Food Quality and Preference 111 (2023) 104992

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Harmony of senses: Exploring the impact of sound aesthetic features' on taste imagery

Carlos Velasco [a,c,*], Guido Corradi [b], Kosuke Motoki [c]

[a] *Centre for Multisensory Marketing, Department of Marketing, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway*
[b] *Department of Psychology, Universidad Camilo José Cela, Madrid, Spain*
[c] *Department of Management, The University of Tokyo, Tokyo, Japan*

ARTICLE INFO

Keywords:
Sound
Aesthetics
Taste imagery
Crossmodal correspondences

ABSTRACT

People reliably associate visual aesthetic features such as curvature and symmetry with tastes. In the present study, considering the transitive hypothesis of crossmodal correspondences, we evaluated whether these findings would extend to the relationship between sound aesthetic features and tastes, and whether feature-based congruency or affective priming would explain the influence of melodies on taste imagery. In Experiment 1, we evaluated how people associate different melody profiles (balanced vs. unbalanced, smooth vs. jagged, symmetrical vs. asymmetrical, simpler vs. more complex) with different tastes (sweet, sour, bitter, salty), as well as the melodies' associated fluency, valence, and arousal. Smooth and complex melodies were perceived as sweeter, jagged and unbalanced melodies as sourer, asymmetrical and jagged melodies as more bitter, and jagged and balanced melodies as saltier. In Experiment 2, we selected the most strongly associated aesthetic sound dimension with tastes, namely contour, and evaluated whether crossmodal congruency, based on crossmodal correspondences, or affective priming would influence people's sensory and hedonic imagery associated with sweet and sour foods. In the imagery tasks, the participants showed higher sour ratings in the sourness task and higher sweetness ratings in the sweetness task. In the sweet imagery task, smooth melodies led to sweeter and less sour food imagery, whereas jagged melodies in the sourness task led to more sour and less sweet food imagery. These results favour the crossmodal congruency explanation rather than the affective alternative, which we ponder in the general discussion.

1. Introduction

Crossmodal correspondences have been defined as the often-surprising associations that exist between features or dimensions across the senses (Spence, 2011; Motoki et al., 2022; Walker, 2016). The key words here are features or dimensions. Different from, say, semantic congruence, which indicates that our brains associate information from different senses as a function of a common identity or meaning (e.g., the way a dog look and its bark, or the way a strawberry looks and its flavour, Laurienti et al., 2004; Letts et al., 2022; Knoeferle et al., 2016), crossmodal correspondences are somewhat more abstract associations between features that may happen beyond identities and meanings (Deroy & Spence, 2016). Take for instance, the case of tastes such as sweetness and sourness and shape features such as curvature or symmetry (Juravle et al., 2022; Salgado-Montejo et al., 2015; Turoman et al., 2018). Tastes and shapes are not specific to a single identity or meaning (e.g., they do not appear to belong exclusively to just one or a specific group of food objects), yet people tend to associate sweet taste with curvature and symmetry and bitter and sour taste with angularity and asymmetry, instead (Velasco et al., 2015; Spence, 2022). Correspondences such as these ones have been documented in many combinations of senses and dimensions (see Marks, 1978; Parise, 2016; Spence, 2011; Spence, 2020; Spence & Levitan, 2021, for reviews).

It has been suggested that some of the aforementioned correspondences may occur as a result of a common affect evoked by the individual sensory dimensions they correspond to (e.g., Marks, 1996; Spence, 2020; Velasco et al., 2015). Taking the same example of taste-shape correspondences (i.e., sweet-curvature/symmetry), it has been shown that people tend to prefer sweet tastes, relative to the other tastes (Velasco et al., 2015) and shape features such curvature and symmetry, relative to their angularity and asymmetry counterparts (Chuquichambi et al., 2022). Whilst evidence is still needed to provide unequivocal

---

* Corresponding author at: Department of Marketing, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway.
E-mail address: carlos.velasco@bi.no (C. Velasco).

https://doi.org/10.1016/j.foodqual.2023.104992
Received 30 May 2023; Received in revised form 9 September 2023; Accepted 10 September 2023

support for this mechanism, it is generally accepted that that affect might, at least, explain in part this correspondence (Lee & Spence, 2022).

Perhaps for this very same idea, research on crossmodal correspondences on the one hand, and empirical aesthetics (i.e., the scientific study of aesthetics, beauty, and artistic experiences through data analysis, e.g., Carbon, 2018), on the other, have shared certain interests and theoretical frameworks (Velasco et al., 2016). Indeed, given that empirical aesthetics deals in part with how and why is that people prefer certain stimuli, aesthetic features (that is, features which appear to evoke tendencies of preference, e.g., Palmer et al., 2013) such as curvature, symmetry, balance, and complexity have been evaluated in the context of crossmodal (affective) correspondences (e.g., Juravle et al., 2022; Salgado-Montejo et al., 2015; Turoman et al., 2018).

Relevant to the present research, aesthetic features are not necessarily exclusive to the visual domain (Diessner et al., 2021). Indeed, researchers have suggested that dimensions such as symmetry, balance, contour, and complexity are not necessarily vision-specific and can be experienced, among others, in auditory stimuli (Clemente et al., 2021). Notably, however, research on crossmodal correspondences and aesthetic features has focused, in most cases, on the visual sense (Juravle et al., 2022; Salgado-Montejo et al., 2015; Turoman et al., 2018). For this reason, the first aim of the present study consisted of evaluating the extent to which what has been documented in terms of taste-(aesthetic) shape crossmodal correspondences, extends to taste-(aesthetic) sound correspondences.

We build on the transitive hypothesis of crossmodal correspondences to develop our predictions. This hypothesis states that if a dimension A (e.g., sweetness) corresponds to a dimension B (e.g., visual symmetry) in another sensory modality, and B, in turn, corresponds to a dimension C (e.g., auditory symmetry) in a third sensory modality, there will also be a crossmodal correspondence between A (e.g., sweetness) and C (e.g., auditory symmetry) (Deroy et al., 2013; Fields et al, 1984). As such, we postulate that those correspondences that have been documented between visual aesthetic features and tastes (e.g., balance, symmetry, curvature, smoothness, and sweetness on the one hand, and unbalance, asymmetry, angularity, roughness, and bitterness and sourness, on the other), will extend to the auditory modality.

