Title: A sprinkle of emotions vs a pinch of crossmodality: Towards globally meaningful sonic seasoning strategies for enhanced multisensory tasting experiences

Abstract

We report a study designed to determine the most efficient means of pursuing sonic seasoning in international marketing. For the first time, music chosen to trigger specific emotional responses was directly and cross-culturally compared with music chosen as crossmodally congruent with specific taste/flavors (the latter usually regarded as ‘sonic seasoning’).

The effects triggered by ‘emotional’ music were more prominent than those triggered by ‘crossmodally-corresponding’ music. Specifically, chocolate was liked more, rated as sweeter, and the purchase intent was higher, when tasted while listening to music that conveyed positive, as compared to negative, emotion. By contrast, the same chocolate was mostly rated as tasting more bitter with the negative music, as compared to the positive music.

Companies looking to use sonic seasoning in marketing strategies, should therefore principally aim at intelligently classifying music based on the likely emotions that they can trigger in their customers (at least when thinking more globally).

Keywords: *crossmodal; emotions; flavors; multisensory; music; purchase intention.*

Introduction

Background music undoubtedly influences consumer behavior in a number of ways. Over the last decade or so, our knowledge concerning how music exerts such an influence on the experience of foods and drinks has become increasingly sophisticated (see Spence et al., 2019, for a review). At the same time, researchers and practitioners have been seeking out new ways in which to better disentangle how the different musical attributes, and the consequent sensations and emotions triggered by such music, can be used to generate specific reactions in listeners prior, during, and after tasting.

Literature review

Thinking of music as a set of combined auditory attributes, one attribute that is usually easy to identify by the listener is the style/genre of music which, for instance, has been shown to bias purchase behavior in the world of food and drink (e.g., North et al., 1997, 1999; Zellner et al., 2017). In this context, North et al. (1999) played French and German music on alternate days, over a period of 2 weeks, in the wine section of a British supermarket. French music led to French wines outselling German ones, whereas German music led to the opposite effect on consumer choice.

The style of music can also affect what people think of a tasting experience (e.g., Yeoh & North, 2010), where the perceived/rated pleasantness of the flavor and the overall impression of food may change depending on the genre of music playing in the background (e.g., Fiegel et al., 2014; Martens et al., 2010). However, if one digs a little deeper, music goes beyond merely its style, being an exhaustive production process, one that involves a combination of particular sonic elements that are combined while composing – think of musical tempo, relative loudness, frequency-ranges/pitch, among many other attributes.

In fact, musical tempo has been shown to affect the shopping and tasting experience of consumers (e.g., Bach & Schaefer, 1979; Caldwell & Hibbert, 2002). For example, the tendency to drink more rapidly when high (rather than low) tempo music is played has been known for decades (e.g., Roballey et al., 1985; cf. Milliman, 1986). When it comes to loudness, numerous studies have shown that consumers tend to drink more when the music is louder (e.g., Guéguen et al., 2004; McCarron & Tierney, 1989). It would also seem that the perceived flavor of alcoholic beverages can be modulated (i.e., enhanced or suppressed) under the influence of loud background music (cf. Stafford et al., 2012; Woods et al., 2010). Concerning the dimension of pitch, several studies have demonstrated that different frequency ranges tend to be associated with specific taste/flavor attributes, such as sweetness, bitterness, and sourness

(e.g., Holt-Hansen, 1968; Knöferle & Spence, 2012; Reinoso Carvalho et al., 2016; Rudmin & Capelli, 1983). In this sense, several musical attributes have been tested for their crossmodal correspondence (the associations that people make between features, attributes or dimensions across the senses; see Spence, 2011, for a review). Here, food and drink can provide a range of useful tools to compose, for example, ‘sweet’, ‘bitter’, ‘salty’, ‘sour’, ‘spicy’, ‘smooth’, or ‘rough’ music (e.g., Crisinel & Spence, 2010; Guetta & Loui, 2017; Mesz et al., 2011). Based on the growing range of crossmodal correspondences that have been documented, music is now being composed, or characterized, specifically to trigger a kind of ‘seasoning’ for different foods/drinks, such as toffee, wine, chocolate, ice-cream, and beer (e.g., Crisinel et al., 2012; Reinoso Carvalho et al., 2015; Wang & Spence, 2018; and see Spence et al., 2019, for an overview). That is, not only can music be composed to correspond with specific taste and flavor attributes, but it is possible to modulate such taste/flavor attributes when eating/drinking while listening to congruent as compared to incongruent music (this is what has been characterized as ‘sonic seasoning’; Spence, 2017; Velasco et al., 2016). Nevertheless, the understanding of the applicability of sonic seasoning may be compromised if we do not consider the fact that our primary relationship with music usually occurs at an emotional level (Huron, 2007; Juslin & Sloboda, 2001). In other words, would the emotions that we experience from, or associate with, the music that happens to be playing in the background while we eat/drink be transferred to the tasting experience itself (see Cheskin, 1972, on the notion of sensation transference effects)? So far, the answer to this question would seem to be in the affirmative, where these emotions also seem to act as an important mediator during sonic seasoning demonstrations (e.g., Kantono et al., 2016, 2019; Reinoso-Carvalho et al., 2019; Ziv, 2018).

Theoretical framework

Intelligent music classification techniques focused on multisensory tasting experiences present a wide range of opportunities, as far as multisensory consumer-computer interaction is concerned. Based on sonic seasoning literature, marketers along with researchers and developers, can reconsider the ways in which they deliver music toward new comprehensive multisensory tasting experiences in order to satisfy digital costumer needs and desires in very different ways. However, before moving forward confidently while investing in, e.g., developing and fine-tuning algorithms to classify music by means of sonic seasoning, one might wonder if the existing literature on this topic is sufficiently robust to approach, say, a multi-cultural consumer base, in a dynamic international market (Knoerfele et al., 2015).

Hence, a priority for research in this area would seem to be to cross-culturally extend the discussion prior to provide to users and consumers with truly engaging, and ultimately comprehensive, multisensory tasting experiences in the near future (Spence, 2019a).

In light of the above considerations, in the present study, it was decided to assess the effects that music can trigger in the tasting experience, where a clear research gap is being addressed for the first time. If companies look to effectively use sonic seasoning techniques as part of their global marketing strategies, how should the music be chosen? Should this selection process be based mainly on crossmodal correspondences? Or perhaps the process should prioritize the emotional effects that the music will likely trigger in listeners?

Hence, the first research goal guiding the present experimental research was to directly compare the effects of crossmodally corresponding music (sonic seasoning) vs emotional music, in order to determine which approach would be more efficient, when tested under similar multisensory tasting conditions.

H1. There will be differences in the multisensory tasting experience of consumers when music is played that has been chosen to trigger particular emotional effects versus music that has been shown to crossmodally correspond with specific taste/flavor attributes. Such differences will be detectable on distinctive taste and flavor ratings (e.g., the bitterness, sweetness, texture), as well as on how much consumers end-up liking the tasting experience, their willingness to pay, and/or purchase intention for the food at stake.

On the one hand, the available research shows that music can be composed, or characterized, in order to correspond to certain specific sensations (e.g., Imschloss & Kuehnl, 2019), including gustatory ones, such as bitterness, sweetness, sourness (e.g., Crisinel et al., 2012; Reinoso Carvalho et al., 2015; Watson & Gunter, 2017; and see Knöferle & Spence, 2012, for an overview). For instance, according to Watson and Gunter (2017), trombones are more strongly associated with bitterness than clarinets. Note that this study was conducted as a partial replication of Crisinel et al.'s (2012) earlier assessment, where the taste of cinder toffee was found to be modulated by 'sweet' vs 'bitter' music. Reinoso Carvalho et al. (2015) also revealed that what people heard (one 'sweet', one 'bitter', and one in-between music clips) enhanced the corresponding taste ratings of chocolate samples.

