

Crossmodal Harmony: Looking for the Meaning of Harmony Beyond Hearing

i-Perception

2022, Vol. 13(1), 1–40

© The Author(s) 2022

DOI: 10.1177/20416695211073817

journals.sagepub.com/home/ipe



Charles Spence 

Crossmodal Research Laboratory, University of Oxford, Oxford, UK

Nicola Di Stefano

Institute for Cognitive Sciences and Technologies, National Research Council of Italy (CNR), Rome, Italy

Abstract

The notion of harmony was first developed in the context of metaphysics before being applied to the domain of music. However, in recent centuries, the term has often been used to describe especially pleasing combinations of colors by those working in the visual arts too. Similarly, the harmonization of flavors is nowadays often invoked as one of the guiding principles underpinning the deliberate pairing of food and drink. However, beyond the various uses of the term to describe and construct pleasurable unisensory perceptual experiences, it has also been suggested that music and painting may be combined harmoniously (e.g., see the literature on “color music”). Furthermore, those working in the area of “sonic seasoning” sometimes describe certain sonic compositions as harmonizing crossmodally with specific flavor sensations. In this review, we take a critical look at the putative meaning(s) of the term “harmony” when used in a crossmodal, or multisensory, context. Furthermore, we address the question of whether the term’s use outside of a strictly unimodal auditory context should be considered literally or merely metaphorically (i.e., as a shorthand to describe those combinations of sensory stimuli that, for whatever reason, appear to go well together, and hence which can be processed especially fluently).

Keywords

harmony, crossmodal, multisensory, audition, pairing, color, flavor, processing fluency

Date received: 29 September 2021; revised: 20 November 2021; accepted: 23 December 2021

Introduction

Harmony is an especially important concept in the fields of music (Lippman, 1964, 1990; Scholes, 1980, pp. 441–454), painting/color theory (e.g., Burchett, 2002; Picabia, n.d.), and aesthetics more

Corresponding author:

Charles Spence, Department of Experimental Psychology, University of Oxford, Anna Watts Building, Oxford OX2 6BW, UK.

Email: charles.spence@psy.ox.ac.uk



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

generally (e.g., Argüelles, 1972; Brain, 2015, pp. 129–131; Tatariewicz, 1980). In musical terms, harmony has been defined as “the clothing of melody” (Scholes, 1980, p. 441) or “the study of simultaneous sounds (chords)” (Schoenberg, 1983, p. 13). Auditory harmonies and rhythms may be consonant or dissonant (e.g., Krebs, 1987; Rehding, 2019; Van de Geer & Levelt, 1962; von Helmholtz, 1954),¹ though rhythmic consonance will not be the primary focus here. The preference for harmonic musical consonance would appear to emerge very early in human development (Masataka, 2006; Schellenberg & Trehub, 1996; Trainor & Heinmiller, 1998; Trainor et al., 2002; Weiss et al., 2020; Zentner & Kagan, 1998; though see also Plantinga & Trehub, 2014). What is more, newly-hatched chicks have been shown to exhibit a preference for consonance (Chiandetti & Vallortigara, 2011), as have Japanese monkeys (Izumi, 2000; though see Wagner et al., 2020). That said, evidence for a role of enculturation has also been suggested on the basis of the results from the small number of studies with those rare and remote populations that do not appear to exhibit a preference for consonance (McDermott et al., 2016; Prete et al., 2020).

When used in the context of color perception, the definition of harmony immediately becomes rather more controversial (Birren, 1967). According to Judd and Wyszecki (1975, p. 390): “When two or more colors seen in neighboring areas produce a pleasing effect, they are said to produce a color harmony.” At the same time, however, it is commonly acknowledged that there are a wide diversity of responses in terms of what harmony means in relation to the aesthetic aspect of color combinations (see Burchett, 2002; Jacobson et al., 1948). As Burchett (2002, p. 28) notes: “Sufficient questions have been raised about color harmony to justify analysis of its meaning. Hardly is anything written about it when some mention is not made of its relative state of misunderstanding and incompleteness or its unruly complexity.” Burchett’s careful analysis of 12 especially influential books on color/color theory written over the last 250 years revealed that an average of close to 5% of the text was given over to discussing the concept of color harmony, thus highlighting the topic’s importance to color theory. Seemingly in contrast to the case of musical harmony though, it has been suggested, albeit without any empirical support, that visual harmony is learnt (Burnham et al., 1963).

While writing/research on harmony in music (e.g., Rameau, 1722; Scholes, 1980; Zarlino, 1558) and painting (i.e., the visual arts) goes back centuries (e.g., Field, 1835, 1845; Hay, 1836), to date, there have been only very limited attempts to extend the notion beyond these two core artistic domains. Writing in 1878, The North American expatriate painter James McNeill Whistler suggested that his paintings be called “harmonies” (see McNeill Whistler, 1978). Intriguingly, however, the notion of harmony has recently been put forward as one of the guiding principles (or outcomes) used by those practitioners wanting to combine flavor experiences (as when pairing the food and wine on a tasting menu; Eschevins et al., 2018, 2019; Spence, 2020c). According to Spence (2020c, 2020d), the search for harmony in such cases should be considered as a perceptual, rather than an intellectual, principle underpinning the decision to pair particular combinations of sensations, regardless of whether they happen to be presented in the same modality, or else in different sensory modalities.

In fact, there has been growing interest in the possibility of crossmodal harmony. It should, though, be noted here how such a suggestion can be seen as presupposing that similarity relations can be established across the senses, something that not all researchers necessarily agree is even possible. For example, long ago, Hermann Helmholtz, the eminent German psychophysicist, famously suggested that: “The distinctions among sensations which belong to different modalities, such as the differences among blue, warm, sweet, and high-pitched, are so fundamental as to exclude any possible transition from one modality to another and any relationship of greater or less similarity. For example, one cannot ask whether sweet is more like red or more like blue. Comparisons are possible only within each modality; we can cross over from blue through violet and carmine to scarlet, for example, and we can say that yellow is more like orange than like blue!” (von Helmholtz, 1878/1971, p. 77; though see also Hartshorne, 1934).

At the same time, however, a less controversial usage of the term harmony is simply to use it to refer to a crossmodal correspondence for a putatively “amodal” feature (though see Spence et al., 2013, on the problematic use of this term), such as the perceived location of paired visual and tactile stimuli (Stratton, 1899). According to Stratton, the spatial harmony between the seen and felt location of visual and tactile stimuli is likely based on associative learning. Stratton demonstrated that spatial harmony can be rapidly re-established in those wearing a pair of distorting prismatic spectacles.

Review Outline

In this review, we will examine attempts to extend the notion of harmony far beyond the unisensory auditory case (hence the “Harmony beyond hearing” of the title). We start by examining the original use of harmony in the domain of music theory and perception, especially referring to the key notion of consonance/dissonance. Then, we summarize the ways in which the term “harmony” has been used in the literature on visual (e.g., color, and color and form) perception, and how it is used in the context of combining vibrotactile frequencies (Yoo et al., 2014), fragrances (e.g., Ackerman, 1990; Piesse, 1891), and even when pairing flavors (e.g., Chartier & Kitano, 2019; Spence, 2020c). All of these examples of harmony are intrasensory. Thereafter, we go on to examine the various ways in which harmony has been introduced in a crossmodal, or multisensory, context. We consider the combination of music and painting (e.g., in the rich literature on the ultimately unsuccessful ideal of “color music”; Plummer, 1915; Rimington, 1915; Sullivan, 1914; Zilczer, 1987), music and scent (e.g., in the occasional development of scent organs; see Spence, 2021, for a review), and the emerging literature on “sonic seasoning” with off-the-shelf, or increasingly bespoke, musical compositions and soundscapes being paired with specific tasting experiences (e.g., see Spence, 2020d, 2022; Spence et al., 2021). One of the key questions here becomes whether harmony can be experienced crossmodally, or whether instead it is only ever experienced within individual senses. As we will see later, though, the answer to this question hinges on which of the various meanings of harmony one happens to be using.

There are several key questions at stake in this review. As we try to address the fundamental question of whether the use of the term “harmony” should be considered as anything more than merely metaphorical when applied outside of the auditory modality, the following related questions will also be addressed: (1) What is harmony? Is it fundamentally an acoustic/musical phenomenon, or does it exist for the other senses, and/or across the senses as well? If so, across which particular pairs of senses does it apply? (2) What are the different possible cognitive components of harmony? Does it refer to those stimuli that can be processed especially fluently? Does it refer to those stimuli that are seen as going well together? And/or does it refer to those stimuli that are pleasurable to experience? (3) Are different components of harmony present for some senses, or crossmodal combinations, but not others? (4) Does this imply, therefore, that people are using the word “harmony” to refer to slightly different phenomena across the different senses? Ultimately, one can ask whether it implies that the use of the term harmony is merely metaphorical for certain senses and, if so, for which ones?

Harmonic Sounds

On the Early History of Harmony

In the Greek tradition, the term “harmony” originally referred to the physical unification of different elements (Lippman, 1963). Deeply rooted in metaphysics and cosmology, the concept started to designate a more general and qualitative agreement, also assuming a positive rather than neutral

meaning. Used in this way, the concept of harmony found a natural application in the domain of music, conceived of as an artistic practice based on the juxtaposition of different auditory elements (i.e., sounds) fitted together to generate pleasant effects. The term “music” became more or less synonymous with “harmony” (Lippman, 1963), and early harmony, or music, theorists started to classify intervals (i.e., the combination of different sounds) in terms of their perceptual effects. They introduced a fundamental distinction which is still key to any discussion of Western harmony today, namely that between consonance (i.e., sounds that go well together), and dissonance (sounds which give rise to harsh, or rough, auditory sensations). According to Turek (1976), consonance tends to suggest a feeling of stability and repose in listeners, whereas dissonance often suggests a feeling of tension instead.

