

i-Comment**On why music changes what (we think) we taste****Charles Spence**Crossmodal Research Laboratory, Department of Experimental Psychology, University of Oxford, Oxford, UK;
e-mail: charles.spence@psy.ox.ac.uk**Ophelia Deroy**Centre for the Study of the Senses, School of Advanced Study, University of London, London, UK;
e-mail: Ophelia.Deroy@sas.ac.uk

Received 7 December 2012, in revised form 7 March 2013; published online 16 April 2013.

Abstract. A pair of recently published studies demonstrate that what we happen to be listening to can sometimes change our perception (or, at the very least, our rating) of what we are eating or drinking. In one recent study, North (2012) showed that the emotional attributes (or connotation) of a piece of music could influence people's perception of red or white wine. Meanwhile, Crisinel et al. (2012) reported that listening to a lower-pitched soundscape can help to emphasize the bitter notes in a bittersweet toffee while listening to a soundscape with a higher pitch tends to bring out its sweetness. Although the most appropriate psychological and neuroscientific explanations for such crossmodal effects are still uncertain, we outline a number of possible alternatives for such intriguing, not to mention surprising, phenomena.

Keywords: music, flavour, taste, crossmodal correspondences, crossmodal effects.

Humans, not to mention other species, make many, what at first may seem surprising, associations between experiences presented in different sensory modalities, such as matching darker objects with lower-pitched sounds and lighter objects with higher-pitched sounds (e.g., Ludwig, Adachi, & Matzuzawa, 2011; Marks, 1978; Spence, 2004). One consequence of the existence of such crossmodal correspondences is that changing what people experience in one modality can sometimes change their experience of the stimuli presented in another.

North (2012) recently published an intriguing study demonstrating just such a surprising crossmodal influence in which the type of music playing in the background was shown to influence students' rating of the qualities of a sample of wine that they were evaluating. The 250 undergraduate students who took part in this study were offered a glass of either Chilean red (a cabernet sauvignon) or white wine (a chardonnay) in return for answering a series of questions about its taste. The students drank the wine while seated in a room in which one of four pre-selected pieces of music was playing at 70 dB (there was also a no-music control condition). The music samples had been chosen on the basis of a pilot study in which they had scored highly on one of several emotional dimensions: "powerful and heavy" (*Carmina Burana* by Orff), "subtle and refined" (*Waltz of the Flowers* from Tchaikovsky's "Nutcracker"), "zingy and refreshing" (*Just can't get Enough* by Nouvelle Vague), and "mellow and soft" (*Slow Breakdown* by Michael Brook).

Having emptied their glass, the students had to rate the wine on four scales (by giving a score from 0 to 10) anchored with the labels "powerful and heavy," "subtle and refined," "zingy and refreshing," and "mellow and soft." A value of 0 indicated that the wine definitely did not have the characteristic, while a score of 10 indicated that the wine definitely did. The participants also had to rate how much they liked the wine. The results demonstrated that the wines were rated as significantly more powerful and heavy when Carl Orff was played in the background than when any of the other pieces of music were playing. Meanwhile, the wines were rated as significantly more "zingy and refreshing" by those listening to the Nouvelle Vague track. A similar effect was reported for the other two musical pieces. Interestingly, however, there was no significant effect on liking, thus suggesting that the music influenced the descriptive, rather than the evaluative, aspects of the tasting experience.

While it is certainly true that, over the years, a number of writers have reached for musical metaphors when trying to describe wines (Spence, [2011b](#); North's [2012](#)) latest findings provide some of the first empirical evidence to demonstrate that what you hear, or listen to, really can affect the sensory evaluation of the wine that one happens to be drinking. But why should that be, given that the music was playing in the background and was not in any obvious way related to the wine itself? And what, exactly, is affected? Before we address these difficult questions, it is worth pointing out that North's findings do not stand alone in terms of demonstrating a significant influence of audition on gustation.

Crisinel et al. ([2012](#)) also recently reported that people's perception of a sample of bittersweet toffee was modified simply by varying the pitch of a soundtrack that was playing over headphones. In this case, the soundtracks had been developed on the basis of prior research showing that people typically pick lower-pitched tones as matching (or corresponding with) bitter tastes while reporting that higher-pitched sounds provide a better match for sweet tastes (Crisinel & Spence, [2010a, b](#); see also Knöferle & Spence, [2012](#)).