On the other hand, alternative evidence suggests that emotions seem to mediate similar effects to those reported above (e.g., Kantono et al., 2016, 2019; Reinoso-Carvalho et al., 2019; Wang & Spence, 2016; Ziv, 2018). For instance, according to Kantono et al. (2016), the perceived

sweetness and bitterness of ice-cream tends to be modulated by whether the music that happens to be playing in the background is liked or not. Meanwhile, Reinoso-Carvalho et al. (2019) demonstrated that discriminative and affective ratings related to the flavor experience when tasting beer are affected by whether people listen to music that happens to induce positive vs negative emotions. Ziv (2018) also reported that biscuits tend to be rated as tasting better when consumers listen to pleasant (as compared to unpleasant) music. Considering that olfactory cues are an important component of flavor perception (Stevenson et al., 2000), it is relevant to note that auditory cues have also been shown to modulate odor pleasantness (Seo & Hummel, 2009; Velasco et al., 2014). For example, Seo and Hummel (2009) reported that the hedonic valence that is associated with auditory stimuli can be transferred to olfactory stimuli, and that such transference effects do not appear to depend on how pleasant people rated the smells in question.

In summary, from the context of an actual business situation, even though there is robust rationale proving that music can prompt specific effects in listeners while experiencing foods/drinks, it is still not necessarily clear how to most effectively pursue such multisensory strategies in, e.g., the context of marketing. For instance, such a lack of knowledge can have important repercussions on convincing the different stakeholders to adopt such techniques during the appropriate stages of business planning.

As a second research goal, it was considered that consumers located in different continents tend to have different cultural backgrounds and, thus, tend to behave somewhat differently (Landes, 2000). For example, Eitan and Timmers (2010) assessed how Western populations were able to perform crossmodal matchings using what they framed as Western (familiar) and non-Western (unfamiliar) metaphors for “high” and “low” auditory pitch. The results highlighted a strong agreement in terms of how such a population used the assessed familiar and unfamiliar metaphors for auditory pitch, across different tasks. Walker (1987) also focused on understanding how participants with different cultural background, age, and musical training, made associations between four acoustic parameters (frequency, waveform, amplitude, duration), and a selection of visual metaphors for each of these parameters. In the latter study, significant cultural and environmental effects were observed in those participants with less exposure to Western lifestyles, on the choices of visual metaphors for sound. In other words, those brands interested in adopting ‘sonic seasoning’ as part of their global multisensory marketing strategies need to consider how these effects operate across diverse cultural backgrounds (Smith et al., 2013; Spence et al., 2019).

With the above being said, in the present study, the aim was also to expand the scope of H1 towards a more international perspective, by conducting a comparison between two very different cultures (in this case, Asian vs Latin-American). Since this study experimentally applies music in different countries, it is relevant to consider that people tend to understand music differently as a function of culture (e.g., Clayton, 2013; Cross, 2011). For instance, some key musical features are used differently by distinct cultures, and such differences may end-up giving rise to different musical mindsets (e.g., Walker, 1987; see also Huron, 2007). Think, for example, about how most Western music is based on diatonic collections of notes, which are basically heptatonic scales (7 tones scale, sub-divided in 12 semi-tones), whereas pentatonic scales (scales composed by 5 tones, rather than 7) are more common musical resources in eastern/Asian cultures, such as traditional Japanese music (Blair & Kondo, 2008).

H2. The effects analyzed in H1 may give rise to different outcomes when tested between Latin-American versus Asian consumers.

H2 is relatively novel in this literature, where most of the existing data on the topic is still based on a western population (e.g., Western Europe, North America). However, there are a couple of previous reports assessing different cultural contexts that support the formulation of H2. For instance, previous research shows that higher-pitched sounds (e.g., when constituting part of voiceover ads) can increase the purchase intent towards sweet food in Japanese consumers, when compared to their purchase intention for bitter food (Motoki et al., 2019). Knoeferle et al. (2015) also demonstrated that sweet, sour, bitter, and salty versions of the same soundscape were similarly matched to the corresponding taste when analyzed between pools of Indian and North-American participants.

In recent and similar reports, variables such as WTP – which are commonly assessed in, i.e., marketing studies – have already provided some interesting, but not always conclusive, insights (e.g., Reinoso-Carvalho et al., 2019). Therefore, as a third research goal, in this new study, it was also decided to quantify the purchase intent, by comparing it with the participant's willingness to pay (WTP). This decision was only made after sampling in Latin-America, and prior to sampling in Asia and, thus, it was only performed in the Asian population.

H3. Purchase intention ratings will show more prominent effects when compared to the WTP of consumers

A number of researchers have worried that asking the WTP as an open question (i.e., asking to rate the WTP in currency, as typically done previously), may be less informative and accurate than researchers may have hoped (O'Donnell & Evers, 2018). For example, the level of accuracy of these ratings may depend on how well the participant understood the referential point and/or context (i.e., "Should I associate this WTP to a specific amount of product? Is this WTP related to the product being offered in a supermarket, in a restaurant, in a bar?" And so on). Purist behavioral economists have also argued for the necessity of engaging a more realistic context while assessing the WTP of a consumer by, e.g., forcing the participants to pay for something during an experimental procedure (e.g., Wertenbroch & Skiera, 2002). Purchase intent, on the contrary, does not necessarily translate into sales (Armstrong et al., 2000; Morwitz et al., 2007).

The present study

In this new assessment, a cross-cultural comparison of the effects of crossmodally corresponding music (sonic seasoning) vs emotional music, on tasting experiences was conducted for the first time. This cross-cultural comparison was carried out across Latin-American (Colombian) and Asian (Japanese) participants. Two methods previously and independently reported only with Western-European consumers, were used as baseline to conduct this original cross-cultural comparison.

First, two songs shown to trigger contrasting emotional effects on listeners (either positive or negative) in Reinoso-Carvalho et al. (2019) were used. The results of the latter showed that while tasting beer under the influence of these two songs, a population of mainly Western-European consumers liked the beer more, and rated it as tasting sweeter, while listening to the song associated with positive emotion. The same beer in this study was rated as tasting more bitter, with a higher alcohol content, and as having more body, when participants listened to a song that was associated with negative emotion instead. Here, the participants were also willing to pay 7–8% more for the beer tasted while they listened to the positive music (see Reinoso-Carvalho et al., 2019, for more details).

Second, two contrasting songs that were previously produced with crossmodal correspondences between music and flavor in mind in Reinoso Carvalho et al. (2017) were also used. One of these songs was validated as corresponding with softness, and the other to hardness (think of the softer song as being congruent with a soft/smooth/creamy food texture while the hard song was more congruent with a hard/rough food texture instead). In this study, while tasting the same chocolate twice, and each time listening to one of these two songs, the

results, based on a population of mainly Western-European consumers, revealed that the ‘soft’ song enhanced the perceived creaminess and sweetness of the chocolate, as compared to the ratings given to the same chocolate while listening to the ‘hard’ song (see Reinoso Carvalho et al., 2017, for more details).

Besides the most salient sensory aspects that a naïve participant may be able to disentangle from a typical tasting experience (i.e., bitterness, sweetness, differences in texture), in this original study, other aspects that influence market-relevant behaviors/judgements were also assessed (Lemon et al., 2016; Spence & Piqueras-Fiszman, 2014). In particular, hedonic ratings, such as the liking of the food, the liking of the music that the participants were being exposed to, how much they thought that the flavor of the food would match the music that they were exposed to, as well as their WTP, and purchase intention for the food, were also assessed. Note that, except for flavor strength and purchase intention, the majority of these variables have been shown to be affected by sonic seasoning (i.e., crossmodal correspondences between sound and taste/flavor) in previous research (flavor/sound liking, flavor-sound matching, bitterness, sweetness, sourness, willingness to pay; e.g., see Reinoso Carvalho et al., 2019).

Materials and methods

Experimental design

In this study, each participant tasted, in total, the same chocolate twice, and no more than two experimental pieces of it. Each time they tasted the chocolate, they were listening to one of the songs from the two available music pairs, depending on which experimental condition they were randomly assigned to (milk-based, or dark-based chocolate, along with emotional or crossmodally-corresponding music pairs). The decision of presenting a contrasting pair of songs in a within-participants design was based on the proven relative nature of crossmodal correspondences, where contrast is often crucial for more effectiveness (Spence, 2019b).