Traditionally, the first explanation of this basic auditory phenomenon has been attributed to Pythagoras (6th Century B.C.E.) who discovered that when auditory frequencies in small-integer ratios were combined, they gave rise to a perception that was harmonious. In particular, he discovered that the first four integers result in perfect consonances—i.e., octave (2:1), fifth (3:2), and fourth (4:3). For Pythagoras, this empirical evidence proved a metaphysical/cosmological assumption, namely, that numbers are the principles of reality and that abstract numerical relationships shape our perception of the world (Riedweg, 2005). To the Pythagoreans, musical harmony represented the paradigm case of a higher-order universal harmony and the creation of the cosmos from the primordial chaos could be understood through number. Pythagorean approaches to harmony were subsequently proposed by Zarlino (1558), who included the numbers 5 and 6 in his model thus admitting the major third (5:4), minor third (6:5), and major sixth (5:3) among consonances. Kepler (1619) considered consonant ratios as the basis of both the cosmos and music. A little over a century later, Euler (1739) formulated a rigorous algorithm to derive the degree of consonance—i.e., *gradus suavitatis*—of a given interval (see also Pesic, 2014).

Neural Underpinnings of Consonance

Scientific investigation of the basis of harmony began in the Nineteenth Century with the empirical work of those experimental psychologists interested in studying music perception (e.g., Stumpf, 1883/1890; see also Hui, 2013).² Over the last couple of decades, a growing body of scientific evidence has converged on the suggestion that the prominent role of auditory consonance in perception might be rooted in properties of the human auditory system, in which consonant stimuli are processed more rapidly than dissonant sounds (e.g., Cousineau et al., 2012; Crespo-Bojorque et al., 2018; Tramo et al., 2001). Researchers have therefore hypothesized that the discrimination of consonance may well have a biological basis (Cousineau et al., 2012; Gill & Purves, 2009; Perani et al., 2010), and that the biological advantage for consonance can be explained in terms of the underpinning neural processing (Bidelman & Heinz, 2011; Bidelman & Krishnan, 2011; Bones et al., 2014; Fishman et al., 2001; Foss et al., 2007; Kadia & Wang, 2003; Minati et al., 2009; Tramo et al., 2001). This account has been supported by research demonstrating that the processing of consonance starts early in information-processing in human auditory cortex and that additional neural resources are needed to encode and discriminate dissonant, as compared to consonant, chords (e.g., Brattico et al., 2009; Crespo-Bojorque et al., 2018; Tervaniemi et al., 2011; Virtala et al., 2013).

Tabas et al. (2019) measured the neuromagnetic activity evoked by dyads with varying degrees of consonance or dissonance. The results of their magneto-encephalography (MEG) study revealed that dissonant dyads evoke a pitch onset response (POR) with a latency that was up to 36 ms longer than consonant dyads. Meanwhile, Anderman et al. (2020) reported that the POR is strongly modulated by the degree of consonance and that the latencies of the transient wave peaks in response to consonant dyads are shorter than those elicited by dissonant stimuli. Such findings help to confirm the

suggestion that the auditory cortex requires more time to process dissonant as compared to consonant dyads. At this point, one would naturally be led to investigate the reason why exactly it is that consonant stimuli should be processed more easily and rapidly than dissonant sounds. At least two possible reasons have been put forward in the literature, and both have garnered empirical support, namely “perceptual coherence” and “processing fluency”. It would certainly be interesting to find out more about whether similar neural differences between the processing of consonant and dissonant auditory stimuli would also be observed in those very rare groups of individuals where a preference for consonance has not been observed (McDermott et al., 2016; Prete et al., 2020).

Perceptual Coherence

Perceptual coherence is a property that can be attributed to perceptual stimuli that form a coherent whole and therefore are more likely to be perceived as unitary (Gurwitsch, 1979, pp. 241 and ff.). This concept was popularly debated amongst the Gestalt Psychologists (Koffka, 1935), where it was primarily applied to the visual domain (e.g., Kanizsa figures). Indeed, various neurophysiological studies have revealed that higher gamma band responses are typically associated with more coherent percepts (e.g., see Csibra et al., 2000; Müller et al., 1997). According to the classification of neural oscillatory patterns, gamma-band activity comprises an electroencephalography (EEG) frequency range, >30 Hz, and is distributed widely throughout cerebral structures (Uhlhaas et al., 2011). Gamma band activity is involved in various perceptual and cognitive functions, and a significant increase in gamma band activity has been observed during several perceptual tasks, such as the perception of gestalt-like (visual) stimuli. For example, the results of an MEG study by Kaiser et al. (2004) revealed an enhancement of induced gamma band activity in response to Kanizsa triangles. Similar findings have also been obtained in those experimental protocols in which perceptual coherence emerged as a result of dynamic motion. By recording electro-encephalography (EEG) from participants viewing either a single moving bar on a screen (coherent motion) or else two identical bars moving in opposite directions (incoherent motion), Muller et al., were able to document that enhanced gamma-band activity was elicited by viewing the coherent as compared to the incoherent visual motion stimuli.

Moving on to the case of unimodal auditory perception, it might be hypothesized that the perception of intervals with pitches that merge coherently (i.e., consonances) would give rise to higher gamma band activity than those tone combinations that do not have such coherent properties (e.g., dissonances). The empirical findings that have been published to date are certainly consistent with such a hypothesis, revealing that consonant sounds elicit higher gamma band activity than do dissonant ones (Park et al., 2011; Passynkova et al., 2007), and confirming that dissonant sounds might produce a breakdown of the perception of a unitary auditory stimulus (e.g., see Csibra et al., 2000; Müller et al., 1997). This hypothesis has received further support from the results of a study considering the object-related negativity (ORN), an electrophysiological measure related to the concept of the holistic grouping of stimuli. In particular, according to Itoh et al. (2003), the minor second elicited a significantly greater ORN than the perfect fifth at the P2 latency (160–180 ms).

In line with Stumpf’s (1883/1890) pioneering intuition, the above findings therefore suggest that consonant sounds elicit higher gamma band activity as they constitute more fused, more unitary, and more coherent percepts as compared to dissonant ones. This has further been confirmed by Palmer and Griscom (2013), who describe people’s “preference for harmony” “as an index of the degree to which a person systematically likes (or dislikes) stimuli that are harmonious, in the sense of being “good gestalts” (Palmer & Griscom, 2013, pp. 453–454). These researchers further clarify that this index is used to stand for harmony in the case music and color, but to stand for perceptual goodness,

“good fit”, “good gestalt”, or “Prägnanz” (Garner, 1974; Palmer, 1991) in the case of assessments of spatial composition and “figural goodness” for shape.³

At this point, one might want to ask whether more coherent percepts are also more fluently processed by the sensory system. However, in order to answer this question, we must first introduce the psychological construct of “fluency” as a second factor possibly relevant to the affective connotations associated with the processing of consonant versus dissonant stimuli.

Processing Fluency

According to the processing fluency account, the more fluently a perceiver can process a given stimulus or object, the more positive their aesthetic response will be (e.g., Reber et al., 2004). Due to the broad convergence of findings suggesting that Western listeners have a preference for listening to consonant rather than dissonant stimuli, one might hypothesize that the pleasantness of consonances represents a perceptual correlate of “processing fluency”. Neurobiological findings would certainly appear to support such a suggestion. For instance, studies of the frequency-following response indicate that the perception of consonance versus dissonance might correspond to processing fluency for pitch at the level of the brainstem (Bidelman & Krishnan, 2009). Furthermore, non-linear approaches to auditory perception have indicated a link between consonance and the dynamical features of the signal, suggesting that dissonance is more demanding for the sensory system to process due to the lack of periodicity at various levels of the auditory information-processing system (e.g., see Lee et al., 2009; Tramo et al., 2001). Hence, harmonious stimuli tend to be processed more fluently, and this increase in processing fluency is thought to be part of what makes harmony in audition positively-valenced (cf. Reber, 2012; Reber et al., 1998, 2004).⁴

Harmony in Hearing: Key Take-Aways

To summarize, neurophysiological studies investigating the processing of harmony in the auditory system (Blood et al., 1999), have evidenced an ability to discriminate between consonance and dissonance as well as the faster and easier processing of consonant as compared to dissonant signals. Additionally, cross-cultural studies have revealed that major music traditions around the world tend to make specific use of many of the same harmonic intervals (e.g., Blacking, 1970; Brown & Jordania, 2013; Sato et al., 2019; though see also Athanasopoulos et al., 2021) and that the most frequently used intervals correspond to those considered more consonant by culturally-diverse listeners (Bowling & Purves, 2015; Burns, 1999). Remarkably, archaeological evidence of bone flutes that are capable of sounding consonance suggest dating the discovery of consonance to earlier than the Greek civilization (~8,000 BP from Zhang et al., 1999; ~35,000 BP from Conard et al., 2009). Taken together, the evidence reviewed here would appear to demonstrate that the harmonic features of music, such as consonance and dissonance, are a permanent trait of the perception, production, and appreciation of music by humans (albeit a couple of groups that lack Western-like tonal system apparently don’t prefer consonance when allegedly exposed to it for the first time; McDermott et al., 2016; Prete et al., 2020). As such, a broad range of literature can be taken to suggest that, at least when used in the context of auditory perception, the term harmony is used to describe fused, coherent, and unitary percepts.

To summarize, in audition, harmonious stimuli seem to be characterized by three different cognitive/affective components, namely pleasantness, what one might call togetherness (i.e., how well stimuli go together), and processing fluency. Having reviewed the literature concerning the origins and nature of harmony perception in audition, we are now in a position to extend the discussion beyond hearing. This will help us to address the question of whether the highlighted properties are found to characterize the experience of harmony in other senses and crossmodally. This will

then help us to address the question of whether the use of the term “harmony” in such cases should be considered as anything more than merely metaphorical.