The second aim of this research consisted of contrasting two possible explanations of congruency effects associated with sound-taste correspondences (Reinoso Carvalho et al., 2020). On the one hand, it has been suggested that crossmodally corresponding stimuli may lead to the crossmodal effect of sound on taste ratings. This is based on the concept of crossmodal (affective) correspondences, and in this particular case, the correspondence between specific sounds and features in other senses, such as taste (e.g., Knöferle & Spence, 2012). If this is the case, following the correspondences documented, presenting, say, symmetrical and asymmetrical sounds would enhance sweetness and sourness ratings, respectively, possibly, due to their shared affective connotations. On the other hand, an alternative explanation may be that the crossmodal effect of sound on taste ratings is not based on crossmodal correspondences but, instead, on affective priming. This idea is based on research suggesting that the hedonic evaluation of sounds (e.g., valence, liked/disliked) is what drives the effect of sound on taste evaluations (e.g., Wang & Spence, 2016), In this scenario, symmetrical sounds, given their positive valence, would lead to increased (more positive evaluations toward) sweetness and bitterness, and asymmetrical sounds would result in lower ratings for both sweetness and bitterness.

In order to evaluate our predictions, we present two experiments. The first experiment was designed to assess the extent to which the same crossmodal correspondences observed between specific visual aesthetic dimensions and tastes extend to those between auditory aesthetic dimensions. The second experiment capitalised on the results of Experiment 1 to evaluate how different types of congruency would influence people's taste imagery judgements.

2. Experiment 1

2.1. Methods and materials

Participants. 102 British participants, with English as their first language, were recruited from Prolific Academic (https://www.prolific.ac/) and took part in the experiment. However, data from three participants were removed as they answered "no" when asked about whether they had normal hearing. The data analysed was based on the remaining 99 participants (73 females, 26 males), aged 18–70 (*Mean* age = 33.97, *SD* = 12.24). The experiment was programmed on Qualtrics (http://qualtrics.com/) and lasted, on average, 27.3 min (SD = 13.62). Each participant was compensated for their time.

Materials. The stimuli consisted of melodies that represented four aesthetic dimensions, namely, balance (balance vs. unbalance), contour (smooth vs. jagged), symmetry (symmetrical vs. asymmetrical), and complexity (simpler vs. more complex). We present a brief definition of each of them, following Clemente and colleagues' work (2021). Balance refers to the distribution of events and the position of the climax within a melody. Balanced melodies exhibit even event distributions and place their climaxes centrally, whereas unbalanced melodies tend to concentrate events towards their ends. Melodic contour is established by the width of intervals, creating either jagged or smooth melodies. Jagged rhythms introduce abrupt changes to the contour. Symmetry encompasses mirror-reversed melodic patterns, resulting in symmetrical melodies functioning as palindromes, whereas asymmetrical melodies lack retrograde repetition. Complexity is associated with the quantity and diversity of notes; more intricate melodies comprise a wider array of notes, while simpler ones rely on repetitive patterns. Stimuli subsets are classified according to distinct attributes to mitigate unintended variations (for further information, refer to Clemente and colleagues' work, 2021). Five melodies of each level of the dimensions were selected from a musical stimuli dataset developed by Clemente and colleagues (2020), for a total of 40 melodies used in the experiment. We selected more than one stimulus for each level of the aesthetic dimensions in order to obtain representativeness. The specific stimuli used were selected considering that they had characteristically salient ratings and included the following: B1, B4, B9, B16, an B40 for balance; B13, B22, B39, B44, and B45 or unbalance; C3, C9, C17, C34, and C40 for smooth; C8, C15, C18, C24, and C27 for jagged; S1, S5, S9, S20, and S31 for symmetry; S8, S19, S21, S37, and S43, for asymmetry; K1, K9, K16, K33, and K45 for simpler; and K4, K12, K18, K43, and K47, for more complex. The stimuli are available at the original authors' OSF page (https://osf.io/bfxz7/).

The melodies were evaluated in 100-point visual analogue scales anchored with "not at all" and "very much" for taste (To what extent do you associate this sound with the following tastes: Bitter, sweet, salty, sour), with "easy" and "hard" for fluency (The process of studying this sound was), with "negative" and "positive" for valence (What this sound makes me feel is:), and with "not aroused at all" and "very aroused" for arousal (When I listen to this sound I feel).

Procedure. The experiment followed a within-participants experimental design with factors aesthetic dimension (balance, contour, symmetry, and complexity) and level (the two poles of the dimension). Before the experiment started, the participants were given the general instructions of the study ("*We are interested in understanding how people evaluate different musical motifs (sounds) related to various tastes. If you decide to take part, we will play different sounds and will ask you to respond to a few questions about them.*") and were asked to agree to a standard consent before taking part in the study. Next, they were asked to report their age and gender. Followed by that, they were asked to report whether they had a normal sense of hearing (answers "*yes*" and "*no*") and to do a sound check task in which the word "carbonation" was pronounced, and they had to write it into a text box. Right after that, each participant was presented with the 40 melodies in random order followed by the individual visual analogue scales for taste, fluency, valence, and arousal in random order, across participants.

Analysis. We analysed the effects of sounds aesthetics features (jagged vs smooth; symmetrical vs asymmetrical, complex vs simple; balanced vs unbalanced) on the taste ratings (sour, salty, sweet and bitter) with linear mixed effects models (Hox, 2010; Snijders & Bolker, 2012). This method accounts simultaneously for the within-subject variability as well as for the item variability (Baayen, Davidson, & Bates, 2008). This modelling is especially suitable for understanding human associations and preferences, which may vary from person to person, and from stimuli to stimuli (Silvia, 2007). In setting the model up, we followed Barr et al. (2013) guidelines. They suggest modelling the maximal random effect's structure justified by the experimental design, which, in addition to avoiding the loss of power and reducing Type-I error, enhances the possibility of generalising the results to other participants and stimuli. However, when using maximal models, the process of parameter estimation will occasionally fail to produce a solution. In these cases, we simplified the model's structure.