The distribution of chocolate types and songs was randomized across participants. Prior to and after tasting the chocolates, the participants were instructed to clean their palates with tap water. They were not informed that they were actually tasting the same chocolate twice. After tasting, they were asked to provide basic demographic data (age, gender) and answer questions related to their chocolate-music experiences (how much they liked the flavor of the chocolate, how much they thought the music matched the flavor of the chocolate, the sweetness/bitterness/flavor-strength/texture of the chocolate, how much they liked the music that they were listening to while tasting the chocolate, their WTP, and purchase intention, for the chocolate). **Table 1** specifies more details on the dependent variables that were assessed

via the electronic questionnaire that was implemented on the Qualtrics online platform. Note that the presentation of the songs was coded with combinations of alphabetical letters, in order to avoid order bias effects (Stone et al., 2012). Moreover, the order of presentation of choices of answers was also fully randomized.

INSERT TABLE 1 ABOUT HERE

Participants

A total of 1390 participants from two different cultural contexts (Asian and Latin-American – Latam – nationals and/or residents, represented by Japanese and Colombian, respectively), took part in this study. They were selected by means of convenience sampling. **Table 2** shows more details of the sample, and how it was sub-divided across the experimental conditions. A sample size simulation conducted via Raosoft sample size calculator (see **Note 1**) suggested that the number of participants gathered for each experimental condition ($n \sim 377$) would be suitable for a behavioral study aiming towards a 95% confidence interval (Hamburg, 1974).

INSERT TABLE 2 ABOUT HERE

In Colombia, the study was conducted inside rooms that were physically similar at the campus of Universidad de los Andes, located in Bogotá (<https://uniandes.edu.co/en>). In Japan, the study took place at one of the sites where expositions are held at the Miraikan National Museum of emerging science and innovation, located in Tokyo (<https://www.miraikan.jst.go.jp/en/>; see also **Note 2**). In both countries, the obtained data was gender balanced, and filtered in order to keep only local participants aged between 15 and 80 years.

Stimuli

Two types of chocolates were used in both countries. One was milk-based (Callebaut n. 823, milk-based with 33.6% minimum cacao solids), and the other dark (Callebaut n. 811, without-milk, with 54.5% minimum cacao solids). Each experimental unit of chocolate had a diameter of around 1 centimeter and weighted approximately 1g, regardless of type.

As aforementioned, two pairs of songs that had been used in previously-published research were used. The first pair of songs (namely, the emotional music) had been used previously in Reinoso-Carvalho et al.'s (2019) study, where they were shown to mainly trigger contrasting emotional effects in listeners (either positive or negative). In this study, the different emotions that this musical selection evoked in participants were validated by means of the PANAS scale (see Crawford & Henry, 2004, and Watson et al., 1988, for details). Based on the score proposed by PANAS, the positive music was significantly evoking positive emotions, and the negative was also significantly evoking negative emotions. Note here also that such a selection of music considered the avoidance of different complex semantic associations that may be primed by the particular style of music, as both songs resemble general classical music (cf. Hutchison, 2003; Labroo et al., 2008; Lucas, 2000). These songs can be heard at <https://tinyurl.com/music-emotions-xcultural>. The song labeled as Clip 1 represented positive emotion, while Clip 4 corresponded to negative emotion (see Reinoso-Carvalho et al., 2019). The second pair (namely crossmodally-corresponding music) was extracted from Reinoso Carvalho et al.'s (2017) study, where two contrasting songs were produced mainly with crossmodal correspondences between music and taste/flavor in mind (Spence, 2011). The bouba-kiki effect reported by Köhler (1947), and the relationship between touch and sound highlighted by Eitan and Rothschild (2010), acted as a baseline for the production of these sounds. These crossmodally-corresponding songs were prepared and validated by means of an online pre-test questionnaire, one corresponding to softness/smoothness (namely soft) and the other to hardness/roughness (namely hard). In this pre-test, and based on previous findings, the authors considered that sweetness and bitterness are usually opposites in terms of valence as compared to, for example, sourness and bitterness (Reinoso Carvalho et al., 2016; Salgado-Montejo et al., 2015; Yarmolinsky et al., 2009). Hence, in this questionnaire, participants rated on two bipolar 7-point scales the sweetness vs bitterness, and the creaminess/softness vs roughness/hardness that they would attribute to each song (with one scale ranging from '1' being very bitter, '4' balanced, and '7' very sweet; and the second scale ranging from '1' being very rough/hard, '4' being balanced, and '7' being very creamy/soft). These results revealed that the 'soft' song was rated as significantly creamier, and sweeter, by the participants, when

compared to the ratings for the ‘hard’ song. Both songs can be heard at <http://tinyurl.com/creaminess-chocolate> (see Reinoso-Carvalho et al., 2017, for more details on this study).

Procedure

At all times, all of the participants were exposed to similar controlled environmental conditions, including comfortable room temperature, low background noise, and stable lighting. In each country, all of the participants used the same computer system and headphones. The playback volume of the music was calibrated in order to be similar (approximately 70 dB +/- 6dB). Those people that were passing near the experimental area during experimental hours, were invited to take part in the study. As an incentive to join, they were told that they would be participating in a short experiment involving tasting chocolate while listening to different sounds. There were 8 to 10 experimental booths operating at all times, where participants were able to participate individually. The experimental setting was organized in order for the participants not to be able to listen to each other, while also being reinforced to perform the experiment individually. During the study, each of the participants ate each of the chocolate samples from a different plate, and were instructed to let the chocolate melt in their mouths (i.e., they were told not to immediately bite and chew it). Prior to the tasting process, they were advised to focus on the most salient flavor attributes of the chocolate while listening to the music.

Data analysis

A multivariate ordinal approach was used to account for non-equally spaced thresholds for the latent (continuous) measurements, which additionally allows for flexibility of non-symmetric ordinal outcomes across the ordinal scale (see Note 3). Hence, a multivariate multinomial ordered probit model (Hirk et al., 2019a) was implemented to measure the differences between ordinal scales, with a dummy intercept vector to account for the within-participant measurement changes in chocolate perception across measured ordinal items, while also accounting for covariate differences (age, gender, order of presentation of stimuli). This jointly-measured – namely intercept – vector represents the distances (in latent measurement space) between the two measurements for each participant across measurement items, while accounting for within-participant correlations between measurements for each item and across measurement items. Positive values represent a higher (latent) valuation of the second measurement.

Let Y_{ijt} be the ordinal response provided by participant i regarding item j (with C_j possible ordinal categories, which may differ across items) at time t , with $t=0$ representing the baseline measurement and $t=1$ representing the second measurement. Similarly, let Z_{ijt} be the latent continuous variable connected to the measurement variable Y_{ijt} through the following representation:

$$Y_{ijt} = c_{ijt} \Leftrightarrow \zeta_{j,c_{ijt-1}} < Z_{ijt} \leq \zeta_{j,c_{ijt}}$$

where $\zeta_{j,1} = 0$ for parameter identifiability, as it is standard practice in the ordinal statistics literature. Then, for a set of covariates X_i (age, gender, order of presentation), which are common across items and time for a given participant i , the model for the latent variable takes the following form:

$$Z_{ijt} = \alpha_j + \beta X_i + \gamma_j t + \varepsilon_{ijt}$$

where ε_{ijt} is assumed to be distributed as a multivariate Gaussian. The main quantity of interest is the vector γ , which accounts for the sound differences between the two measurements, and whether the value is significantly different from zero. Bonferroni-adjusted results are reported for both conditions, representing an overly-conservative approach, since measurements across items are highly correlated. The Akaike Information Criterion (AIC) is used for model discrimination. Note that covariate-related parameter information is also provided. This statistical analysis was performed in R, using the mvord package (Hirk et al., 2019b).

In order to measure the relative willingness to pay, the ratio of quantitative measurements was log-transformed to represent a percent difference between tasting (with the baseline tasting in the denominator), and regressed against the same set of covariates defined above assuming a Gaussian relationship. The latter statistical analysis was performed using the core syntax of R.

Results

General comparison: The effects of emotional vs crossmodally-corresponding music

In general, the results were more stable for the effect of emotional music, where ratings for most of the variables were similar across both types of chocolates and cultures. The ratings that replicated similarly across both conditions related to how much the chocolate flavor, and the music, were enjoyed, and how well the music matched the flavor of the chocolate. Concerning the effects of the participants' age on the results, there was a general pattern across both conditions suggesting that the older the participant, the higher the rating they gave. Regarding the effects of gender, and the order of presentation of the songs, the results revealed no consistent significant outcomes (see **Table 3** for more details on the aforementioned results).