Unisensory Harmony Perception Beyond Audition

Early History of Visual Harmony

The Pythagorean foundation of music harmony inspired the aesthetic theorization of visual artists, sculptors, and architects (Tatarkiewicz, 1980). The link between music and architecture was made explicit in the 1st Century B.C.E. by Vitruvius, who was convinced that “the architect should know music in order to have a grasp of canonical and mathematical relations” (Vitruvius, 1999, I 1:8). For Vitruvius, the harmony of proportions was the unifying principle that linked architecture with sculpture, writing that: “Just as in the human body there is a harmonious quality of shapeliness expressed in terms of the cubit, foot, palm, digit, and other small units; and so it is in completing works of architecture” (Vitruvius, 1999, I 2:4). In the Renaissance, one of the most famous and literal readers of Pythagorean principles was the Italian architect and theorist Leon Battista Alberti. In his treatise *De Re Aedificatoria*, Alberti acknowledged Pythagoras’s role in helping to shape his view of architecture, explicitly declaring that architecture should get its inspiration from numbers and proportions, that are best represented in music: “I am every day more and more convinced of the truth of Pythagoras’s saying, that nature is sure to act consistently, and with a constant analogy in all her operations: From whence I conclude that the same numbers, by means of which the agreement of sounds affects our ears with delight, are the very same which please our eyes and mind. We shall therefore borrow all our rules for the finishing our proportions, from the musicians, who are the greatest masters of this sort of numbers, and from those things wherein nature shows herself most excellent and compleat” (Alberti, 1485/1726, IX:5; see also Siefkes & Arielli, 2015).

Harmony, conceived of as the proportions amongst different parts, has also been demonstrated to affect the perception of visual beauty. One of the most universal principles that has been used to provide some of the world’s greatest art and architecture with balance and harmony is the golden ratio. The ratio (an irrational number approximately equal to 1.618) denotes the division of a segment in two sections in such a way that the longer section is as many times larger than the other, as many times smaller than the whole segment. Systematically adopted across several disciplines, the golden ratio, also known as the “divine proportion” (Livio, 2008), was conceptualized at the time of Pythagoras, but it was used earlier in Ancient Egypt and Babylon (e.g., in the pyramids, see Akhtaruzzaman & Shafie, 2011). Renaissance architects and artists largely used the Golden ratio proportions in eminent works of art, such as sculptures, paintings, and architecture, with the aim of increasing the harmony of artifacts and hence allegedly enhancing their perceived beauty. Such expectations were confirmed empirically by Di Dio et al. (2007), when they investigated the aesthetic effect of modifying the golden ratio of sculptures selected from masterpieces of Classical and Renaissance art. Their findings show that original sculptures were aesthetically evaluated more positively than the modified ones. Besides human-created beauties, the golden ratio has also been found in the natural world, e.g., in the body proportions of living beings and the growth patterns of many plants and insects (see Akhtaruzzaman & Shafie, 2011, for a review).

Color Harmony and Beyond

Harmony has frequently been used to describe the combination of colors (e.g., Allen & Guilford, 1936; Basten, Salvatore, & Kaufman, 1995; Field, 1835; Granger, 1953, 1955; Ou, 2015; Ou & Luo, 2006; Ou et al., 2004; Schloss & Palmer, 2011; Shen et al., 1996, 2000; Westland et al., 2007). In his unpublished treatise on painting, Leonardo da Vinci wrote that harmony can be

“produced by a judicious arrangement of colors” (Leonardo, 1802, 271). In the 19th Century, several treatises were published on the harmony of colors (e.g., Field, 1835, 1845; Hay, 1836). In his pioneering book *Chromatics*, the English chemist Field (1835) talked extensively about colors in musical terms. After presenting tints and hues, he wrote that: “Upon these gradations and successions depend the sweetest effects of colors in nature and painting, so analogous to the melody of musical sounds, that we have not hesitated to call them the melody of colors” (Field, 1835, p. 19), seemingly suggesting that the term harmony is technically, and not merely metaphorically, used in such contexts. Meanwhile, Baudelaire (1889, p. 89) once wrote that: “In color there is harmony, melody, counterpoint”.

More recently, color harmony has been defined as a satisfying (i.e., comfortable, favorite) human response to two or more juxtaposed colors (Burchett, 1991). According to Burchett (2002, p. 28): “Colors seen together to produce a pleasing affective response are said to be in harmony.” In order to try and resolve some of the disagreement in the color harmony literature, Schloss and Palmer (2011, p. 551) distinguished between two similar-sounding concepts, writing that: “We define *pair preference* as how much an observer *likes* a given pair of colors as a Gestalt, or whole. We define *pair harmony* as how strongly an observer experiences the colors in the combination as *going or belonging together*, regardless of whether the observer likes the combination or not” [italics in original].⁵ Their empirical results (see Figure 1), while highlighting the strong correlation between these measures, at least in the case of pairs of foreground and background colors, also shows that certain color pairs may be more preferred without necessarily being judged as more harmonious, and vice versa. Notice the distinction between mere preference for the whole (or Gestalt) and the goodness of the relationship

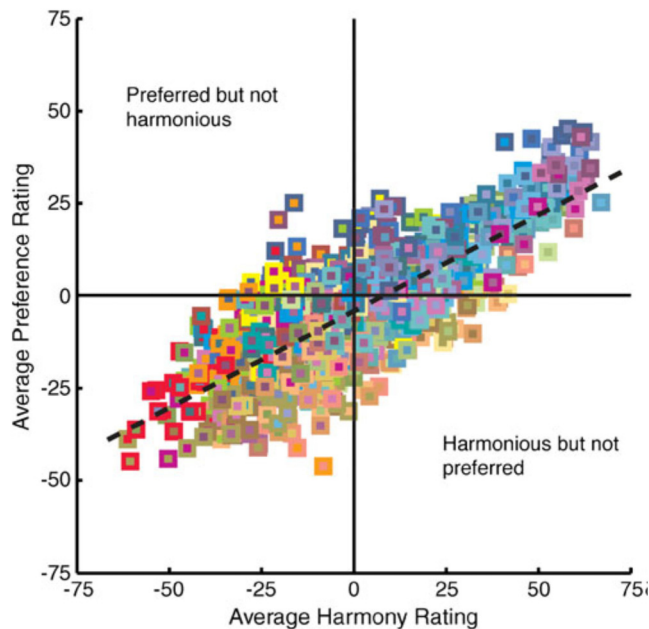


Figure 1. Data from Schloss and Palmer (2011) highlighting the close, but by no means perfect, relationship between people’s preference ratings for color pairs plotted as a function of their harmony ratings. Each of the 992 data points depicts an approximation of the figural color (small square) and ground color (large square behind the figure). The dashed line shows the best fitting regression line relating preference to harmony ($y = -7.93 + 0.52x$). [Figure reprinted from Schloss and Palmer (2011, Figure 8), with permission].

between the component parts. According to Palmer et al. (2013) processing fluency is also an important component of visual harmony.

While the majority of the research on color harmony has addressed the question in relation to static stimuli, it is worth noting that those interested in the color music (delivered by a color organ) that was so popular in the decades around the end of the 19th Century were also concerned with the possibility of achieving harmony by sequentially presenting colored lights (Rimington, 1895, 1911, 1915; Zilczer, 1987). Rimington (1895), for example, uses auditory language (e.g., talking of chords), when discussing the likely consequence of combining colors, or tints, as he describes them, on his color organ. The art movement known as “Synchronism” was also interested in the harmonic connections between colors (see Russell, 1913; South, 2001).

Here, it is important to note that the way in which superimposed (rather than adjacently-presented) wavelengths of light are perceived/combined/mixed is fundamentally different from what happens in the case of soundwaves (Kubovy, 1981, 1988). Interestingly, the term harmony, when used in the context of color perception, therefore tends to refer to spatially separated patches of color, whereas the harmony of combinations of sounds can be determined no matter whether the component stimuli happen to be presented together or separately (either spatially or temporally). In a curiously ambiguous study (ambiguous because of the peculiar method of stimulus delivery) related to Guilford’s research on the harmony of color pairs (e.g., Allen & Guilford, 1936), Spence and Guilford (1933) presented pairs of olfactory stimuli, one to either nostril.⁶ In this case, the pleasantness of the combination (be it of colors or odors) could be predicted by the weighted mean of the pleasantness of the individual stimuli. In the latter case, the equation predicting the affective value of the combination of two odors (C, also referred to as the pleasantness of the affective whole), one pleasant (P) and the other unpleasant (U) was: $C = 0.54 P + 0.69 U - 0.21$. Intriguingly, however, the study authors do not mention harmony once in the olfactory paper, despite mentioning the term frequently in the context of the pleasantness of color pairs.⁷ Rather, they note how sometimes the pairing of odors led to a fusion whereas for other combinations there was a dominance of the pleasantness of one element over the other. Indeed, they talk of the unification or fusion of adjacent odors (one presented to each nostril) as likely been much more common than in case of adjacent colors.