All analyses were carried out within the R environment for statistical computing (R Core Team, 2016). We used the *lmer* function from the lme4 package, with Kenward-Roger approach to compute degrees of freedom and p-values corrected using the multivariate t-distribution adjustment to produce the inferential statistics and p-values as recommended by Luke (2016). The models were primarily set up to study the impact of the sound aesthetic features on the taste ratings. However, in order to account for the effect of arousal, valence, and fluency in each participant, we included them in the models. We set up a model to predict each taste category (sour, salty, sweet, and bitter) for each aesthetic feature class (contour, complexity, balance and symmetry), summing 16 models. We coded the smooth, symmetrical, complex, and balanced predictors as 0.5, and their opposites as -0.5. A positive coefficient is understood to indicate a positive correlation between smooth, symmetric, complex, and balanced stimuli with the predicted value. Conversely, a negative coefficient indicates a connection between jagged, asymmetric, simpler, and unbalanced auditory features and the predicted value. The models which converged took participants and stimuli as random effects. Continuous variables were mean-centred.

3. Results and discussion

We included the score of each participant rating of valence, fluency, and arousal ratings as additional variables in the models. Here, we describe their general values. Valence mean scores for each category were placed at about the middle of the scale (Range = [49.93; 57.91]) with standard deviations covering about the 11% of the scale range. Fluency scores were lower than valence ones (Range = [36.94; 41.09]) with higher standard deviations covering about the 20% of the possible scale range. Arousal mean scores were the lowest (Range = [23.83; 27.52]) with standard deviations resembling those of fluency and covering about the 21% of the scale possible range (see Table 1 for detailed statistics). Regarding the taste ratings, mean scores were lower for sour taste (Range = [23.96; 29.65]) and highest for sweet (Range = [31.57; 49.13]). Salty (Range = [26.38; 32.14]) and bitter had similar scores (Range = [21.36; 30.59]). All standard deviations covered about 20% of the scales' possible range.

In Table 2, we report the coefficients for the 16 models predicting each taste by the aesthetic feature and additional variables (valence, fluency and arousal) as well as 95% confidence intervals (CI) and *p*-values.

Our results showed that auditory stimuli involving symmetry and complexity were not statistically significantly related to any taste ratings once the additional variables were considered. We found that balanced (vs. unbalanced) stimuli are positively related to sourness (b = 4.21, 95% CI [0.78; 7.63], p =.016). When we considered contour, we found that smooth ones were positively related to sweet taste evaluations (b = 10.35, 95% CI [0.41; 20.29], p =.041), but also negatively related to the sour (b = -5.80, 95% CI [-11.09; −0.50], p =.032) and bitter taste ratings (b = -6.96, 95% CI [-13.00; −0.93], p =.024). Regarding the additional variables, valence was the predominant predictor of each taste, as it was statistically significantly related to all taste evaluations.

The results of Experiment 1 are visually summarised in Fig. 1. We inspected the coefficients magnitude and their CI, which revealed a number of differences. For sweetness, a difference was observed in contour and complexity, whereby the smooth melodies were judged as sweeter than the jagged and the more complex stimuli as sweeter than the simpler. For sourness, a difference was observed for melody balance and contour such that the jagged stimuli were sourer than the smooth melodies, and the balanced melodies sourer than the unbalanced. For bitterness, differences were observed in relation to contour, symmetry, and complexity. In particular, the asymmetrical, more complex, and jagged melodies were more strongly associated with bitter than their counterparts. For salty, differences were observed in terms of balance and contour. The jagged and balanced stimuli were perceived as saltier than their counterparts. A difference in fluency was only observed for balance.

Note: Panels shows the models' mean predictions with 95% CI. Asterisks denote a statistically significant coefficient.

The results of Experiment 1 suggest that not all sound aesthetic features, as captured in the melodies used in the experiment, are related in the same way with tastes, as visual aesthetic features relate to them. Indeed, the most consistent aesthetic feature matching all tastes, that also matched research on vision-taste correspondences, was contour (e.g., Velasco et al., 2015). As such, in Experiment 2, we moved on to assess whether congruency would influence taste ratings, using a taste imagery task (Schifferstein, 2009).

4. Experiment 2

4.1. Methods and materials

Participants. The analyses were based on 104 British participants (55 males, 48 males, 1 preferred not to answer, with English as their first language), aged 18–68 (*Mean* age = 30.0 *SD* = 9.0) recruited from Prolific Academic (https://www.prolific.ac/). There were no discarded data from the participants as all of them correctly responded to the sound test (see Experiment 1). The experiment was programmed on Qualtrics (http://qualtrics.com/) and lasted, on average, 17 min (SD = 14.35). Each participant was compensated according to their time.

Table 1
Descriptive statistics for Experiment 1. We present the means and the standard deviations of each rating scale as a function of the features of the melodies.