Table 4 includes a summary of the declared WTP (in percentage difference terms) as a function of age, gender, and order. In most cases, there is insufficient information to claim statistical significance for most WTP associations.

Tasting chocolate while listening to emotional music

In general, the flavor of the chocolate was liked more, and the purchase intention was higher (the latter only measured amongst the Asian participants), when tasted alongside the positive emotional song. The chocolate was also consistently rated as sweeter when tasted alongside the positive emotional song, and more bitter when presented together with the negative emotional song. The latter was not significant only for the Asian participants while tasting the dark chocolate. The texture of the dark chocolate was also rated as softer while listening to the positive emotional song, but only by the Latam participants. The positive emotional song was also liked more, and was rated as a better match for the flavor of the chocolate, when compared to the negative emotional song (see **Figure 1**).

INSERT FIGURE 1 ABOUT HERE

Tasting chocolate while listening to crossmodally-corresponding music

The chocolate flavor was liked more when accompanied by the ‘soft’ song (except for Latam participants while tasting the dark chocolate). Furthermore, the ‘soft’ song was liked more, and it was considered a better match to the flavor of the chocolate, when compared to the ‘hard’ song. None of the differences concerning the sensory evaluation of the chocolates was significant with this music. Finally, the purchase intention for the chocolate was higher for Asian participants while listening to the ‘soft’ song, when compared to their purchase intention while listening to the ‘hard’ song, but only for the milk chocolate (see **Figure 2**).

INSERT FIGURE 2 ABOUT HERE

The corresponding dataset is publicly available, and can be accessed via <http://dx.doi.org/10.17632/rrm7sbfpcy.1>

Discussion

The main purpose of this large-scale study was to innovate in terms of pursuing universal sonic seasoning effects on consumers' tasting experiences. For the first time, music chosen to principally trigger an emotional effect was compared to music chosen primarily to be crossmodally congruent with specific flavor attributes, during a multisensory tasting experience. Here, each participant tasted the same chocolate twice, and each time while listening to a pair of contrasting songs (positive vs negative emotions, or crossmodally-corresponding with soft vs hard sensations, depending on the experimental condition to which they were randomly assigned to). This is also the first time that these effects have been cross-culturally assessed far away from the usual Western-European cultural scope, and while using two different types of chocolates (milk and dark).

In general, the results were clearly more prominent under the effect of the emotional music, where ratings for most of the variables assessed were fully replicated (see H1). Specifically, the chocolate flavors were liked more, and the purchase intention was higher (though the latter was only assessed in Asia), when tasted along with the positive emotional music. The chocolates were also rated as sweeter when tasted along with the positive emotional music (see **Figure 1**). These results suggest that there was positive emotion induced by this positive music. Additionally, the positive music was preferred over the negative musical selection. As expected, the aforementioned effects give support to a positive sensation transference, from the positive emotion music, to the tasting experience (Cheskin, 1972; Skaczkowski et al., 2016).

In contrast, the chocolates were mostly rated as tasting more bitter when accompanied by the music associated with negative emotion. The latter was not, however, the case for the dark chocolate in Asia (see **Figure 1**). As expected, the negative music drove negative emotions. Here, it could be argued that the negative emotions brought out by the music resulted in most consumers trying to explain the unpleasantness of the experience (e.g., Schwarz & Clore, 1983) by associating such unpleasantness with taste/flavor aspects that may not be usually regarded as synonyms of quality in chocolate, such as an excess of bitterness (as in a kind of attentional shift/redirection effect; e.g., Johnson & Proctor, 2004).

Concerning the soft vs hard pair of crossmodally-corresponding musical stimuli, the chocolate was generally liked more when accompanied by the soft music. The latter was also preferred, and was thought to be a better match to the flavor of the chocolate (as compared to the hard music; see **Figure 2**). As with the emotional music, these results suggest, once again, that the preference for the soft music had a hedonic effect on the chocolate tasting experience. However, when comparing to the results obtained with emotional music, in the crossmodally-corresponding musical condition none of the effects triggered by the soft vs hard songs significantly affected the sensory evaluation of the chocolates. Here, the crossmodally-corresponding songs did not seem to deliver the expected ‘seasoning’ effects in Latam and Asian consumers, when compared to what Western-Europeans reported a couple of years ago in Reinoso Carvalho et al.’s study (2017); see H2.

While music appreciation tends to differ across cultures (Blair & Kondo, 2008), differences in dietary manners also seem to be relevant when thinking on the crossmodal correspondences assessed in this study. For example, besides general differences in flavor pairing between East and West, Westerners tend to share ingredients across food recipes more often than Eastern consumers do (Ahn et al., 2011; Mouritsen et al., 2017). A study conducted in different regions of China has also suggested that a history of farming rice makes cultures more interdependent, whereas wheat farming cultures tend to be more independent (Talhekm et al., 2014). Here, a parallel can be drawn between Japanese (more rice based) vs Western-European and/or Latam (more wheat based) cultures. Concerning the particularities of Japanese culinary culture, and their potential differences with Latin-American cuisine, Japanese traditional diet has been framed as more balanced in terms of nutrition (thus, providing a wider variety of taste and flavor sensations; Ikeda, 2002; Nakamura, 2011). In other words, if different cultures ‘understand’ music differently, and have their palates trained in different ways, it is plausible to assume that the related associations between music and taste/flavor will likely exhibit different correspondences. Interestingly, Talhekm et al.’s (2014) aforementioned study also suggests that a more collective farming culture (e.g., rice farming) can lead to more global manner of thinking. Like this, it is possible to imagine the musical culture of each country as a potential mediator for more/less global thinking, and this may have interesting consequences when strategically applied into multisensory tasting experiences.

The results related to purchase intent were more prominent when compared to the null results obtained for the WTP ratings, which may inspire future similar studies to continue assessing purchase intention over WTP (see H3). Note that, in the case of the crossmodally-corresponding music, the purchase intention of the chocolate was higher while listening to the

soft song, but only for the milk chocolate. The latter results may be more associated to the type of chocolate than to the music itself, since dark chocolate may be usually regarded as a more sophisticated – and perhaps more exclusive – type of chocolate.

In general, the importance of cultural variance in marketing has been shown to affect the perception of consumers in different ways, where most of these assessments have focused on understanding this matter visually, i.e., by means of product labelling (e.g., Caporale & Monteleone; 2001; Herlesth et al., 2011; Verbeke & Roosen, 2009). When thinking about how managers could be further inspired by this new study, an opportunity to explore customized crossmodal correspondences between what we listen and what we taste as a way to, for example, innovate while approaching the way culture nudges consumption, could emerge.

Implications for multisensory consumer-computer interaction

To date, multisensory consumer experiences involving music are gaining more traction in the context of promotion (i.e., via digital branding and/or communication; see Campari [see **Note 4**], Godiva [see **Note 5**]). In 2019, for instance, Jägermeister designed and launched a multisensory consumer experience with music being produced to enhance specific flavor notes of this liquor (see **Note 6**). Based on music streaming technology, Tramontina – a Brazilian company dedicated to cookware and cutlery – recently supported the development of an algorithm that could turn any song on Spotify into a recipe (see **Note 7**).

Compared to previous similar studies, and from the context of an actual business situation, it is now clearer that the most effective way to pursue this type of multisensory effect passes principally by means of the emotions triggered by the music that consumers listen while experiencing food and drink. Hence, this study provides novel and strategic knowledge while, for instance, negotiating with the different stakeholders in terms of adopting such techniques during the proper stages of business planning. Based on the results obtained here, those companies who may be keen to adopt (or, perhaps, continue using more appropriately) sonic seasoning techniques as part of their marketing strategies, should aim towards intelligently classifying music based on the emotions that they can trigger in consumers for greater efficacy. That is, whereas crossmodal correspondences are usually about mapping potential congruent associations that people tend to make between features across the senses (in this case, between music and tasting features), music chosen to trigger emotions do not necessary rely on congruency.

From a theoretical viewpoint, this study adds original evidence of the key mediating role of emotions for cross-cultural sonic seasoning (cf. Kantono et al., 2016, 2019; Reinoso-Carvalho

et al., 2019; Ziv, 2018). For example, the results obtained here point towards emotions as a potential distinctive class of crossmodal correspondence. It also seems appropriate to integrate both factors (emotion along with crossmodality), where music can be produced/classified to be crossmodally congruent with specific taste/flavor notes, and, at the same time, let's say, pleasant to the global listener whoever they might be. The latter could be achieved by, e.g., considering his/her own personal musical preferences.