Inner, or Affective, Harmony

The Russian-born artist Kandinsky (1977) wrote extensively on the theme of “inner harmony”⁸ of paintings and music. In language that is more literary than scientific (at least in the English translation), Kandinsky would appear to be suggesting that visual features, such as color and form should coincide in terms of their affective consequences, or effects, in the viewer, rather than necessarily because of any particular perceptual similarity. In this, he can perhaps be seen as prefiguring the development of the “semantic differential technique” (Osgood et al., 1957; Snider & Osgood, 1969) and its use to establish the connotative meaning of stimuli (cf. Parncutt, 2014; Van de Geer & Levelt, 1962; Walker, 2012; Walker & Walker, 2012). The semantic differential technique relates to the suggestion that the connotative meaning of stimuli can be ascertained by means of people’s responses to a selection of line scales anchored by pairs of adjectival opposites, such as “good-bad”, “active-passive”, and “dominant-submissive” (Osgood et al., 1957). At one point, Kandinsky (1977, p. 46) writes: “From the nature of modern harmony, it results that never has there been a time when it was more difficult than it is today to formulate a complete theory,¹ or to lay down a firm artistic basis.” In the quoted footnote, Kandinsky goes on to note that: “Attempts have been made. Once more emphasis must be laid on the parallel with music. For example, of “Tendence Nouvelles,” No. 35, Henri Ravel: “The laws of harmony are the same for painting and music.”“

The notion of inner harmony was key to Kandinsky's (1977, p. 51) ideas around delivering the total work of art, one that incorporated multiple sensory elements (cf. Smith, 2007). He writes that:

“The achievement of the dance-art of the future will make possible the first ebullition of the art of spiritual harmony—the true stage-composition.

The composition for the new theater will consist of these three elements:

- (1) Musical movement
- (2) Pictorial movement
- (3) Physical movement

And these three, properly combined, make up the spiritual movement, which is the working of the inner harmony. They will be interwoven in harmony and discord as are the two chief elements of painting, form and color.”

Unlike Wagner, Kandinsky contends that the three elements of his *Gesamtkunstwerk*—sound, color, and movement—should not have external or narrative connections with one another (Cardullo, 2017). He stressed, instead, that those elements are necessarily unified by an inner unity, which arises out of each one's inner connections with itself as well as with the other elements. Kandinsky stressed the importance of basing these inner connections, not on plot or dramatic action, but on the “inner sounds” of every art and the “inner vibrations” of the audience. As suggested by Cardullo, such concept of “inner sounds” seems closely interwoven with the symbolist theory of synesthesia or “correspondences”. However, given Kandinsky's highly metaphorical use of language, it is difficult to be sure of quite what exactly what Kandinsky is getting at, we have preferred to present his position primarily by means of quotes, so as not to misinterpret his position (though see Cardullo, 2017 for some helpful insights).

On the Contemporary Study of Color-Form Harmony

Beyond the harmony of simultaneously-presented pairs of colors, researchers have also studied the harmony of colors with simple shapes (Kimura et al., 2005). Discussion of preferable combinations of color and shape can be seen as linking to Kandinsky's early studies at the Bauhaus (Jacobsen, 2002, 2004; Kandinsky, 1925). For instance, Kimura et al., conducted a couple of experiments in which they presented various combinations of color and form. The Japanese participants in a first study rated a selection of 12 color patches and 12 simple shapes using a series of semantic differential scales (cf. Hogg, 1969). Thereafter, in a second experiment, a new group of participants rated the harmoniousness of various combinations of color and form ($N = 36$). Taken together, the results of these studies suggested that the combinations of visual features that were judged as being most harmonious were those that were individually associated with a common affective meaning (e.g., according to the semantic differential terms of “lightness” and “activity”; Osgood et al., 1957; Snider & Osgood, 1969). Note that according to the semantic differential approach, these terms would be arranged as polar opposites, namely “active-passive” and “light-dark”. In other words, according to this line of research, pairs of visual stimuli are be rated as harmonious (or not) because of their affective (rather than perceptual) similarity (Collier, 1996; Osgood, 1962; Oyama, 2003; Oyama et al., 1998).

Alternate Definitions of Harmony

Several other definitions of harmony have been proposed in the unisensory visual, and by extension auditory, literature over the years. For example, in a study by Palmer and Griscom (2013), the participants ($N = 90$) rated the harmony of a range of visual and auditory stimuli (color, shape, spatial

location, and music)—127 in total. The visual stimuli consisted of color pairs, dot patterns, and framed dot images (see Figure 2). The auditory stimuli consisted of a range of brief classical music selections. Intriguingly, individual preferences for harmony were found to be strongly correlated across all four stimulus dimensions, and further analysis suggested a common factor underlying the preference for harmony in the various different domains.⁹

Meanwhile, in another study, reported thus far only as a conference abstract, Griscom and Palmer (2011) had their participants rate visual stimuli (color pairs, and single and multiple dot patterns) and short clips of classical piano music in terms of their harmony, and the emotional associations (using happy-sad and angry-calm semantic differential scales). The results revealed strong correlations between ratings of harmony and ratings of positive emotional associations for both the music and color pairs. In other words, the music that was judged as harmonious was also rated as happy and calm rather than as sad or angry. This led Griscom and Palmer to suggest that the consistent cross-domain preferences for harmony that have been reported previously may, in part, reflect a preference for the positive emotional associations evoked by harmonious stimuli (see also Spence, 2020a).

As Palmer and Griscom (2013, p. 460) note, the concept of “harmony” implies a relational aspect.¹⁰ Kandinsky (1977, p. 49) writes that: “The need for coherence is the essential of

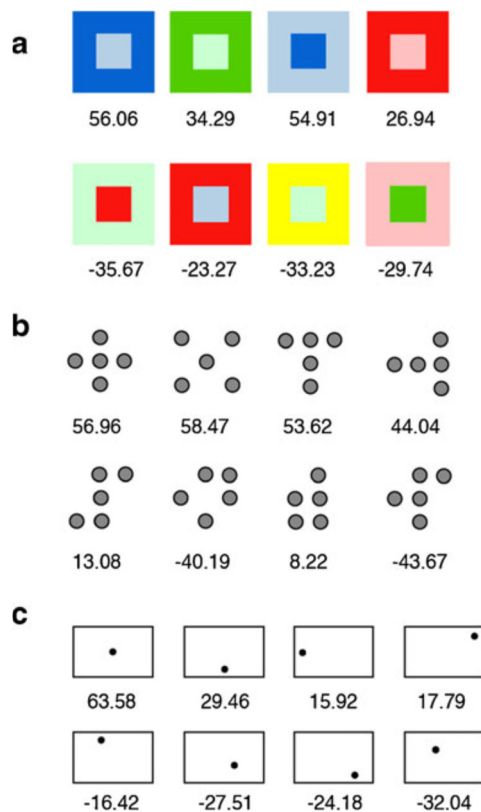


Figure 2. Examples of the visual stimuli used in Palmer and Griscom’s (2013) study of individual differences in the preference for harmonious designs in the auditory and visual modalities: (a) color pairs, (b) dot patterns, and (c) framed-dot images. The numbers below each display indicate its average rated harmony on a scale from -100 to $+100$. [Figure reprinted from Palmer and Griscom (2013), Figure 1].

harmony—whether founded on conventional discord or concord. The new harmony demands that the inner value of a picture should remain unified whatever the variations or contrasts of outward form and color.” The art historical literature on color harmony has been divided between those color theorists who have taken preference to be synonymous with harmony (e.g., Chevreul, 1839/1967; Itten, 1970) and those who have wanted to suggest that harmonious combinations need not be liked, nor dissonant combinations necessarily disliked (Albers, 1963/2013). Schloss and Palmer (2011) demonstrated a strong positive correlation between the average preference (how much participants *liked* a particular combination of colors) and harmony ratings (how well the two colors *went together*, regardless of preference) of 992 pairs of 32 colors rated by 48 participants ($r = +0.79$). At the same time, however, these researchers also found that the corresponding correlations for individuals varied widely, from -0.03 to $+0.70$.

Elsewhere, Japanese researchers have studied textural harmony using a Kansei design approach (Qiao et al., 2014). Kansei refers to an approach to affective design that emerged in Japan and is sometimes known as affective engineering (Nagamachi, 1995; Schütte et al., 2004). The term harmony has therefore been used in a wide variety of unisensory contexts when describing visual stimuli, either individually or, more commonly, in combination. However, the term’s meaning would itself also appear to carry an even broader range of connotations depending on the particular sensory modality to which it is applied.

Vibrotactile Harmony

In many ways, the most natural extension of the concept of harmony in audition is to the vibrotactile sub-modality of touch given that the same soundwaves (albeit with a different perceptible range) give rise to the vibrotactile sensation of touch (cf. von Békésy, 1959a, 1959b; Yau et al., 2009). Indeed, researchers have worked on developing vibrotactile chords, looking specifically for the harmonic consonance of multiple simultaneously-presented tactile stimuli (Yoo et al., 2014).¹¹ For instance, the latter researchers examined the perception of complex vibrotactile stimuli in which a few sinusoidal vibrations with different frequencies were superimposed. Notice how such vibrotactile signals are analogous to musical chords in which multiple notes are played simultaneously.

Yoo et al. (2014) designed a set of “vibrotactile chords” on the basis of musical chords, and the degree to which the participants perceived consonance (harmony) was evaluated. Specifically, 40 participants evaluated the degree of consonance of 80 vibrotactile dyads (i.e., chords consisting of two notes; designed on the basis of musical dyads chords) using a 0–100 scale. The results revealed that the participants were able to rate the degrees of consonance of vibrotactile chords reliably. On the basis of their findings, Yoo et al. succeeded in establishing a well-defined function relating the degree of consonance to the base and chordal frequency of a vibrotactile chord. The subjective impressions associated with vibrotactile consonance and dissonance were smooth vibrational feelings and a rough, fluttering sensation, respectively. This is analogous to what is experienced in the auditory case (cf. von Helmholtz, 1863). That said, in contrast to the study of auditory and visual harmony, it is noticeable how few studies have been published to date on the theme of vibrotactile harmony, perhaps reflecting the fact that there is simply less artistic interest in vibrating the skin senses. What is also noticeable is that it is the pleasantness interpretation that is dominant.

