primary visual cortex following brief exposure to bimodal audiovisual stimuli. *Neuropsychologia*, *48*(2), 591–600. https://doi.org/10.1016/j.neuropsychologia.2009.10.022

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494

Zhang, D. (2022). *R-squared and related measures* (version 2.5) [Computer software]. https://cran.r-project.org/package=rsq

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supplementary Materials

15516709, 2024, 3, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/cogs.13421 by Copenhagen Business School, Wiley Online Library on [26/03/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License