It is also intriguing to foresee how automated/intelligent sonic seasoning strategies may end-up influencing consumer choice/acceptance in the near future. For instance, one could foresee the studies reported by North and colleagues, back in the 90's, soon being replicated through, i.e., virtual/augmented retail environments (see North et al., 1997, 1999).

Multisensory design, together with technology, may eventually also be used to nudge consumers into healthier behaviors. For example, it is plausible to imagine an intelligent selection of background music, or soundscapes, gently inviting consumers to choose more fruits/vegetables or sustainably sourced protein, while looking for their next meal through online food ordering/delivering platforms (cf. Youssef et al., 2019; Zellner et al., 2017). Moreover, as seen in this exercise, positive emotion background music may be an interesting path to gently persuade consumers from different parts of the world to add less sugar to their food or drink (cf. Blecken, 2017; Lowe et al., 2018), being excessive sugar consumption a major health problem (Lustig et al., 2012).

Due to its reduced sugar and milk content, dark chocolate can also be considered as strategic while branding a healthier lifestyle. Interestingly, in this study, the chocolate's bitterness was enhanced by the mediation of the least preferred and negatively-valenced type of music in Latin America. Hence, depending on the desired outcomes, negativeness and/or multisensory incongruency may be as useful as the opposite — and perhaps more common — strategy (at least in some cultures).

Finally, in this study, we are introducing the multivariate multinomial ordered probit model as an alternative/more-robust way to analyze this type of ordinal data, when compared to standard approaches, such as ANOVA. The multivariate multinomial ordered probit model provides a very powerful statistical tool that is already widely used in the areas of machine learning and statistics (e.g., Hirk et al 2019a, b). Such a model is especially relevant for conducting multidimensional classification tasks for ordered data. Since this study is mostly aimed at finding the best way to intelligently classify music (with multisensory tasting experiences in mind), exploring the use of this statistical model seems pertinent to those topics related to multisensory computer-consumer interaction.

Limitations and future work

This study was designed primarily to determine the most efficient means of pursuing sonic seasoning strategies, while cross-culturally comparing music that was chosen to trigger a specific emotional response, vs music that had been chosen to be crossmodally congruent with specific taste/flavor attributes. However, it is also plausible to assume that emotions were being triggered by the crossmodally-congruent music. Hence, future similar assessments could further inquire into this by, for instance, including other techniques for measuring the affection (emotional/mood state) of participants during this type of multisensory tasting experiences, along with the behavioral methods that were here applied. Techniques that help better locating the sensations reported by the participants over time could also be of support (e.g., Temporal check-all-that-apply - TCATA, Temporal dominance of sensations - TDS). It may be also important to verify if these effects are easily replicable in a between-participants experimental design. So far, most of these studies seem to be designed in a within-participants context (see Table 1 of Spence et al., 2019, where it was summarized that 15 out of 19 of these studies were, in fact, designed as within-participants). The latter is most likely due to the relative nature of crossmodal correspondences, where contrast is often decisive for more effectiveness (Spence, 2019b). Finally, it is also necessary to clarify that, even though an important part of this work is being positioned in the cross-cultural domain, H3 was fully tested only with Asian consumers.

Ethics statement

This study was carried out in accordance with the recommendations of the Social and Societal Ethics Committee at KUL (SMEC, protocol registered as G-2015 07 281). All of the participants gave their written informed consent prior to starting the study.

Notes

1. <http://www.raosoft.com/samplesize.html>; retrieved February, 2020.
2. <https://www.miraikan.jst.go.jp/event/1907031024390.html>; retrieved February, 2020.
3. ANOVA methods are oftentimes used for analysis of ordinal scales under strong assumptions, such as symmetry of the underlying distribution and equal-spacing of ordinal scales to match a continuous equivalent in the real line (e.g., the unobservable distance measure of the variables to be equal to the distance measure represented by the chosen numerical representation of those ordered categories). While this approach is valid in restricted cases, it

should not be understood as a blanket method for the analysis of ordinal data. Since there is no theoretical basis to support equally-spaced interpretation by respondents of the ordinal scales within the questionnaires, a conservative approach is needed that does not rely on the aforementioned assumptions and remains valid regardless of whether the ordinal scales are integer mappings of a continuous underlying scale or not. Additionally, in this study, the data collected strong non-symmetric behavior among scales for some items and regions, with large numbers of respondents selecting extreme values within the rating scales.

4. <https://www.softespace.com/en/portfolio/campari/>; retrieved August, 2019.
5. <https://www.youtube.com/watch?v=rph6oyIEJ9o>; retrieved August, 2019.
6. <https://www.jagermeister.com/en/taste-remastered>, retrieved October, 2019.
7. <https://www.youtube.com/watch?v=dykfJI2fLTw>; retrieved October, 2019.

References

- Ahn, Y. Y., Ahnert, S. E., Bagrow, J. P., & Barabási, A. L. (2011). Flavor network and the principles of food pairing. *Scientific Reports*, **1**:196.
- Armstrong, J. S., Morwitz, V. G., & Kumar, V. (2000). Sales forecasts for existing consumer products and services: Do purchase intentions contribute to accuracy? *International Journal of Forecasting*, **16**(3), 383-397.
- Bach, P. J., & Schaefer, J. M. (1979). The tempo of country music and the rate of drinking in bars. *Journal of Studies on Alcohol*, **40**, 1058-1059.
- Biswas, D., Lund, K., & Szocs, C. (2019). Sounds like a healthy retail atmospheric strategy: Effects of ambient music and background noise on food sales. *Journal of the Academy of Marketing Science*, **47**(1), 37-55.
- Blair, D. V., & Kondo, S. (2008). Bridging musical understanding through multicultural musics. *Music Educators Journal*, **94**(5), 50-55.
- Blecken, D. (2017). Hold the sugar: A Chinese café brand is offering audio sweeteners. *Campaign*, **February 13th**. <https://www.campaignasia.com/video/hold-the-sugar-a-chinese-cafe-brand-is-offering-audio-sweeteners/433757>.
- Caldwell, C., & Hibbert, S. A. (2002). The influence of music tempo and musical preference on restaurant patrons' behavior. *Psychology & Marketing*, **19**, 895-917.
- Camargo, J., González, M., Guzmán, A., ter Horst, E., & Trujillo, M. A. (2018). Topics and methods in economics, finance, and business journals: A content analysis enquiry. *Heliyon*, **4**(12):e01062.

- Caporale, G., & Montelone, E. (2001). Effect of expectations induced by information on origin and its guarantee on the acceptability of a traditional food: Olive oil. *Sciences des Aliments*, **21**(3), 243-254.
- Casarin, R., Camargo, J., Correa, J. C., ter Horst, E., Dakduk, S., & Molina, G. (2019). What makes a tweet be retweeted? A Bayesian trigram analysis to understanding tweet propagation during the 2015 Colombian political campaign. *Journal of Information Science*.
<https://doi.org/10.1177/0165551519886056>
- Cheskin, L. (1972). *Marketing success: How to achieve it*. Boston, MA: Cahnners Books.
- Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, **43**, 245-265.
- Clayton, M. (2013). *The cultural study of music: A critical introduction*. New York, NY: Routledge.
- Cross, I. (2001). Music, cognition, culture, and evolution. *Annals of the New York Academy of Sciences*, **930**(1), 28-42.
- Crisinel, A.-S., Cosser, S., King, S., Jones, R., Petrie, J., & Spence, C. (2012). A bittersweet symphony: Systematically modulating the taste of food by changing the sonic properties of the soundtrack playing in the background. *Food Quality and Preference*, **24**, 201-204.
- Crisinel, A.-S., & Spence, C. (2010). As bitter as a trombone: Synesthetic correspondences in non-synesthetes between tastes and flavors and musical instruments and notes. *Attention, Perception, & Psychophysics*, **72**, 1994-2002.
- Eitan, Z., & Rothschild, I. (2010). How music touches: Musical parameters and listeners' audiotactile metaphorical mappings. *Psychology of Music*, **39**(4), 449-467.
- Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, **114**(3), 405-422.
- Fiegel, A., Meullenet, J. F., Harrington, R. J., Humble, R., & Seo, H.-S. (2014). Background music genre can modulate flavor pleasantness and overall impression of food stimuli. *Appetite*, **76**, 144-152.
- Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1, No. 10). New York, NY: Springer.
- Guéguen, N., Le Guellec, H., & Jacob, C., (2004). Sound level of background music and consumer behavior: An empirical evaluation. *Perceptual and Motor Skills*, **99**, 34-38.