Olfactory Harmony

Harmony is a notion that the chemist and perfumer Piesse (1891) famously evoked in the context of trying to explain why it was that certain combinations of olfactory stimuli appear to combine better than others. In a treatise first published in 1855, Piesse explicitly noted that sounds and odors blend

together similarly, producing different degrees of “a nearly similar impression” in the sensory nerves (Piesse, 1867, p. 39). For example, the mixture needed to prepare the odors for the handkerchief evokes effects on the smelling nerve “similar to that which music or the mixture of harmonious sounds produces upon the nerve of hearing, that of pleasure” (Piesse, 1867, p. 219). Piesse suggests that creating a mixture of scents is like creating a mixture of sounds, i.e., chords: “We have citron, lemon, orange peel, and verbena, forming a higher octave of smells, which blend in a similar manner” (Piesse, 1867, p. 39). The pleasantness of musical harmony resembles that of perfumes. Piesse presented a scale of correspondence between sounds and odors, the so-called *Gamut of Odors* (see Figure 3) being convinced that “there is, as it were, an octave of odors like an octave in music” (Piesse, 1867, p. 38). Note here the underpinning suggestion that there are similar mechanisms within audition as within olfaction, without there necessarily being any claim about the existence, or otherwise, of harmonious sensations that extend across the senses (in this case from olfaction to audition, or vice versa).

When interviewed by Ackerman (1990), the perfumer Sophia Grojsman, suggested that the link between music and perfume, mediated through the notion of harmony, was key to modern perfumery: “Perfumery is closely related to music. You will have simple fragrances, simple accords made from two or three items, and it will be like a two- or three-piece band. And then you have a multiple accord put together, and it becomes a big modern orchestra. In a strange way, creating a fragrance is similar to composing music, because there is also a similarity in finding the “proper” accords. You

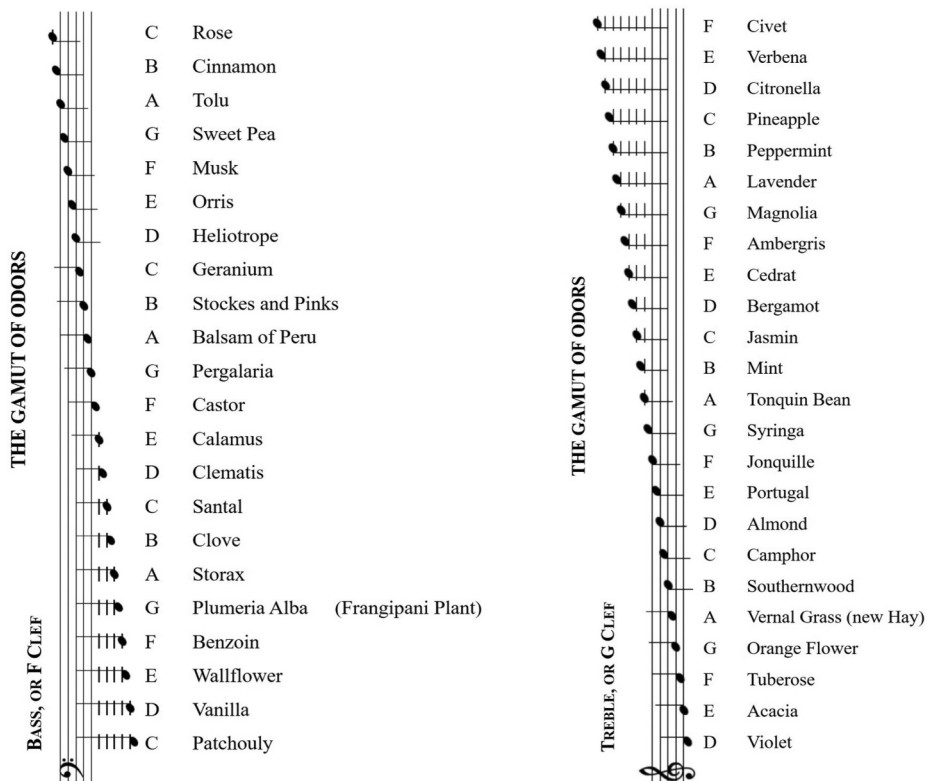


Figure 3. Scale of crossmodal correspondences between sound and odors reproduced from Piesse (1867, pp. 42–43).

don't want anything being overpowering. You want it to be harmonious. One of the most important parts of putting a creation together is harmony" (Ackerman, 1990, p. 49). Spence (2021) also highlighted these and other similarities between musical composition and the creation of perfume.¹² In the olfactory case, then, the sense of harmony would seem to imply a balanced relationship between the component parts. What is more, the suggestion is also that similar perceptual grouping rules apply in the olfactory modality as have been documented previously in the case of audition. At the same time, however, one fundamental difference that is worth stressing here is how the researchers have thus far been unable to discover the numbers, or mathematical relationships, underpinning the harmonious mixing of olfactory stimuli. Instead, harmonious combinations are established on the basis of perceptual experience and trial-and-error. Notice, once again, the focus here is on the more formal aspects of creating harmonious compositions rather than necessarily processing fluency, pleasantness, unity, or any of the other perception-related components (or interpretations) of harmony that have been put forward in the case of auditory and visual harmony.

Harmonious Tastes/Flavors

There has been a recent explosion of interest in the principles underlying the pairing of flavors in recent years (see Eschevins et al., 2019; Spence, 2020c).¹³ Indeed, the desire to harmonize flavor sensations has been put forward as one of the many different principles/strategies behind the pairing of, e.g., food and drink. Analysis of the rationale used by professionals working in the field of gastronomy has recently highlighted the existence of a wide range of intellectual, perceptual, and idiosyncratic reasons for wanting to pair flavors (Eschevins et al., 2019; see also Laudamiel, 2008). At the same time, however, various food scientists have also been known to invoke the concept of harmony when describing those flavor sensations that give rise to what might well be described as a well-balanced dish. For instance, at a flavor conference held in Denmark some years ago, one Japanese researcher gave the following musical metaphor for umami: "Umami is just like the bass note in music. No one listens only to the bass, but the sounds of the bass notes has an important role to give depth and presence to music. Umami likewise creates balance and harmony in dishes." (cf. Ninomiya, 2015; Ninomiya et al., 2010).¹⁴

Harmony in the context of flavor pairing has been defined as "the pleasant effect made by parts being combined into a whole" (Bullon, 1978), or as "how well sensations go together" (Eschevins et al., 2019). It has often been shown to correlate with how much a particular food-beverage pairing is liked (Eschevins et al., 2018; Paulsen et al., 2015; cf. Figure 1 for the analogous correlation observed for color pairs). The idea of pairings giving rise to a particularly harmonious (and thus pleasing) perceptual experience has also featured in the work of those sensory scientists interested in pairing olive oil with leafy greens (Cerretani et al., 2007; Cichelli et al., 2020). In the latter case, the approach is very much based on the notion that congruent tastes and flavors will likely combine harmoniously.

People's ratings of harmony and homogeneity of flavor combinations tends to correlate highly, though the researchers concerned admit that this may reflect nothing more than the fact that their participants may have struggled to discriminate between these terms (see Eschevins et al., 2018; Meillon et al., 2010). Eschevins et al., also make much of the fact that aromatic similarity between food and drink also tends to result in food and beverage pairings that are rated as harmonious.

Elsewhere, according to Kim and Lecat (2017): "Harmony with food" is one of the main reasons given by Koreans to explain why they drink wine. Meanwhile, top chef, Heston Blumenthal, was, at one time interested in investigating the harmoniousness of specific food pairings, as captured by the following quote from Segnit's (2010) *Flavor Thesaurus*: "This led Blumenthal to *research* the harmoniousness of pairings with ... chocolate is like a heightened version of Alain Senderens' *pairing*

of lobster and vanilla. ... or horseradish with mackerel, cucumber with *salmon*, and *watercress* with trout.” In the flavor pairing case though, just like the harmonious combination of odorants, there is no code to predict harmonious pairs of flavors. Instead, successful (or harmonious) combinations would appear to be established on the basis of extensive experience of the distinctive flavor qualities of different foods and ingredients.

The North American winemaker Clark Smith (2010) also invokes the notion of harmony. So, for example, when describing wine texture/astringency, he writes that “great wines are tuned in to harmony” ... “fine nuances of discord and harmony, to which we relate without being told, show up in wines as they do in music.” Smith (2013, p. 140) also suggests that: “A sense of harmony and dissonance is strongly shared.” And, when describing his attempts to systematically adjust the alcohol level of wine, Smith (2010) writes that: “Every wine has very discrete balance points we can all identify, where the astringency abates and the flavors are married and harmonious. The adjacent wines in such a series taste especially disharmonious 0.1% alcohol too high was hot and bitter, and 0.1% too low was harsh and sour.” Here, the term harmony appears to refer to distinctive yet balanced elements. In his 2013 book, *Postmodern winemaking*, Smith suggests that wines with finesse have “unified flavors ... creating a sense of harmony” (Smith, 2013, p. 25). Smith, then, would appear to be using the term harmony in the context of wine, what he is fond of calling “liquid music”, to describe both a unified flavor sensation as well as a balanced combination of elements. This, he contrasts with disharmonious wines that express a fault, such as “corkiness”, or which are overly astringent/harsh, hence suggesting that the elements are somehow out of balance.

That said, it is interesting to note that not everyone who is interested in food or flavor pairing necessarily invokes the notion of harmony. The term is absent from Coucquyt et al. (2020) recent book on *The art and science of Foodpairing*, making no appearance in the Index nor, as far as we can see, in the main text either. Reference to harmony is also pretty minimal in Chartier’s (2012) *Taste buds and molecules: The art and science of food, wine, and flavor*, at least when compared to the color books analyzed by Burchett (2002), mentioned earlier. Only at one point in the text does Chartier briefly talk of “harmonic proof”, and “harmonic families” (Chartier, 2012; pp. 43–44). Chartier (2012, p. 44) further reports that he has surveyed “the greatest possible number of wines and foods with the goal of uncovering new harmonic families This set the stage for a harmonious combination of ingredients ... The process culminated in wine and food pairings that were just as harmonious and yet even more precise than ever before.” However, the suggestion that Chartier is heavily invested in the explanatory value of harmony is suggested by a recent presentation together with Sony announcing the launch of a new SONY AI approach targeted at gastronomy entitled “The aromatic theory of molecular harmonies” (Chartier & Kitano, 2019). To summarize, harmony as the unity of the component stimuli appears to the most common use of the term in the literature on flavor perception.