Now, a quick comparison of these two studies reveals how different the approaches and interpretations of the music-taste/flavour relation are. In Crisinel et al. ([2012](#)), the mapping of specific notes (or soundscapes) onto particular tastes (or, for that matter, flavour, aroma, or oral-somatosensory textural) or qualities was framed in terms of the notion of crossmodal correspondences (Palmer, Schloss, Xu, & Prado-León, [in press](#); Spence, [2011a, 2012](#)). As argued elsewhere (Deroy & Spence, [in press](#); Spence, [2011a](#)), the term "crossmodal correspondence" refers to the general tendency for our brains and/or minds to match features or dimensions across sensory modalities, but is often restricted to those cases where the matching seems arbitrary or surprising. Indeed, the idea that crossmodal correspondences are arbitrary explains why they are often associated (or confused) with synaesthesia (e.g., Ludwig et al., [2011](#); Marks, [1978](#); Martino & Marks, [2001](#); Spence & Deroy, [2012](#)).

By contrast, North ([2012](#)) did not look for surprising perceptual associations between wines and music, but instead selected his musical pieces on the basis that they had attributes that could be applied to both music and wine. In this sense, his starting point was metaphor; that is, it was based on the fact that people use, or find it natural to use, the words "heavy" or "mellow" to refer both to music and wine. Notice also that the words chosen were *not* straightforwardly capturing a musical qua auditory characteristic of the piece, but were themselves examples of metaphorical transfer (Williams, [1976](#)) from the tactile modality (concerning temperature, texture, and weight). One worry here, then, is that North might have confirmed the attraction of certain metaphorical mappings between music, wines, and tactile attributes, more than necessarily having demonstrated a genuine perceptual effect. As North himself puts it, the results mostly show that: "participants *appeared* to perceive the taste of the wine in a manner consistent with the connotations of the music" (North, [2012](#), p. 298, emphasis our own).

Although North ([2012](#)) is correct in stressing that the demonstrated influence of music on participants' ratings of the wine might be attributable to crossmodal priming, it is reasonable to ask whether the effect is really perceptual. The fact that it was not semantic, because music and wine do not have a shared semantic field, led North to suggest that his results might reflect a form of connotative or symbolic priming instead; that is, where the presentation of one stimulus (here music) triggers some interpretation or symbolic associations (not a set of referents), which, in turn, primes a characteristic of another object that happens to be perceived at around the same time. Although this idea is certainly intriguing, one might reasonably ask whether the notion of symbolic priming is theoretically robust and whether other empirical evidence can be brought to bear in support.

Nothing here necessarily implies that the basis for this crossmodal transfer meant that the flavour of the wine was affected, or that it is in virtue of a perceived commonality or correspondence between the wine and the music that North's results were obtained. Certainly, nothing rules out the possibility that the softness of the wines' tannins or their sweetness were associated with the perceived softness of musical tones, and perceived as being more intense as a result. However, in the absence of an independent sensory evaluation of the wines by expert panelists (rather than a bunch of thirsty undergraduates), it is hard to assess whether the change in sensory descriptors by the participants (with the appropriate music) could be related to the boost in a sensory dimension really possessed by the wine.

It is, then, worth underscoring the differences between North ([2012](#)) and Crisinel et al. ([2012](#)), which start with crossmodal correspondences (rather than metaphors). Crisinel et al. based their research on earlier findings demonstrating that the majority of people match bitter tastes with lower-pitched sounds while matching sweet tastes with higher-pitched sounds (Crisinel & Spence, [2010a, b](#)). Now, although the interpretations in term of semantic priming and emotional effects can still apply

(see Crisinel & Spence, 2012), it is harder to apply the former as the metaphor of sweet as high pitch certainly is not common in the English language (and, indeed, contradicts the rule/description of metaphorical transfer originally suggested by Williams, 1976, and thereafter confirmed by Shen & Aisenman, 2008).

As such, Crisinel et al.'s (2012) results suggest another explanation, namely one in terms of the lower-level influence of crossmodal correspondences on perception. Audiovisual correspondences, for instance between pitch, brightness, and elevation (see Marks, 2004; Spence, 2004, for reviews) give rise to rapid and seemingly automatic compatibility effects in a variety of tasks such as speeded detection tasks or variants of the implicit association test. Such associations can be formalized in Bayesian decision theory as coupling priors (see Spence, 2011a).

One expectation that comes with such a hypothesis is that crossmodal correspondences have some form of ecological validity—that is, they have been internalized at some point because they pick up on the statistical regularities of the environment, and therefore facilitate multisensory integration (or prediction). The critical question to ask here then is whether pitch–taste correspondences can be interpreted in this way, or whether instead they represent just another form of metaphorical transfer? According to Spence (2012), the pitch–taste correspondence may have its origins in the stereotypical orofacial gestures that the newborns of many species make in response to the presentation of a tastant. In particular, the tongue normally goes out and up in response to sweet tastes as the newborn tries to ingest the calories signified by the sweetness. By contrast, the tongue is normally pushed out of the mouth and down in response to bitter tastes (presumably an evolutionarily adaptive strategy given that many bitter tasting foods are poisonous). The pitch of any sound (or utterance) made by the neonate would likely be lower in the latter case than in the former.