- Guetta, R., & Loui, P. (2017). When music is salty: The crossmodal associations between sound and taste. *PLoS ONE*, **12**(3):e0173366.
- Hamburg, M. (1974). *Basic statistics: A modern approach*. Boston, MA: Houghton Mifflin Harcourt.
- Hersleth, M., Lengard, V., Verbeke, W., Guerrero, L., & Næs, T. (2011). Consumers' acceptance of innovations in dry-cured ham: Impact of reduced salt content, prolonged aging time and new origin. *Food Quality and Preference*, **22**(1), 31-41.
- Hirk, R., Hornik, K., & Vana, L. (2019a). Multivariate ordinal regression models: An analysis of corporate credit ratings. *Statistical Methods & Applications*, **28**(3), 507-539.
- Hirk, R., Hornik, K., Vana, L., & Genz, A. (2019b). Multivariate ordinal regression models [computer software]. GPL-3 licensing. Available from <https://cran.r-project.org/web/packages/mvord/mvord.pdf>
- Holt-Hansen, K. (1968). Taste and pitch. *Perceptual and Motor Skills*, **27**, 59-68.
- Huron, D. (2007). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: MIT Press.
- Hutchison, A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, **10**, 785-813.
- Ikeda, K. (2002). New seasonings. *Chemical Senses*, **27**(9), 847-849.
- Imschloss, M., & Kuehnl, C. (2019) Feel the music! Exploring the crossmodal correspondence between music and haptic perception of softness. *Journal of Retailing*, **95**(4), 158-169.
- Johnson, A., & Proctor, R. W. (2004). *Attention: Theory and practice*. Thousand Oaks, CA: Sage.
- Juslin, P. N., & Sloboda, J. A. (Eds.). (2001). *Music and emotion: Theory and research*. New York, NY: Oxford University Press.
- Kantono, K., Hamid, N., Shepherd, D., Lin, Y. H. T., Skiredj, S., & Carr, B. T. (2019). Emotional and electrophysiological measures correlate to flavor perception in the presence of music. *Physiology & Behavior*, **199**, 154-164.
- Kantono, K., Hamid, N., Shepherd, D., Yoo, M. J. Y., Carr, B. T., & Grazioli, G. (2016). The effect of background music on food pleasantness ratings. *Psychology of Music*, **13**, 1-15.
- Knoeferle, K. M., Woods, A., Kppler, F., & Spence, C. (2015). That sounds sweet: Using crossmodal correspondences to communicate gustatory attributes. *Psychology & Marketing*, **32**, 107-120.
- Knferle, K. M., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic Bulletin & Review*, **19**, 992-1006.

- Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York, NY: Liveright.
- Labroo, A. A., Dhar, R., & Schwartz, N. (2008). Of frog wines and frowning watches: Semantic priming, perceptual fluency, and brand evaluation. *Journal of Consumer Research*, **34**, 819-831.
- Landes, D. (2000). Culture makes almost all the difference. In L. E. Harrison (2002). *Culture matters: How values shape human progress* (Chapter 1, pp. 2-13). New York, NY: Basic books.
- Lemon, K. N., & Verhoef, P. C. (2016). Understanding customer experience throughout the customer journey. *Journal of Marketing*, **80**(6), 69-96.
- Lowe, M., Ringler, C., & Haws, K. (2018). An overture to overeating: The cross-modal effects of acoustic pitch on food preferences and serving behaviour. *Appetite*, **123**, 128-134.
- Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, **7**, 618-630.
- Lustig, R. H., Schmidt, L. A., & Brindis, C. D. (2012). The toxic truth about sugar. *Nature*, **482**(7383), 27-29.
- Martens, M., Skaret, J., & Lea, P. (2010). Sensory perception of food products affected by different music genres In: EuroSense, Vitoria-Gasteiz, Spain, 5-8th September. P1.124.
- McCarron, A., & Tierney, K. J. (1989). The effect of auditory stimulation on the consumption of soft drinks. *Appetite*, **13**, 155-159.
- Mesz, B., Trevisan, M., & Sigman, M. (2011). The taste of music. *Perception*, **40**, 209-219.
- Milliman, R. E. (1986). The influence of background music on the behavior of restaurant patrons. *Journal of Consumer Research*, **13**, 286-289.
- Morwitz, V. G., Steckel, J. H., & Gupta, A. (2007). When do purchase intentions predict sales? *International Journal of Forecasting*, **23**(3), 347-364.
- Motoki, K., Saito, T., Nouchi, R., Kawashima, R., & Sugiura, M. (2019). A sweet voice: The influence of cross-modal correspondences between taste and vocal pitch on advertising effectiveness. *Multisensory Research*, **32**(4-5), 401-427.
- Mouritsen, O. G., Edwards-Stuart, R., Ahn, Y. Y., & Ahnert, S. E. (2017). Data-driven methods for the study of food perception, preparation, consumption, and culture. *Frontiers in ICT*, **4**:15.
- Nakamura, E. (2011). One hundred years since the discovery of the “umami” taste from seaweed broth by Kikunae Ikeda, who transcended his time. *Chemistry—An Asian Journal*, **6**(7), 1659-1663.

- North, A. C., Hargreaves, D. J., & McKendrick, J. (1997). In-store music affects product choice. *Nature*, **390**, 132.
- North, A. C., Hargreaves, D. J., & McKendrick, J. (1999). The influence of in-store music on wine selections. *Journal of Applied Psychology*, **84**, 271-276.
- O'Donnell, M., & Evers, E. R. (2018). Preference reversals in willingness to pay and choice. *Journal of Consumer Research*, **45**(6), 1315-1330.
- Reinoso Carvalho, F., Van Ee, R., Rychtarikova, M., Touhafi, A., Steenhaut, K., Persoone, D., Spence, C., & Leman, P. (2015). Does music influence the multisensory tasting experience? *Journal of Sensory Science*, **30**, 404-412.
- Reinoso Carvalho, F., Wang, Q. (J.), De Causmaecker, B., Steenhaut, K., Van Ee, R., & Spence, C. (2016). Tune that beer! Listening to the pitch of beer. *Beverages*, **2**:31.
- Reinoso Carvalho, F., Wang, Q. (J.), Van Ee, R., Persoone, D., & Spence, C. (2017). "Smooth operator": Music modulates the perceived creaminess, sweetness, and bitterness of chocolate. *Appetite*, **108**, 383-390.
- Reinoso-Carvalho, F., Dakduk, S., Wagemans, J., & Spence, C. (2019). Not just another pint! The role of emotion induced by music on the consumer's tasting experience. *Multisensory Research*, **32**(4-5), 367-400.
- Roballey, T. C., McGreevy, C., Rongo, R. R., Schwantes, M. L., Steger, P. J., Wininger, M. A., & Gardner, E. B. (1985). The effect of music on eating behavior. *Bulletin of the Psychonomic Society*, **23**, 221-222.
- Rudmin, F., & Cappelli, M. (1983). Tone-taste synesthesia: A replication. *Perceptual & Motor Skills*, **56**, 118.
- 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.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality Social Psychology*, **45**, 513-523.
- Seo, H.-S., & Hummel, T. (2009). Effects of olfactory dysfunction on sensory evaluation and preparation of foods. *Appetite*, **53**(3), 314-321.
- Skaczkowski, G., Durkin, S., Kashima, Y., & Wakefield, M. (2016). The effect of packaging, branding and labeling on the experience of unhealthy food and drink: A review. *Appetite*, **99**, 219-234.