Sound feature	Assessment Sourness	Sweetness	Saltiness	Bitterness	Valence	Fluency	Arousal
Smooth	23.96 (18.17)	49.13 (22.65)	26.38 (19.22)	21.36 (17.19)	57.91 (11.56)	39.18 (20.19)	27.53 (21.83)
Jagged	31.53 (20.51)	34.68 (17.84)	30.40 (19.84)	31.32 (20.80)	51.33 (11.96)	39.18 (20.97)	26.07 (21.65)
Complex	26.81 (19.14)	41.04 (20.16)	29.57 (21.08)	25.65 (18.90)	55.01 (11.43)	40.31 (21.85)	26.17 (21.00)
Simpler	28.21 (18.57)	31.57 (18.02)	30.49 (20.35)	30.59 (20.37)	49.93 (11.35)	39.95 (20.96)	23.83 (19.36)
Symmetric	26.49 (18.61)	43.83 (20.00)	27.22 (19.61)	26.59 (17.65)	55.55 (12.18)	37.44 (21.23)	25.42 (20.73)
Asymmetric	25.15 (18.63)	46.07 (21.51)	27.92 (19.40)	27.92 (19.40)	58.67 (11.46)	39.23 (21.98)	25.42 (20.73)
Balanced	29.65 (20.06)	41.79 (20.33)	32.14 (21.75)	27.15 (19.33)	56.94 (13.61)	36.94 (20.58)	29.15 (23.64)
Unbalanced	26.06 (19.74)	42.72 (18.91)	27.89 (19.57)	26.05 (18.10)	55.65 (11.21)	41.09 (21.47)	27.52 (21.83)

Table 2
Inferential models for Experiment 1 showing 16 models in which tastes are predicted by auditory features.

Models predicting sweetness	**Smooth**	**Balanced**	**Complex**	**Symmetric**
	Beta **95 % CI** ***p* value**	**Beta** **95 % CI** ***p* value**	**Beta** **95 % CI** ***p* value**	**Beta** **95 % CI** ***p* value**
Coefficient of the feature	**10.35** **[0.41; 20.29]** ***p* =.041**	−2.01 [-16.07; 12.05] *p* =.779	6.59 [-2.22; 15.41] *p* =.143	−0.74 [-14.02; 12.54] *p* =.913
Additional variables				
Valence	**0.61** **[0.53; 0.69]** ***p* <.001**	**0.47** **[0.39; 0.55]** ***p* <.001**	**0.53** **[0.45; 0.62]** ***p* <.001**	**0.49** **[0.41; 0.57]** ***p* <.001**
Fluency	−0.02 [-0.09; 0.05] *p* =.543	**−0.08** **[-0.14; −0.01]** ***p* =.028**	−0.02 [-0.09; 0.05] *p* =.560	**−0.12** **[-0.19; −0.05]** ***p* =.001**
Arousal	0.06 [-0.02; 0.14] *p* =.154	**0.1** **[0.01; 0.18]** ***p* =.021**	0.07 [-0.01; 0.16] *p* =.096	**0.11** **[0.03; 0.20]** ***p* =.008**
Models predicting sourness	**Smooth**	**Balanced**	**Complex**	**Symmetric**
	Beta 95 % CI *p* value	Beta 95 % CI *p* value	Beta 95 % CI *p* value	Beta 95 % CI *p* value
Coefficient of the feature	**−5.80** **[-11.09; −0.5]** ***p* =.032**	**4.21** **[0.78; 7.63]** ***p* =.016**	−0.22 [-2.85; 2.42] *p* =.871	0.62 [-3.05; 4.28] *p* =.742
Additional variables				
Valence	**−0.29** **[-0.37; −0.21]** ***p* <.001**	**−0.29** **[-0.37; −0.21]** ***p* <.001**	**−0.29** **[-0.37; −0.21]** ***p* <.001**	**−0.3** **[-0.38; −0.23]** ***p* <.001**
Fluency	**0.08** **[0.01; 0.14]** ***p* =.023**	**0.12** **[0.06; 0.19]** ***p* <.001**	**0.07** **[0.00; 0.14]** ***p* =.049**	**0.08** **[0.02; 0.15]** ***p* =.014**
Arousal	**0.11** **[0.03; 0.19]** ***p* =.010**	**0.17** **[0.08; 0.25]** ***p* <.001**	**0.11** **[0.03; 0.2]** ***p* =.010**	0.05 [-0.03; 0.14] *p* =.188
Models predicting saltiness	**Smooth**	**Balanced**	**Complex**	**Symmetric**
	Beta 95 % CI *p* value	Beta 95 % CI *p* value	Beta 95 % CI *p* value	Beta 95 % CI *p* value
Coefficient of the feature	−3.07 [-9.19; 3.05] *p* =.325	4.83 [-2.5; 12.16] *p* =.196	−0.7 [-4.02; 2.61] *p* =.677	−0.74 [-14.02; 12.54] *p* =.913
Additional variables				
Valence	**−0.16** **[-0.25; −0.08]** ***p* <.001**	**−0.17** **[-0.25; −0.08]** ***p* <.001**	**−0.09** **[-0.17; 0.00]** ***p* =.041**	**0.49** **[0.41; 0.57]** ***p* <.001**
Fluency	**0.08** **[0.02; 0.15]** ***p* =.015**	**0.16** **[0.09; 0.23]** ***p* <.001**	0.05 [-0.02; 0.12] *p* =.181	**−0.12** **[-0.19; −0.05]** ***p* =.001**
Arousal	0.09 [0; 0.17] *p* =.046	0.17 [0.08; 0.26] *p* < .001	0.09 [0; 0.18] *p* =.060	0.11 [0.03; 0.2] *p* =.008
Models predicting bitterness	**Smooth**	**Balanced**	**Complex**	**Symmetric**
	Beta 95 % CI *p* value	Beta 95 % CI *p* value	Beta 95 % CI *p* value	Beta 95 % CI *p* value
Coefficient of the feature	**−6.96** **[-13.00; −0.93]** ***p* =.024** *p* value	1.67 [-4.79; 8.12] *p* =.613	−3.15 [-7.45; 1.15] *p* =.151	2.04 [-4.57; 8.66] *p* =.545
Additional variables				
Valence	**−0.48** **[-0.56; −0.41]** ***p* <.001**	**−0.4** **[-0.48; −0.32]** ***p* <.001**	**−0.39** **[-0.47; −0.31]** ***p* <.001**	**−0.39** **[-0.46; −0.31]** ***p* <.001**
Fluency	0.04 [-0.02; 0.1] *p* = 0.176	**0.07** **[0; 0.14]** ***p* = 0.041**	−0.03 [-0.09; 0.04] *p* = 0.471	0.04 [-0.03; 0.10] *p* = 0.238
Arousal	**0.13** **[0.05; 0.2]** ***p* =.001**	**0.15** **[0.06; 0.23]** ***p* <.001**	0.08 [-0.01; 0.17] *p* =.079	**0.10** **[0.02; 0.18]** ***p* =.014**

Note: Each column represents a mixed model and the estimates of each predictor with inferential p value and confidence interval. Statistically significant results (p <.05) are shown in bold. Models account for item and stimuli variation. Smooth, balanced, complex. and symmetrical auditory features are coded as positive value.