Crossmodal Harmony

Romanticism and the Harmony of the Senses

Talk of the harmony of the senses has often appeared historically in more of a literary (e.g., von Erhardt-Siebold, 1932; see Di Stefano et al., 2022, for a review), or even spiritual, mystical, or divine, context (Argüelles, 1972; Kandinsky, 1914; Kiltinavičiūtė, 2020; cf. Zhang, 2007). According to Mahvash (2007, p. 57): “The medieval cathedral is another example of such a multi-sensory spatial experience where the acoustic quality of the material and space, together with the massiveness of the structure, dramatic play of light and shadow, and the feel and touch of materials provide a very powerful sense of spirituality through a harmonic manipulation of our sensory experiences.” (cf. Schneider, 1955). Intriguingly, the notion of harmonizing music and architecture also

emerges repeatedly from the collection of essays in a book entitled *Resonance: Essays on the intersection of music and architecture*, edited by Muecke and Zach (see, for example, Sands, 2007; Sterken, 2007).

It is relevant here to consider how in the closing decades of the 19th Century and the opening years of the 20th, it was common for writers, artists, and even scientists to explicitly link their study of the harmony between the senses to the phenomenon of synaesthesia (e.g., Argüelles, 1972; Kandinsky, 1977; Kiltinavičiūtė, 2020). Charles Henry (1859–1926), the influential late 19th Century French scientist, was interested in trying to develop a psychophysical approach to understanding harmony. However, he, like so many of the artists of that period, tended to get sidetracked by searching for solutions in the phenomenon of synaesthesia (e.g., see Argüelles, 1972, p. 96, p. 122, p. 135).¹⁵

For example, in 1891, a pioneering adaptation of the *Cantique des cantiques* of Solomon by Paul-Napoleon Roinard was performed in Paris at the Theater d'Art, to present the novel idea of theater as multisensory art by engaging the audience's senses of sight, hearing, and smell. Inspired by symbolist aesthetics, Roinard conceived a synaesthetic multisensory work in which original music, words, vowel sounds, colors, and scents were to be harmonized (Halperin, 1988, p. 199; Stokes, 1972, p. 167). For each poetic section, Roinard provided the details for the exact combination of music (e.g., "in C"), color (e.g., "pale purple"), and scents (e.g., "frankincense") (Roinard et al., 1976). A total of nine scents (namely, frankincense, white violets, hyacinth, lilies, acacia, lily of the valley, syringa, orange blossom, and jasmine) were released into the theater, while the audience simultaneously listened to words and music.

Roinard wanted to realize the ideal of a synthetic and perfect union of all the arts and the senses, though in this case the *Cantique des cantiques* was savaged by the critics and audience alike (Fleischer, 2012; Shepherd-Barr, 1999). In hindsight it would seem fair to say that part of the problem with these performances stemmed from technical issues associated with the appropriate release (and subsequent removal or dissipation) of the various scents. That said, there was presumably no particular reason to think that the specific idiosyncratic inducer-concurrent relations experienced by a given synaesthete should necessarily be judged meaningful, pleasant, or harmonious in the minds of those other non-synaesthetes who happen to be exposed to them (see Spence, 2012).

According to Argüelles (1972, p. 143): "Psychophysics, as Henry conceived and worked with it, is essentially a metaphor for harmony or the pursuit of harmony, whether this pursuit be defined as a science or an art. Harmony is a general condition which Henry was fond of evoking, and one might well consider him a harmonist: one, who in his own way, espouses the doctrine of harmony." Elsewhere, Argüelles (1972, p. vii) writes of how Henry: "reformulates the doctrine of harmonic unity as psychophysics and the doctrine of harmonic work as the psychophysical aesthetic." Many of the creative individuals of the period were, in other words, interested in both intramodal and crossmodal harmony.

Audiovisual Harmony

In his theorization concerning the senses, Aristotle hypothesized the existence of a link between color harmonies and musical proportions. Aware of the Pythagorean theory of consonances, the philosopher put forward the pioneering speculation that the color combinations most acceptable to the eye might depend upon the same numerical proportions as the musical consonances (Aristotle, 1908, III 439b–440a). Although he was unable to test such a hypothesis, the claim continued to fascinate artists and theorists for a long time thereafter. For example, the works of the Italian painter Arcimboldo (1527–1593) have been interpreted by contemporary critics as being directly inspired by Pythagorean musical harmony.

Caswell (1980, p. 156) writes that Arcimboldo's paintings are a "thoroughly scientific excursion into the twin realms of sight and sound conducted for the purpose of accurate mathematical measurements of the intervals found in both sensory domains by means of the Pythagorean ratios derived from music". Caswell suggested that the linear structure of an entire polyphonic composition can quite literally be "seen" as a pattern of interwoven lines of light and color, and quotes a significant passage in which Comanini (1590), a humanist who was very familiar with Arcimboldo's work and commented upon it extensively, stated that Arcimboldo "has located the tones, semitones, the diatesseron, the diapente, the diapason, and all the other musical consonances in colors, using the formulae which Pythagoras invented to define the same proportions in harmony" (Caswell, 1980, p. 157).

Arcimboldo's paintings therefore represent a case in which a visual artist uses musical criteria as a basis for his/her creations. In other cases, a composer might be inspired by aesthetic proportions that pertain to visual art, such as happened for the *Miserere Mei Dei* by Gregorio Allegri. Traditionally, the *Miserere Mei Deus* was only ever presented as part of the Easter performances given in the Sistine Chapel, whose ceiling was painted by Michelangelo (Grovier, 2012). In the arrangement of Psalm 51, Allegri incorporates this ratio by pairing two chapel choirs of four and five singers, respectively. According to a commentary by Kelly Grovier, such an arithmetical ratio is also mirrored in the design of the chapel's ceiling:

"The ratio of voices that invigorates the structure of the choral work also mirror's Michelangelo's complex design for the chapel's ceiling, as though it had been deliberately conceived as a soundtrack for the space. The compositional recurrence of 5:4 imposes schematic sense on the artist's celebrated illustrations from the Bible. The ceiling's central spine is divided into nine scenes from Genesis, and these nine in turn organize themselves into an alternating sequence of five minor frescos separated by four major ones. Further amplifying the asymmetry of Michelangelo's plan is the painter's decision to surround each of the five major panels with four so-called Ignudi—reclining nude figures which many have surmised to be angels."

One might presumably want to consider the common use of the 5:4 ratio in the musical arrangement and spatial composition of the painting as presenting a kind of crossmodal structural correspondence (e.g., Abbado, 1988; Caivano, 1994; Pridmore, 1992; Schöffner, 1985; Sebba, 1991; Spence, 2020a; Wells, 1980; see also Beck, 1999, 2005; Walker, 1967).¹⁶ Like Newton (and Piessé) before him, Wells suggested that fruitful parallels could be drawn between the mixing of colors (or scents), and of tones, in terms of the deriving of combinations that are perceived as harmonious. Handel (1988a, 1988b, p. 315) is therefore right to point out that: "The correspondences among the senses and among perceptual experiences have been an intriguing issue throughout the history of philosophy and psychology."

The structural correspondence between the 5:4 ratio found in the choral arrangement for the *Miserere Mei* sung in the Sistine Chapel in Italy under Michelangelo's frescos where the 5:4 arrangement is multiply repeated has been described as presenting a kind of secret (possibly cross-modal, or multisensory) harmony, that would have been lost were the *Miserere* to have been presented elsewhere (see Grovier, 2012).¹⁷ However, here, it is worth stressing that the notion of genuinely multisensory Gestalts, at least amongst the spatial senses (e.g., audition, vision, and touch) have proved exceedingly difficult to document convincingly (though see Huang et al., 2012, for one of the only examples, involving audiotactile integration in the perception of musical meter). Kubovy and Van Valkenburg (2001) suggest that the only place where one comes across what they describe as "transmodal audiovisual Gestalts" is in the case of speech (see Sato et al., 2007).

As such, it is by no means certain that the analogous structure presented to eye and ear would necessarily have been recognized by the majority of those who were fortunate enough to attend the performances of the *Misere Mei* that were performed during the *Tenebrae* liturgy on the Wednesday and Friday of Holy Week. More generally, it has certainly remained an open question amongst those experimental psychologists interested in Gestalt psychology as to what role attention may play in helping to establish crossmodal relations between unimodal patterns and/or the emergence of patterns that are perceptible only once the inputs from the various senses have been combined, or grouped (see Aksentijević et al., 2001; Handel & Buffardi, 1968, 1969; Kimchi & Razpurker-Apfeld, 2004; Kubovy & Pomerantz, 1981; see Spence, 2015, for a review). Part of the problem here for theorists has been to try and figure out the most appropriate crossmodal mapping between audition and vision. According to Handel, it is likely context dependent, such that while the visual spatial to auditory-temporal mapping might work in the case of form perception, there are likely other situations in which a different mapping, or equivalence is more appropriate (Handel, 1988a, 1988b; Kubovy, 1988). Over the years, psychologists have proposed various plausible equivalences between the senses (Julesz & Hirsh, 1972; Kubovy, 1981; Marks, 1978), with Kubovy fond of the suggestion that visual space-time may be analogous to auditory pitch-time. Of course, if one takes the space to time analogy, this necessarily means that it is mostly the grouping of a temporally-presented sequence of elements, and hence brings us back to the rhythmic harmony which is not the primary concern of the present review (though see Rozin & Rozin, 2018, for a meditation on the various analogies between the temporal sequencing of auditory stimuli and flavor perception/cuisine).