- Smith, P. B., Fischer, R., Vignoles, V. L., & Bond, M. H. (2013). *Understanding social psychology across cultures: Engaging with others in a changing world*. Sussex, UK: Sage.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review, *Attention, Perception, & Psychophysics*, **73**, 971-995.
- Spence, C. (2017). Sonic seasoning. In L. Minsky & C. Fahey (Eds.), *Audio branding: Using sound to build your brand* (pp. 52-58). London, UK: Kogan Page.
- Spence, C. (2019a). Multisensory experiential wine marketing. *Food Quality & Preference*, **71**, 106-116.
- Spence, C. (2019b). On the relative nature of (pitch-based) crossmodal correspondences. *Multisensory Research*, **32**, 235-265.
- Spence, C., & Piqueras-Fiszman, B. (2014). *The perfect meal: the multisensory science of food and dining*. Oxford, UK: John Wiley & Sons.
- Spence, C., Reinoso-Carvalho, F., Velasco, C., & Wang, Q. J. (2019). Extrinsic auditory contributions to food perception & consumer behaviour: An interdisciplinary review. *Multisensory Research*, **32**(4-5), 275-318.
- Stafford, L. D., Fernandes, M., & Agobiani, E. (2012). Effects of noise and distraction on alcohol perception. *Food Quality & Preference*, **24**, 218-224.
- Stevenson, R. J., Boakes, R. A., & Wilson, J. P. (2000). Counter-conditioning following human odor–taste and color–taste learning. *Learning and Motivation*, **31**(2), 114-127.
- Stone, H., Bleibaum, R. N., & Thomas, H. A. (2012). *Sensory evaluation practices* (4th Ed.). London, UK: Academic Press.
- Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., & Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science*, **344**(6184), 603-608.
- Velasco, C., Balboa, D., Marmolejo-Ramos, F., & Spence, C. (2014). Crossmodal effect of music and odor pleasantness on olfactory quality perception. *Frontiers in Psychology*, **5**:1352.
- Velasco, C., Reinoso Carvalho, F., Petit, O., & Nijholt, A. (2016). A multisensory approach for the design of food and drink enhancing sonic systems. In *Proceedings of the 1st Workshop on Multi-sensorial Approaches to Human-Food Interaction* (p. 7). Tokyo, Japan: ACM.
- Verbeke, W., & Roosen, J. (2009). Market differentiation potential of country-of-origin, quality and traceability labeling. *The Estey Centre Journal of International Law and Trade Policy*, **10**(1), 20-35.
- Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics*, **42**, 491-502.

- Wang, Q. (J.) & Spence, C. (2016). "Striking a sour note": Assessing the influence of consonant and dissonant music on taste perception. *Multisensory Research*, **30**, 195-208.
- Wang, Q., & Spence, C. (2018). Assessing the influence of music on wine perception among wine professionals. *Food Science & Nutrition*, **6**(2), 295-301.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales *Journal of Personality Social Psychology*, **54**(6), 1063-1070.
- Watson, Q. J., & Gunter, K. L. (2017). Trombones elicit bitter more strongly than do clarinets: A partial replication of three studies of Crisinel and Spence. *Multisensory Research*, **30**, 321-335.
- Wertenbroch, K., & Skiera, B. (2002). Measuring consumers' willingness to pay at the point of purchase. *Journal of Marketing Research*, **39**(2), 228-241.
- Woods, A. T., Poliakoff, E., Lloyd, D. M., Dijksterhuis, G. B., & Thomas, A. (2010). Flavor expectation: The effects of assuming homogeneity on drink perception. *Chemosensory Perception*, **3**, 174-181.
- Yarmolinsky, D. A., Zuker, C. S., & Ryba, N. J. (2009). Common sense about taste: From mammals to insects. *Cell*, **139**(2), 234-244.
- Yeoh, J. P. S., & North, A. C. (2010). The effects of musical fit on choice between two competing foods. *Musicae Scientiae*, **14**, 127-138.
- Youssef, J., Keller, S., & Spence, C. (2019). Making sustainable foods (such as jellyfish) delicious. *International Journal of Gastronomy & Food Science*.
<https://doi.org/10.1016/j.ijgfs.2019.100141>.
- Zellner, D., Geller, T., Lyons, S., Pyper, A., & Riaz, K. (2017). Ethnic congruence of music and food affects food selection but not liking. *Food Quality & Preference*, **56**(Part A), 126-129.
- Ziv, N. (2018). Musical flavor: The effect of background music and presentation order on taste. *European Journal of Marketing*, **52**, 1485-1504.

Tables

Table 1. Experimental design, including the variables sampled, and the system used to measure each variable. The presentation of the variables was fully randomized.

Section of questionnaire	Variable	Measurement system
1. Demographics	Age	Open numerical
	Gender	1=Male; 2=Female
2. Rating scales	Like flavor of chocolate	7-point rating scale (with 1 being ‘not at all’ and 7 being ‘very much’).
	Match flavor of chocolate and music	
	Sweetness of chocolate	
	Bitterness of chocolate	
	Like music	
	Flavor strength	
	Purchase intention ¹	7-point rating scale (with 7 being ‘very high’ and 1 being ‘very low’).
	Texture of chocolate	7-point rating scale (with 1 being ‘very hard’ and 7 being ‘very soft’).
3. Open question	Willingness to pay (WTP)	Open numerical in local currency, including message with local reference.

1- This question was only presented to the Asian participants.

Table 2. Sample sizes for each experimental condition by dates, location, and music (n_{total}=1390), as well as average age and corresponding standard deviation, SD, both in years).

Dates of collection	Region of collection	n _{Emotional music}	n _{Crossmodally-corresponding music}	n _{Total} Mean age (SD)
March-April, 2019	Latin America (Latam)	340	317	20 (2) ¹
June-July, 2019	Asia	383	350	37 (12) ²

¹ Median Age in Colombia: 30 yrs (source - World Fact book: <https://www.cia.gov>; retrieved August, 2019).

² Median Age in Japan: 47 yrs (source - World Fact book: <https://www.cia.gov>; retrieved August, 2019).

Table 3. Summary of outcomes for the rating scales concerning the chocolate experience under the influence of the emotional and crossmodally-corresponding music. When it comes to the intercept (*SE*), ratings are reported in white cells when significant at the 95% confidence level (Bonferroni adjusted), whereas the grey cells are those ratings where $p \geq 0.05$. A positive value of the intercept means that the dominant effect comes from the positive/soft song, and vice versa. When it comes to those effects related to the covariates (age, gender, and presentation order of the songs, each one with corresponding *SE*), values in bold indicate a significant value at the 95% confidence level.