Materials and procedure. The experiment followed a 2 × 2 within participants design experimental design with melody contour (smooth vs. jagged) and imagined food taste (sweet and sour) as factors. The stimuli consisted of melodies that represented the aesthetic dimensions of contour used in Experiment 1. In total, 10 auditory stimuli were used, five smooth and five jagged. All stimuli were presented in two conditions, that is, sweet and sour imagery conditions, for a total of 10 trials in each, randomly presented. In each trial, the participants were told to "Please listen to the following short music while imagining eating something that is characteristically sweet/sour", and then asked to evaluate the imagined food in 100-point visual analogue scales (e.g., "How vividly do you imagine the following tastes in this food") anchored with "not at all" and "very much" for sweetness or sourness, respectively. They were also asked to evaluate the extent to which the process of imagining the food was easy or difficult (fluency), and the extent to which the imagined food made them feel positive or negative (valence) and aroused (arousal). All but one of these questions were presented in 100-point visual analogue scales. A software error recorded fluency as a 10-point visual analogue scale. To overcome this caveat, we multiplied the results associated with this scale by 10 to make it a comparable scale. This procedure is known to control the potential bias introduced, especially in our case (10 anchors, no floor or ceiling effect) (Kuhlmann et al, 2017).

Analysis. We performed the same analytic strategy used in Experiment 1, however, we changed the predictors and outcomes. Here, we performed one model for each imagery task (imagine sweet or imagine sour task). The models predicted the assessments of taste by the interaction of the aesthetic feature (jagged or smooth) with the type of assessment (sweet and sour). We added easiness of imagery, valence, and arousal as additional variables.

5. Results and discussion

We included the scores of each participant's ratings of valence, fluency (easiness of imagery), and arousal as additional variables in the models. Valence mean scores for each category were placed at about the end range of the scale (Range = [4.9; 87.1]) with standard deviations covering about the 17 % of the scale range. Easiness of imagery scores spanned the entire range (Range = [0.0; 100.0]), with standard deviations covering about 19% of the possible scale range. Arousal mean scores exhibited some floor effects (Range = [0.0; 77.8]), indicating that a significant portion of the data had the lowest value, with standard deviations covering around 14% of the scale's possible range. While mean valence and easiness of imagery were positioned in the middle of

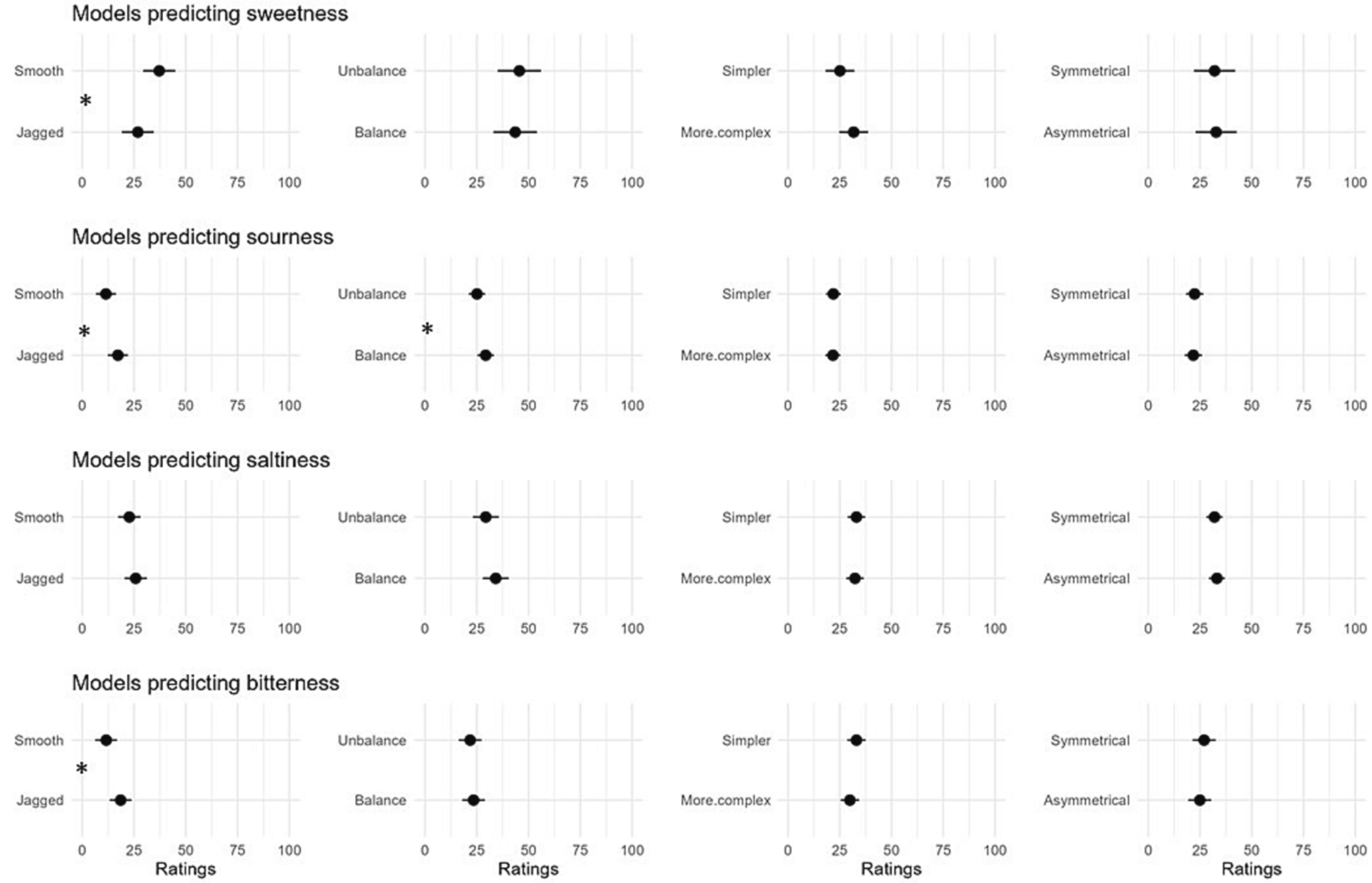

Fig. 1. Visualisation of the results of Experiment 1.

the range, mean arousal was below the midpoint of the range.