Another important point to be addressed here is whether the presentation of matching, similar, or harmonious combinations to the eye and ear gives rise to mere repetition, or pleonasm (cf. Banes, 2001, critical assessment of the pleonastic use of scent to simply re-present some of what is seen on stage in the context live performance).¹⁸ More optimistically, however, the harmonious presentation of stimuli to different senses might lead to some kind of modulation (Spence, 2015) or resonance (Muecke & Zach, 2007; cf. Smith, 2013) instead (see Spence, 2020a, 2020b, for reviews of a number of the outcomes that have been documented when music and visual stimuli are combined). It must remain an open question as to which of these responses would be elicited today, or in the centuries gone by, by listening to the performance of *Misere Mei* while seated under Michelangelo's frescos. The experience of some sort of crossmodal, or multisensory, Gestalt, or harmony, certainly should not be assumed.

Inspired by Newton's *Opticks* (1704), in which the scientist proposed a distribution of the seven colors found in prismatic light into a scale based on the Pythagorean ratios, several theorists conceived color painting based on musical criteria. One of the most interesting case is the already mentioned *Chromatics* by Field (1835), in which the author not only talked in general about the harmony of colors, but went farther defining harmony as "the complex accordances of three or more sounds or colors in consonance, opposition or contrast" (Field, 1835, p. 20). Field defines consonant colors as the ones that "when compounded or placed together produce a third pleasing tone of color or agreeable effect on the eye" (Field, 1835, p. 21). By contrast, dissonant colors produce a "disagreeable effect to the eye" (Field, 1835, p. 22) and are accordingly dissonant. Then, he strengthened the link between colors and sounds, by relating the classification between primary, secondary, and tertiary colors with the positions of sounds as treble, tenor, and bass (see Field, 1835, p. 25). Such an effort at theorization results in the conceptualization of color combinations in terms of consonance and dissonance (see Figure 4a) thus leading to the definition of the general mapping between sound and color (see Figure 4b), one of many possible physical correspondences that have been suggested over the years.

The systematic use of musical notions such as keys, melody, intervals, consonance, dissonance, inversion, and scale for describing harmony of colors, likely make Field's system the most rigorous

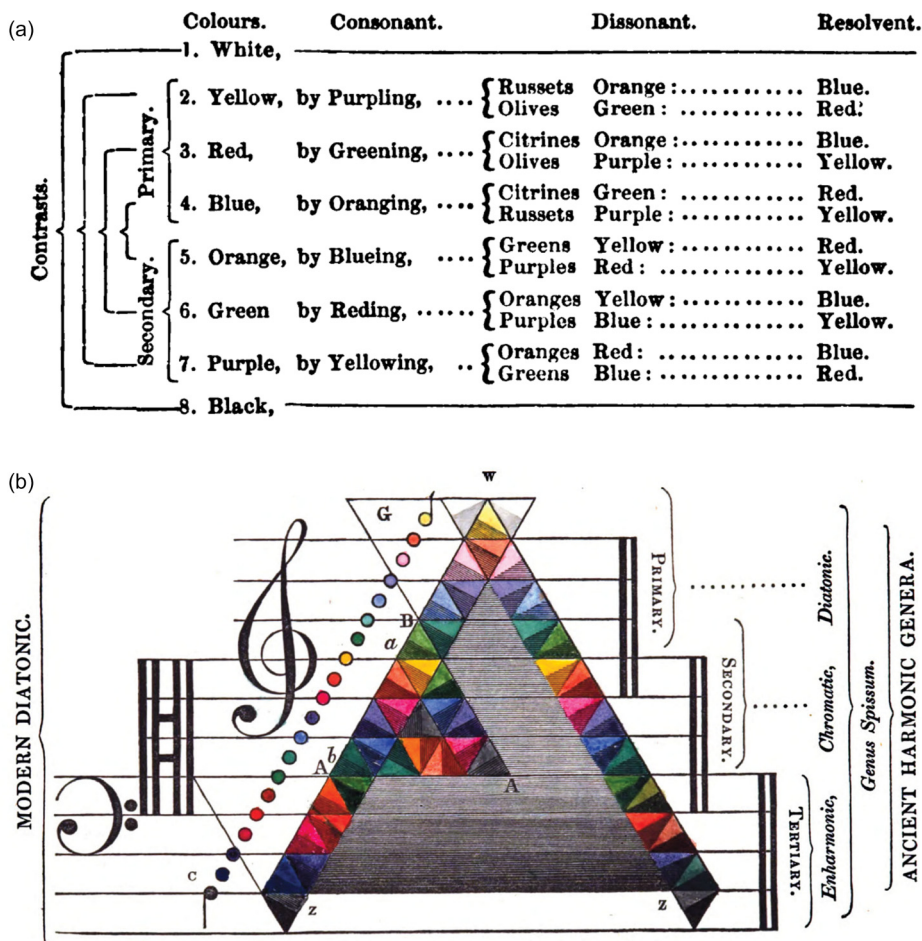


Figure 4. (a) Table of chromatic consonances and dissonances from Field (1835, p. 36). Going far beyond the general claim of the similarity between colors and sound, Field wanted to provide a musical classification of color combinations based on the notion of consonance/dissonance. Color combinations are thus listed depending on their consonant or dissonant effect therefore obtaining “the concords, discords, and expression of colors in the harmonic relations of musical sounds” (Field, 1835, p. 36). (b) The correspondence between colors and musical sounds as theorized by Field (1835, p. 79). Such an “entire analogy and perfect correspondence of the chromatic and harmonic systems” (Field, 1835, p. 79) leads Field to highlight the link between specific key elements of music theory (e.g., diatonic and modal scales) and color theory (e.g., primary, secondary, and tertiary colors). Each colored triangle is divided into two equal triangles of slightly different hues that correspond to the chromatic intervals (represented also in circles on the left). Pitches are ordered from low to high, with darker and lighter hues, respectively.

and musically informed attempt to achieve a general theory of harmony which applies equally to music and sounds. In addition to theorists, artists also imagined that ensuring color harmony, to match the underpinning musical harmony, would be an important element of audiovisual performances based on the use of color organs (Plummer, 1915; Rimington, 1911, 1915; Sullivan, 1914; Zilczer, 1987; see Spence, 2020a, 2020b, for reviews). For example, Rimington insisted that musical harmonies have their analogue in color and proposed that the same consonant ratios in music will produce pleasing color combinations when applied to frequencies of the spectrum (Rimington, 1911).¹⁹

At around the same time as the popularity of the color organ was at its peak, Scriabin suggested introducing a colorful accompaniment to his musical score Prometheus—*Poem of Fire (Opus 60)*. Although never realized in his lifetime, subsequent analysis has been taken to suggest that Scriabin's intended use of color had been to disambiguate the music itself, rather than being an expression of his possible synaesthesia (see Gawboy & Townsend, 2012; Secor, 2013).²⁰ Wells (1980, p. 103) writes that: "One has only to trace the remarkable unfolding of harmonic color in Scriabin's later works to realize that it was harmony above all that prompted him to seek a visual counterpart to his music. An equivalence of aural harmony and visual is so strongly implied in Scriabin's later works that for me they seem to emanate from a single source."²¹

Relevant evidence here may also come from the fields of film music, as well as certain forms of theater, in which extensive work has asked how best to combine dynamic auditory-visual stimuli (e.g., movies with musical scores; Bashwiner, 2013, pp. 109–110; Tan et al., 2013). Argüelles (1972, p. 154) writes of Walt Disney's *Fantasia* that it "may yet be the most perfect, the most successful, and the most universal in its appeal of all the harmonic artistic efforts in the twentieth century ... creating startling combinations of synchronized visual/musical abstracts interwoven with fairytales. The work may be considered close to the fulfillment of the symbolist *Gesamkunstwerk* ideal."

Oskar Fischinger is also well-known for his performance of color music and for his animations that were meticulously designed to accompany recorded music. Fischinger attempted to provide a visualization of the music, through a very close correspondence to the formal properties of beats and note onsets. Moreover, size correlated with amplitude; color and shape with timbre and instrumentation. A notable example of this work is "An Optical Poem" in which animated object accompany Franz Liszt's Hungarian Rhapsody No. 2 (see Brougher et al., 2005).

Visual music is not only composed for live performance and color music performances, but is now available in recorded form through film, video, and computer technologies. For example, the films of Mary Ellen Bute echo the synaesthetic spirit of performance combining color with music, relying on both color and form to accompany and intertwine with music (see e.g., Basquin, 2020). Today, a popular form of visual music involves the algorithmic visualizations that are generated by media players, in real-time correspondence with live or recorded music. However, looking to the future, developing harmonious combinations of multisensory stimuli will likely feed into the emerging interest in "Sensploration" (Leow, 2015) and multisensory experiential events such as, for examples the Tate Sensorium in 2015 (Pursey & Lomas, 2018).

Meanwhile, the composer and writer Bill Alves, in several of his works, e.g., *Static Cling* (2000) and *Stellation* (2008), systematically combined Pythagorean ratios and symmetry to create what he claimed was harmony in visual music (Alves, 2005, 2012; see also Whitney, 1980). While much of the work in the area broadly defined under the header "color music" (see Zilczer, 1987) has indeed been concerned with the harmonious mapping of auditory and visual stimuli, this has by no means always been the sole motivation for those wanting to combine the senses (see Spence, 2020b). Indeed, in terms of the contemporary literature on crossmodal correspondences involving the emotion-based mapping of music to visual stimuli, such as color patches, it is noticeable how "harmonious-disharmonious" is merely one of the many semantic differential scales on which people have been asked to rate the stimuli (e.g., short music clips and color patches in a study by Whiteford et al., 2018). Although not stated explicitly, one might therefore consider the emotional-mediation of such crossmodal correspondences between complex and affectively-valenced stimuli, as a crossmodal variant of Kandinsky's inner harmony. Or as Marks (1978, p. 181) once put it, sensory qualities "talk over their common feeling."