Running head: Sonic seasoning in multisensory marketing

	MUSIC TYPE	POPULATION	CHOCOLATE TYPE	VARIABLE / RATING	INTERCEPT	SE	AGE	SE	GENDER	SE	ORDER	SE
EMOTIONS	ASIA	MILK		LIKE FLAVOR	1.08	0.10	0.03	0.00	0.12	0.13	-0.05	0.13
				MATCH FLAVOR-MUSIC	1.62	0.12	0.01	0.00	0.02	0.11	0.04	0.11
				SWEETNESS	0.40	0.09	0.02	0.00	0.53	0.13	0.16	0.14
				BITTERNESS	-0.49	0.11	0.00	0.00	-0.39	0.14	0.46	0.14
				LIKE MUSIC	1.86	0.13	0.01	0.00	0.37	0.12	-0.01	0.12
				FLAVOR STRENGTH	0.23	0.09	0.02	0.00	0.31	0.14	0.34	0.13
				TEXTURE	0.26	0.09	0.04	0.00	0.31	0.14	-0.05	0.13
				PURCHASE INTENTION	0.77	0.11	0.02	0.00	0.31	0.11	0.15	0.11
		DARK		LIKE FLAVOR	0.87	0.10	0.02	0.00	0.34	0.14	0.15	0.13
				MATCH FLAVOR-MUSIC	1.10	0.11	0.01	0.00	0.21	0.11	0.06	0.12
				SWEETNESS	0.47	0.09	0.02	0.00	0.51	0.14	0.08	0.13
				BITTERNESS	-0.26	0.09	0.01	0.00	0.18	0.13	0.48	0.13
				LIKE MUSIC	1.61	0.12	0.01	0.00	0.16	0.12	-0.05	0.12
				FLAVOR STRENGTH	0.14	0.10	0.02	0.00	0.30	0.13	0.41	0.13
				TEXTURE	0.26	0.08	0.02	0.00	0.22	0.14	0.37	0.15
				PURCHASE INTENTION	0.65	0.11	0.01	0.00	0.38	0.13	0.18	0.12
	LATAM	MILK		LIKE FLAVOR	0.65	0.10	0.08	0.01	0.24	0.17	-0.27	0.16
				MATCH FLAVOR-MUSIC	1.48	0.15	0.04	0.01	-0.04	0.15	-0.43	0.14
				SWEETNESS	0.41	0.12	0.09	0.01	0.15	0.15	-0.31	0.14
				BITTERNESS	-0.56	0.12	0.02	0.01	-0.25	0.17	0.14	0.16
				LIKE MUSIC	1.16	0.11	0.04	0.01	-0.15	0.16	0.00	0.16
				FLAVOR STRENGTH	0.14	0.11	0.09	0.01	0.11	0.16	-0.24	0.14
				TEXTURE	0.37	0.11	0.08	0.01	0.17	0.16	-0.08	0.16
		DARK		LIKE FLAVOR	0.75	0.10	0.07	0.01	0.26	0.12	0.02	0.12
				MATCH FLAVOR-MUSIC	0.65	0.11	0.05	0.01	0.28	0.10	-0.14	0.11
				SWEETNESS	0.69	0.10	0.05	0.01	-0.06	0.12	0.29	0.12
				BITTERNESS	-0.57	0.09	0.06	0.01	0.27	0.12	0.29	0.11
				LIKE MUSIC	1.47	0.11	0.04	0.01	-0.04	0.11	-0.10	0.12
				FLAVOR STRENGTH	0.06	0.10	0.06	0.01	0.33	0.12	0.35	0.12
CROSSMODAL	ASIA	MILK		TEXTURE	0.45	0.08	0.07	0.01	0.15	0.12	0.07	0.12
				LIKE FLAVOR	0.69	0.09	0.03	0.00	0.08	0.12	0.25	0.13
				MATCH FLAVOR-MUSIC	1.18	0.10	0.02	0.00	0.23	0.12	0.35	0.13
				SWEETNESS	0.14	0.07	0.03	0.00	0.49	0.14	0.28	0.14
				BITTERNESS	-0.09	0.12	0.00	0.00	-0.35	0.14	-0.25	0.14
				LIKE MUSIC	1.37	0.10	0.02	0.00	0.45	0.13	0.25	0.13
		DARK		FLAVOR STRENGTH	0.12	0.07	0.02	0.00	0.27	0.14	0.42	0.14
				TEXTURE	0.11	0.08	0.03	0.00	0.32	0.13	0.06	0.14
				PURCHASE INTENTION	0.64	0.11	0.03	0.00	0.43	0.11	0.38	0.12
				LIKE FLAVOR	0.53	0.10	0.02	0.00	0.59	0.15	0.28	0.16
				MATCH FLAVOR-MUSIC	1.04	0.10	0.01	0.00	0.37	0.15	0.51	0.15
				SWEETNESS	-0.10	0.10	0.02	0.00	0.47	0.15	0.33	0.16
	LATAM	MILK		BITTERNESS	0.03	0.11	0.01	0.00	0.31	0.16	0.07	0.16
				LIKE MUSIC	1.05	0.12	0.01	0.00	0.42	0.14	0.36	0.16
				FLAVOR STRENGTH	0.22	0.10	0.02	0.00	0.51	0.16	0.25	0.17
				TEXTURE	-0.09	0.10	0.03	0.00	0.04	0.17	0.56	0.17
				PURCHASE INTENTION	0.15	0.11	0.02	0.00	0.47	0.13	0.17	0.14
		DARK		LIKE FLAVOR	0.38	0.10	0.09	0.01	0.16	0.18	0.19	0.16
				MATCH FLAVOR-MUSIC	0.96	0.12	0.06	0.00	0.09	0.15	0.41	0.13
				SWEETNESS	0.04	0.08	0.07	0.01	0.21	0.19	0.24	0.16
				BITTERNESS	-0.14	0.13	0.00	0.01	-0.31	0.18	0.02	0.17
				LIKE MUSIC	1.25	0.12	0.06	0.01	0.02	0.15	0.22	0.14
				FLAVOR STRENGTH	0.32	0.11	0.07	0.01	0.20	0.18	0.40	0.13
		DARK		TEXTURE	0.24	0.09	0.09	0.01	0.18	0.17	0.11	0.17
				LIKE FLAVOR	0.26	0.10	0.08	0.01	0.10	0.13	0.27	0.12
				MATCH FLAVOR-MUSIC	0.67	0.11	0.05	0.01	0.25	0.13	0.36	0.12
				SWEETNESS	0.12	0.11	0.06	0.01	-0.01	0.13	0.38	0.12
				BITTERNESS	-0.14	0.10	0.05	0.01	0.13	0.13	0.02	0.12
				LIKE MUSIC	1.37	0.11	0.06	0.01	0.30	0.13	0.40	0.11
				FLAVOR STRENGTH	0.16	0.10	0.07	0.01	0.31	0.13	0.19	0.13
				TEXTURE	0.19	0.08	0.07	0.01	0.17	0.14	0.42	0.13

Table 4. Summary of outcomes for the declared WTP (as percentage in relative terms) under the influence of emotional and crossmodally-corresponding music. These results include the parameter estimates for the multiple linear regression of the relative WTP (percentage change for a given individual with respect to his/her original WTP) as a function of the set of three covariates (age, gender, and order of presentation of sounds). Hence, the intercept represents the estimated relative WTP when all three proposed covariates equal zero. The remainder of the parameter estimates represent the estimated impact on the relative WTP per unit increase in the covariate, while holding all the other covariates constant.

<i>MUSIC</i>	<i>POPULATION</i>	<i>CHOCOLATE</i>	<i>VARIABLE</i>	<i>ESTIMATE</i>	<i>SE</i>	<i>P-VALUE</i>
EMOTIONS	ASIA	MILK	Intercept	-0.16	0.15	0.28
			Age (years)	0.00	0.00	0.78
			Gender	-0.07	0.09	0.44
			Order	-0.02	0.09	0.81
		DARK	Intercept	-0.09	0.12	0.45
			Age (years)	0.00	0.00	0.93
			Gender	0.03	0.09	0.75
			Order	-0.08	0.09	0.36
	LATAM	MILK	Intercept	-0.44	0.44	0.32
			Age (years)	0.01	0.02	0.58
			Gender	0.14	0.10	0.16
			Order	-0.03	0.10	0.72
		DARK	Intercept	-0.65	0.27	0.02
			Age (years)	0.02	0.01	0.06
			Gender	-0.02	0.08	0.79
			Order	0.00	0.08	0.99
CROSSMODAL	ASIA	MILK	Intercept	0.19	0.15	0.21
			Age (years)	0.00	0.00	0.97
			Gender	0.20	0.09	0.02
			Order	-0.05	0.09	0.56
		DARK	Intercept	0.10	0.14	0.47

		Age (years)	0.00	0.00	0.63
		Gender	-0.02	0.10	0.83
		Order	0.02	0.10	0.84
LATAM	MILK	Intercept	-0.12	0.17	0.47
		Age (years)	0.01	0.01	0.45
		Gender	0.22	0.07	0.00
	DARK	Order	0.10	0.07	0.15
		Intercept	0.03	0.23	0.90
		Age (years)	0.01	0.01	0.46
		Gender	-0.11	0.09	0.24
		Order	-0.05	0.09	0.62

Figures

Figure 1. Ratings concerning the chocolate experience with the emotional music, where only the ratings that showed a significant difference, at 95% confidence and Bonferroni-corrected, are represented. There are four groups of data represented, depending on the combination between chocolate type (dark or milk-based) and the region where the participants were located (Asia or Latam). The y-axis shows the intercept (average mean difference between positive vs negative song ratings), and corresponding standard error (SE) bars. A positive value means that the dominant effect comes from the positive song, and vice versa. The x-axis shows each type of rating.

Figure 2. Ratings concerning the chocolate experience with the crossmodally-corresponding music, where only those ratings that revealed a significant difference, at 95% confidence and Bonferroni-corrected, are represented. Four groups of data are shown, depending on the combination between chocolate type (dark or milk-based formulae) and the continent in which the participants were located (Asia or Latam). The y-axis shows the intercept (average mean difference between soft vs hard songs ratings) and corresponding standard error (SE) bars. A positive value means that the dominant effect comes from the soft song. The x-axis shows each type of rating.