In the sourness imagery task, both sour and sweet assessments covered nearly the entire range of possible values ($\text{Range}_{\text{sourness}} = [2.8; 96.6]$, $\text{Range}_{\text{sweetness}} = [0.0; 82.8]$). Regarding the sweetness imagery task, we found a similar pattern ($\text{Range}_{\text{sourness}} = [0.0; 77.0]$, $\text{Range}_{\text{sweetness}} = [5.6; 95.2]$). All standard deviations were around 20% of the scale's possible range. We found a consistent pattern in which the sour ratings were higher in the sourness imagery task and sweetness ratings higher in the sweetness imagery task (see Table 3). Moreover, we found that in the sourness imagery task the highest means was in the jagged-sourness evaluation while the lowest was in the sweetness-jagged. As for the sweet imagery task, we observed higher means in smooth-sweet pairings and the lowest means in smooth-sour pairings.

Table 3
Descriptive statistics for Experiment 2.

Feature	**Sourness imagery task** **Sourness**	 **Sweetness**	**Sweetness imagery task** **Sourness**	 **Sweetness**
	M ***(SD)***	***M*** ***(SD)***	***M*** ***(SD)***	***M*** ***(SD)***
Jagged	53.3 (18.6)	25.3 (17.7)	26.1 (20.9)	51.7 (19.1)
Smooth	44.3 (21.1)	35.9 (22.4)	18.7 (17.7)	64.3 (18.7)
Additional variables	**Sourness imagery task** ***M*** ***(SD)***		**Sweetness imagery task** ***M*** ***(SD)***	
Valence	52.1 (17.5)		58.9 (16.4)	
Fluency	49.7 (20.5)		45.2 (22.2)	
Arousal	33.8 (27.9)		37.4 (23.8)	

Table 4. presents the coefficients associated with the two models (one for sour imagery task, one for sweet imagery task) predicting each taste evaluation (sweet and sour) as a function of the interaction between the aesthetic feature (jagged or smooth) and additional variables (valence, fluency, and arousal), as well as 95% CIs and *p*-values.

Regarding the results for the sweet imagery task, we found a significant interaction effect (see Table 4). In the sourness imagery task, the interaction between assessment type and feature was also statistically significant. We calculated predicted means (see Fig. 2) to clarify the interaction. Results showed that in the sourness task, jagged auditory stimuli significantly resulted in higher sourness imagery and less sweetness imagery than the smooth. In the sweetness imagery task, smooth stimuli resulted in significantly higher sweetness imagery and less sourness imagery than the jagged ones. Regarding the additional variables, we found that fluency was negatively, and significantly, associated with the assessment in both models. All other additional variables (valence and arousal assessments) were statistically significant associated to the taste assessment in both models.

Note: Panels shows the models' mean predictions with 95% CI.

6. General discussion

We conducted two studies designed to assess the extent to which the crossmodal correspondences between visual aesthetic features and tastes would extend to sound aesthetic features and tastes, as well as possible mechanisms underlying these correspondences. In Experiment 1, we studied whether the crossmodal correspondences between specific visual aesthetic dimensions and tastes would also apply to auditory aesthetic dimensions.

Based on the transitive hypothesis of crossmodal correspondences, as

Table 4
Inferential models for Experiment 2 predicting taste assessments by auditory feature in each task.

	Model for each task	
	Sweetness task	**Sourness task**
Coefficient	Beta 95 % CI *p* value	Beta 95 % CI *p* value
Taste [Sweet is coded with positive value]	**35.61** **[27.51; 35.04]** ***p* <.001**	**−18.16** **[-20.30; −16.60]** ***p* <.001**
Feature [Smooth is coded with positive value]	1.17 [-0.93; 3.27] *p* =.275	0.51 [-1.67; 2.69] *p* =.645
Taste × Feature	**19.90** **[15.80; 24.00]** ***p* <.001**	**19.63** **[15.34; 23.92]** ***p* <.001**
Additional variables	**Sweetness task**	**Sourness task**
Valence	**0.07** **[0.01; 0.13]** ***p* =.020**	**0.10** **[0.03; 0.17]** ***p* =.004**
Fluency	**−0.09** **[-0.13; −0.04]** ***p* <.001**	**−0.12** **[-0.17; −0.07]** ***p* <.001**
Arousal	**0.12** **[0.07; 0.18]** ***p* <.001**	**0.10** **[0.04; 0.17]** ***p* <.001**

Note: each column represents a mixed model and the estimates of each predictor with inferential p value and confidence interval. Statistically significant results ($p < .05$) are shown in bold. Models account for item and stimuli variation.

escribed by Deroy et al. (2013), we posited that the previously documented correspondences between visual aesthetic features (such as balance, symmetry, curvature, smoothness, and sweetness) and tastes (such as bitterness, sourness, and sweetness) would extend to the auditory domain (Experiment 1) (see also Fields et al. 1984).We did not find evidence in support of the association between sound symmetry and complexity, on the one hand, and taste words, on the other, when controlling for factors such as valence and arousal. However, melody balance was found to be statistically significant related to sourness, while smooth melodies were positively related to sweet taste but negatively related to sour and bitter tastes. Among the various features, contour (smooth vs. jagged) consistently matched the tastes, aligning with previous research on correspondences between vision and taste. Valence was found to be the main predictor of taste ratings. Based on these results, we suggest that not all melody aesthetic features match tastes in the same way as visual aesthetic features, at least, as embodied in the stimuli used in Experiment 1. Indeed, it might be the case that, while the features studied here are direct descriptors of visual properties, they might function as metaphors in the context of melodies (Marks, 1996). Consequently, people use these attributes to describe sounds; however, in melodies without additional information, other sound-specific characteristics might influence the participants' taste evaluations, at least for those attributes where the visual–taste correspondence does not extend to the auditory-taste correspondence.