Although speculative, Spence's (2012) hypothesis generates a number of testable predictions, which could also help to extend North's (2012) wine and music study. For instance, if perceptually grounded, the crossmodal effects of music on taste/flavour perception should be (relatively) culture-independent (although linguistic conventions and cultural differences in musical appreciation might cover the initial perceptual correspondence, this could be tested with pre-verbal infants... though probably best not with wine!); or that changing certain low-level features of the pieces of music utilized by North, such as their pitch, should lead to a specific boost of a certain perceived characteristic present in the wine (for instance, its bitterness). To conclude, as one of the key challenges in this area of research currently lies in trying to determine the appropriate explanation, both psychological and neuroscientific, for such surprising crossmodal effects and influences of audition on gustation and flavour perception, we would argue that researchers will need to thoroughly test those that have been reliably documented to date.

References

- 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. doi:10.1016/j.foodqual.2011.08.009
- Crisinel, A.-S., & Spence, C. (2010a). A sweet sound? Exploring implicit associations between basic tastes and pitch. *Perception*, 39, 417–425. doi:10.1068/p6574
- Crisinel, A.-S., & Spence, C. (2010b). 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. doi:10.3758/APP.72.7.1994
- Crisinel, A.-S., & Spence, C. (2012). The impact of pleasantness ratings on crossmodal associations between food samples and musical notes. *Food Quality and Preference*, 24, 136–140. doi:10.1016/j.foodqual.2011.10.007
- Deroy, O., & Spence, C. (in press). Weakening the case for “weak synaesthesia”: Why crossmodal correspondences are not synaesthetic. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-013-0387-2
- Knöferle, K. M., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic Bulletin & Review*, 19, 992–1006. doi:10.3758/s13423-012-0321-z
- Ludwig, V. U., Adachi, I., & Matzuzawa, T. (2011). Visuoauditory mappings between high luminance and high pitch are shared by chimpanzees (*Pan troglodytes*) and humans. *Proceedings of the National Academy of Sciences USA*, 108, 20661–20665. doi:10.1073/pnas.1112605108
- Marks, L. (1978). *The unity of the senses: Interrelations among the modalities*. New York: Academic Press. doi:10.1121/1.382733
- Marks, L. E. (2004). Cross-modal interactions in speeded classification, In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes* (pp. 85–105). Cambridge, MA: MIT Press.

-
- Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, 10, 61–65. [doi:10.1111/1467-8721.00116](https://doi.org/10.1111/1467-8721.00116)
- North, A. C. (2012). The effect of background music on the taste of wine. *British Journal of Psychology*, 103, 293–301. [doi:10.1111/j.2044-8295.2011.02072.x](https://doi.org/10.1111/j.2044-8295.2011.02072.x)
- Palmer, S. E., Schloss, K. B., Xu, Z. X., & Prado-León, L. (in press). Color, music, and emotion. *Proceedings of the National Academy of Sciences USA*.
- Shen, Y., & Aisenman, R. (2008). “Heard melodies are sweet, but those unheard are sweeter”: Synaesthesia and cognition. *Language and Literature*, 17, 101–121. [doi:10.1177/0963947007088222](https://doi.org/10.1177/0963947007088222)
- Spence, C. (2011a). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971–995. [doi:10.3758/s13414-010-0073-7](https://doi.org/10.3758/s13414-010-0073-7)
- Spence, C. (2011b). Wine and music. *The World of Fine Wine*, 31, 96–104.
- Spence, C. (2012). Managing sensory expectations concerning products and brands: Capitalizing on the potential of sound and shape symbolism. *Journal of Consumer Psychology*, 22, 37–54. [doi:10.1016/j.jcps.2011.09.004](https://doi.org/10.1016/j.jcps.2011.09.004)
- Spence, C., & Deroy, O. (2012). Crossmodal correspondences: Innate or learned? *i-Perception*, 3, 316–318. [doi:10.1068/i0526ic](https://doi.org/10.1068/i0526ic)
- Williams, J. M. (1976). Synthetic adjectives: A possible law of semantic change. *Language*, 52, 461–478. [doi:10.2307/412571](https://doi.org/10.2307/412571)