In Experiment 2, we used the findings from Experiment 1 to assess how congruency vs. affective priming would impact people's taste imagery judgments. In both the sweet and sour imagery tasks, there were significant interaction effects between assessment type and feature, with jagged melodies being perceived as more sour and less sweet than smooth ones in the sourness task, and smooth stimuli being perceived as more sweet and less sour than jagged ones in the sweetness task. The ease of imagery was negatively associated with ratings in both tasks, while valence and arousal were significant additional variables in both tasks. These results favour the crossmodal correspondence congruency explanation rather than the affective alternative, as the data indicates a stronger association between specific sounds and taste ratings. This provides evidence in support of the idea that melodies that crossmodally correspond to a given taste can influence taste imagery, rather than the affective properties of the melody influencing them via affective priming (see Reinoso-Carvalho et al., 2020, for alternative results in the context of music).

Our results reveal sound aesthetic properties-taste correspondences. Prior research has demonstrated that visual aesthetic features such as curvature, symmetry, balance, and complexity have been evaluated in the context of crossmodal correspondences (e.g., Juravle et al., 2022; Salgado-Montejo et al., 2015; Turoman et al., 2018). We extend these findings to the relationship between aesthetic sound features of melodies and tastes. For example, smooth melodies are more rated as sweeter than jagged sounds, while jagged sounds are more associated with sourness than smooth melodies. The findings are also extended to taste imagery. That is, smooth (jagged) sounds can induce imagining eating something that is characteristically sweet (sour) than jagged (smooth) auditory stimuli. Together, our findings reveal the role of

Fig. 2. Visualisation of the results of Experiment 2 depicting the interaction between task and auditory features.

melody aesthetics on taste association and taste imagery.

One limitation is that our research did not test the downstream effects on consumer preference. Prior research has shown that mental imagery can influence consumer preference (e.g., Elder, & Krishna, 2022). For example, in one study, advertisements describing multi- (vs. single-) sensory experiences induce positive consumer experiences (e.g., positive taste evaluations). Given our findings, smooth melodies can induce sweetness imagery, which may possibly enhance consumer preference (e.g., intention to eat sweet foods). Moreover, our findings can be applied to soundscapes when a store sell sweet or sour foods. Our results reveal that smooth and jagged sounds are congruent with sweet and sour foods, respectively.

Our findings contribute to literature on sound-taste correspondences. Research on sound-taste correspondences has revealed that sound features are reliably matched with tastes (Knöferle & Spence, 2012; Guedes et al., 2023). For example, pitch (Crisinel & Spence, 2009, 2010a, 2010b, 2012; Reinoso Carvalho et al., 2016; Velasco et al., 2014; Wang et al., 2016), timbre (Crisinel & Spence, 2010b; Qi et al., 2020), speech sounds (Motoki et al., 2020; Pathak, & Calvert, 2020), musical pieces (Kontukoski et al., 2015; Guedes et al., 2022; Wang et al., 2015), and voice quality (Motoki, Pathak, & Spence, 2022) are associated with specific tastes. Extending this line of research, our results demonstrate that certain sound aesthetic characteristics are reliably associated to tastes, and that, in turn, congruency can influence taste imagery.

Declaration of Competing Interest

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

Data availability

Data is available on OSF

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Carbon, C. C. (2018). Empirical aesthetics: In quest of a clear terminology and valid methodology. In Z. Kapoula, E. Volle, J. Renoult, M. Andreatta, *Exploring transdisciplinarity in art and sciences*, pp. 107-119.

Chuquichambi, E. G., Vartanian, O., Skov, M., Corradi, G. B., Nadal, M., Silvia, P. J., & Munar, E. (2022). How universal is preference for visual curvature? A systematic review and meta-analysis. *Annals of the New York Academy of Sciences, 1518*(1), 151–165.

Crisinel, A. S., & Spence, C. (2009). Implicit association between basic tastes and pitch. *Neuroscience Letters, 464*(1), 39–42.

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

Crisinel, A. S., & Spence, C. (2010b). A sweet sound? Food names reveal implicit associations between taste and pitch. *Perception, 39*(3), 417–425.

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.

Clemente, A., Vila-Vidal, M., Pearce, M. T., Aguiló, G., Corradi, G., & Nadal, M. (2020). A Set of 200 musical stimuli varying in balance, contour, symmetry, and complexity: Behavioral and computational assessments. *Behavior Research Methods, 52*, 1491–1509.

Clemente, A., Pearce, M. T., Skov, M., & Nadal, M. (2021). Evaluative judgment across domains: Liking balance, contour, symmetry and complexity in melodies and visual designs. *Brain and Cognition, 151*, Article 105729.

Deroy, O., Crisinel, A. S., & Spence, C. (2013). Crossmodal correspondences between odors and contingent features: Odors, musical notes, and geometrical shapes. *Psychonomic Bulletin & Review, 20*, 878–896.

Deroy, O., & Spence, C. (2016). Crossmodal correspondences: four challenges. *Multisensory Research, 29*(1–3), 29–48.

Diessner, R., Genthôs, R., Arthur, K., Adkins, B., & Pohling, R. (2021). Olfactory and gustatory beauty: Aesthetic emotions and trait appreciation of beauty. *Psychology of Aesthetics, Creativity, and the Arts, 15*(1), 38–50.

Elder, R. S., & Krishna, A. (2022). A review of sensory imagery for consumer psychology. *Journal of Consumer Psychology, 32*(2), 293–315.

Fields, L., Verhave, T., & Fath, S. (1984). Stimulus equivalence and transitive associations: A methodological analysis. *Journal of Experimental Behavior, 42*, 143–157.

Guedes, D., Garrido, M. V., Lamy, E., Cavalheiro, B. P., & Prada, M. (2023). Crossmodal interactions between audition and taste: A systematic review and narrative synthesis. *Food Quality and Preference, 107*, Article 104856.

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*, 1–20.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York: Routledge.

Juravle, G., Olari, E. L., & Spence, C. (2022). A taste for beauty: On the expected taste, hardness, texture, and temperature of geometric shapes. *i-Perception, 13*(5): 20416695221120948.

Knöferle, K., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic bulletin & Review, 19*, 992–1006.

Knoeferle, K. M., Knoeferle, P., Velasco, C., & Spence, C. (2016). Multisensory brand search: How the meaning of sounds guides consumers' visual attention. *Journal of Experimental Psychology: Applied, 22*(2), 196–210.

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

Kuhlmann, T., Dantlgraber, M., & Reips, U. D. (2017). Investigating measurement equivalence of visual analogue scales and Likert-type scales in Internet-based personality questionnaires. *Behavior Research Methods, 49*, 2173–2181.

Laurienti, P. J., Kraft, R. A., Maldjian, J. A., Burdette, J. H., & Wallace, M. T. (2004). Semantic congruence is a critical factor in multisensory behavioral performance. *Experimental Brain Research, 158*, 405–414.

Lee, B. P., & Spence, C. (2022). Crossmodal correspondences between basic tastes and visual design features: A narrative historical review. *i-Perception*, *13*(5): 20416695221127325.

Letts, E., Basharat, A., & Barnett-Cowan, M. (2022). Evaluating the Effect of Semantic Congruency and Valence on Multisensory Integration. *Multisensory Research, 35*(4), 309–334.

Luke, S. G. (2016). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*, 1–9.

Marks, L. E. (1978). *The unity of the senses.* Academic Press.

Marks, L. E. (1996). On perceptual metaphors. *Metaphor and Symbol, 11*(1), 39–66.

Motoki, K., Saito, T., Park, J., Velasco, C., Spence, C., & Sugiura, M. (2020). Tasting names: Systematic investigations of taste-speech sounds associations. *Food Quality and Preference, 80*, Article 103801.

Motoki, K., Pathak, A., & Spence, C. (2022). Tasting prosody: Crossmodal correspondences between voice quality and basic tastes. *Food Quality and Preference, 100*, Article 104621.

Motoki, K., Saito, T., & Velasco, C. (2022). Spontaneous crossmodal correspondences grounded in contexts. *Food Quality and Preference, 100*, Article 104619.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology, 64*, 77–107.

Parise, C. V. (2016). Crossmodal correspondences: Standing issues and experimental guidelines. *Multisensory Research, 29*(1–3), 7–28.

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

Qi, Y., Huang, F., Li, Z., & Wan, X. (2020). Crossmodal correspondences in the sounds of Chinese instruments. *Perception, 49*(1), 81–97.

R Core Team. (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.Rproject.org.

Reinoso Carvalho, F., Wang, Q., De Causmaecker, B., Steenhaut, K., Van Ee, R., & Spence, C. (2016). Tune that beer! Listening for the pitch of beer. *Beverages, 2*(4), 31.

Reinoso-Carvalho, F., Gunn, L., Molina, G., Narumi, T., Spence, C., Suzuki, Y., … Wagemans, J. (2020). A sprinkle of emotions vs a pinch of crossmodality: Towards globally meaningful sonic seasoning strategies for enhanced multisensory tasting experiences. *Journal of Business Research, 117*, 389–399.

Salgado-Montejo, A., Alvarado, J. A., Velasco, C., Salgado, C. J., Hasse, K., & Spence, C. (2015). The sweetest thing: The influence of angularity, symmetry, and the number of elements on shape-valence and shape-taste matches. *Frontiers in Psychology, 6*, 1382.

Schifferstein, H. N. (2009). Comparing mental imagery across the sensory modalities. Imagination. *Cognition and Personality, 28*(4), 371–388.

Silvia, P. J. (2007). An introduction to multilevel modeling for research on the psychology of art and creativity. *Empirical Studies of The Arts, 25*(1), 1–20.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis. An introduction to basic and advanced multilevel modeling* (2nd ed.). London: SAGE Publications.

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

Spence, C. (2020). Assessing the role of emotional mediation in explaining crossmodal correspondences involving musical stimuli. *Multisensory Research, 33*(1), 1–29.

Spence, C. (2022). The form of taste: On the origins, implications, and applications of shape-taste crossmodal correspondences. *UOU Scientific Journal, 4*, 150–163.

Spence, C., & Levitan, C. A. (2021). Explaining crossmodal correspondences between colours and tastes. *i-Perception, 12*(3).

Turoman, N., Velasco, C., Chen, Y.-C., Huang, P.-C., & Spence, C. (2018). Tasting transformations: Symmetry and its role in the crossmodal correspondence between shape and taste. *Attention, Perception, & Psychophysics, 80*, 738–751.

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

Velasco, C., Woods, A. T., Deroy, O., & Spence, C. (2015). Hedonic mediation of the crossmodal correspondence between taste and shape. *Food Quality and Preference, 41*, 151–158.

Velasco, C., Woods, A. T., Petit, O., Cheok, A. D., & Spence, C. (2016). Crossmodal correspondences between taste and shape, and their implications for product packaging: A review. *Food Quality and Preference, 52*, 17–26.

Walker, P. (2016). Cross-sensory correspondences: A theoretical framework and their relevance to music. *Psychomusicology: Music, Mind, and Brain, 26*(2), 103–116.

Wang, Q. J., & Spence, C. (2016). 'Striking a sour note': Assessing the influence of consonant and dissonant music on taste perception. *Multisensory Research, 29*(1–3), 195–208.

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.

Wang, Q., 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.