Harmonizing Music and Taste/Flavor

Those scientists and practitioners working in the emerging field of "sonic seasoning", where flavors/aromas are deliberately paired with music (see Spence, 2020d), is sometimes explicitly based on the

perceived harmony across the chemical senses and audition. As a starting point for discussion in this regard, it is interesting to note how wine writers have, on occasion, attempted to describe the harmonious taste/ flavor of a particularly pleasant wine in terms of musical harmony. Just take, for example, the following quote from the famous British wine writer Hugh Johnson: “I have tasted first-attempt Chardonnays that were like Dizzy Gillespie’s solos: all over the place. And the color of his trumpet, too. On the other hand, a Stony Hill Chardonnay recently had the subtle harmonies and lilting vitality of Bix Beiderbecke.” (Johnson, 2005, p. 253). Or how about the following from North American winemaker Clark Smith: “there is strong evidence that the qualities of harmony and dissonance are as mutually perceived in wine as they are in music.” (Smith, 2013, p. 25).

That being said, it is important to note that just because similar aesthetic, or Gestalt, phenomena exist within different senses that does not necessarily imply that an equivalent phenomenon will necessarily be experienced between, or across, the senses as well. One of the few suggestions of a genuine crossmodal experience of harmony appears in Holt-Hansen’s (1968, 1976) early research on the “pitch of harmony” between the taste of beer and a pure tone.²² In his seminal early work in Copenhagen, Kristan Holt-Hansen demonstrated that people would match Carlsberg Elephant lager to a higher pitch than regular Carlsberg lager (perhaps because of its higher alcohol content). When participants tasted a beer while a matching sound (a tone at the pitch of harmony) was played, they reported a variety of sensational experiences such as “outspoken delight, bodily harmony, and relaxation.” Interestingly, Holt-Hansen (1976) explicitly talks of the sensations of sound and taste becoming unified at the pitch of fit/harmony (though see also Rudmin and Capelli, 1983, for only a partial replication of Holt-Hansen’s early findings).

Smith (2013, pp. 139–140), though, would seem to be convinced that the music played while tasting a wine can help to upset its balance, making it appear smoother or harsher. Wang and Spence (2016) conducted a study that was designed specifically to investigate the consequences of manipulating the harmonic content of background music on taste perception. The participants had to evaluate samples of mixed fruit juice whilst listening to soundtracks that had either been harmonized with consonant or dissonant musical intervals.²³ The participants consistently matched the consonant soundtracks with sweetness and the dissonant soundtracks with sourness. What is more, the juices were rated as tasting significantly sweeter in the consonant than in the dissonant music condition. These results therefore support the claim that the crossmodal correspondence between a higher level musical attribute (namely, harmony) and basic taste can be used to modify the evaluation of the taste/ flavor of a drink. That said, it is possible that the crossmodal influence of auditory harmony on taste might well be mediated by the relative pleasantness of the two sounds, rather than necessarily having anything specifically to do with consonance/ dissonance *per se*.

Nevertheless, a growing body of empirical research now demonstrates the significant impact that music can have on the tasting experience of a wide range of food and beverage products. That said, the matching of the emotional tone of music and wine, might perhaps be described in terms of Kandinsky’s “inner harmony”. In another example of sonic seasoning, synaesthetic composer, Nick Ryan, created soundscapes to match three wines (a cava and two rojas).²⁴ As the composer noted, sweet fruity wines “need more harmonious and complex compositions” (Knapton, 2015). This took place as part of the Campo Viejo Streets of Spain festival in London. According to Knapton (2015, p.7), one of the journalists who reported on the event “The first volunteers to try listening to the scores while drinking said that they felt physically transported to a different place. Others wept.”

Harmonizing Scent and Music

In recent decades, there have been a very limited number of more-or-less artistic attempts to combine scent and music, often inspired by Piesse’s scent scale (see Spence, 2021, for a review). Such

creations, and the associated multisensory performances, once again raise the question of whether auditory harmony can be meaningfully captured olfactorily, and whether any crossmodal, or multisensory, Gestalt may emerge (Gilbert, 1938; Spence, 2015). According to the aroma jockey (this is someone who synchronizes the release of various fragrances in time with the music) Erich Berghammer: “In order to interpret popular songs with scent, you need to be able to listen to a song with your nose, meaning you naturally choose scents that harmonize with what you’re hearing.” (as quoted in Chesters, 2017). Elsewhere, scent and music have, on occasion, been deliberately paired in order to help illustrate the correspondences, or perceptual similarity between scent and sound (e.g., see Spence, 2021, for a review). However, it is worth stressing here that the crossmodal mapping that has been used in the majority of such cases is based on the untested suggestions found in Piesse’s *Gamut of Odors* (see Figure 3) rather than on the scientifically-validated crossmodal correspondences that have been established by researchers over the last half a century (see Velasco & Spence, in press).

For example, as part of an ongoing collaboration between Sean Francis Conway and Brian Goeltzenleuchter a performance for scent and chamber ensemble, going by the name of “Odophonics”, took place at the San Diego Art Institute, on May 14th, 2016 (Goeltzenleuchter, 2017). The auditory component of the performance involved Minimalist structures such as consonant harmony, drones and polyrhythms to create gradual chord transformations. All the notes in this ambient soundscape are represented in Piesse’s scale (see Figure 1). As the performers played, the corresponding scent notes were released in synchrony. Goeltzenleuchter describes how: “Together, the musical and olfactory harmonics gradually shift. Specific to the performance is the question: What relationships exist between concurrent perceptions of smell and sound?” According to the description, the performance can be considered as a jumping off point to explore Piesse’s *Odophone* to test new propositions about how one experiences smell, particularly in relation to sound. However, given that there is, as yet, no empirical evidence supporting the consensuality of the crossmodal correspondences outlined in Piesse’s *Gamut of Odors* (see Figure 3), one can only imagine how much more powerful such scented musical performances might one day be were the mapping between scent and sound to be based on commonly-shared crossmodal associations (e.g., as documented in the work of Belkin et al., 1997; Crisinel & Spence, 2012), rather than an idiosyncratic mapping such as that of Piesse, or worse still, a synaesthete’s inducer-concurrent relations.

Akin to Kandinsky’s (1977) notion of the inner harmony of color and form, Kenneth (1923, p. 77) has similarly referred to “indirect associations” between smells and music that result from “the affect produced by smell (being) similar to the affect produced by some other stimulus.” Interestingly, however, when Velasco et al. (2014) assessed the impact of playing consonant or dissonant music (versus playing white noise or silence) on participants’ ratings of fragrances, there was no effect of consonance or dissonance (contrary to Wang and Spence’s, 2016, findings with taste stimuli mentioned earlier). Intriguingly, only the white noise exerted any impact on olfactory ratings in Velasco et al.’s study.

Approaching a rather different problem, Cavazos Quero et al. (2021) attempted to harmonize sound and scent in order to help convey/represent a third modality, namely vision. In particular, these researchers proposed a prototype sensory substitution device for the exploration of the color content of visual art. The device implemented a multisensory color code that was designed to combine sounds and scents harmonically in order to represent (or substitute for) an absence of color vision. In particular, the code decomposed a specific color into a hue and a set of color dimensions (saturated, light, and dark) for each hue. The VIVALDI (see Cho et al., 2020) color code was used to facilitate hue identification (for example, red and orange are represented by string instruments; yellow and green by brass instruments; blue and purple by percussion instruments). To express the saturated, light, and dark color dimensions, VIVALDI uses a different set of pitches for each dimension and fragments of Vivaldi’s *Four Seasons* (Spring, Autumn, and Summer,

respectively). Regarding the pitch, the saturated dimension is represented by an A major chord (medium pitch), the light dimension using F major (high-pitch), and dark using E minor (low-pitch).

Besides sound, the proposed multisensory color code simultaneously integrates an olfactory component in order to make the association with colors easier and stronger. In particular, scents were used to express the saturated, lightness (light-dark), and temperature (warm-cool) color dimensions. In order to match the color dimension to scent, the authors performed a semantic differential survey in which the participants smelled different scents and chose which semantic adjectives relate most to that particular scent (cf. Dalton et al., 2008). Using those adjectives, scents were then matched with the color dimension (cf. Spence, 2020e). The system was tested on a group of 18 participants, allowing the researchers to evaluate the efficiency of correct color identification based on the multisensory approach. The preliminary results of this study suggested that the multisensory-based prototype improved people's confidence in exploring the color content of visual artworks, thus potentially making it suitable for those visually impaired individuals wanting to experience colorful artworks.

Olfactory-Visual (Color) Harmony

It is interesting that none of the many researchers working on the matching of colors with odors ever appear to use the term harmony, instead talking in terms of congruent or crossmodally corresponding color-odor mappings instead (see Spence, 2020e, for a recent review). The assumption would seem to be that the senses are so different, phenomenologically-speaking, that intramodal harmony (e.g., between colors or component scents) would always trump any kind of crossmodal harmony. At the same time, however, if one adopted the affective alignment account of harmony (see Kimura et al., 2005, 2012; and Kandinsky, 1977, on the notion of inner harmony) then there would appear to be no reason not to talk of the harmony of a particular combination of visual and olfactory stimuli. The affective alignment account of harmony refers to the idea that combinations of stimuli may be rated as harmonious just so long as the component stimuli are matched in terms of their affective valence—this possibly what Kandinsky had in mind. Similarly, definitions of harmony in terms of a pleasant combination of stimuli (Burchett, 1991, 2002; Judd & Wyszecki, 1975), at least as adopted by certain of those researchers working on color harmony, would also appear to allow for the possibility of olfactory-visual harmony. As such, the fact that researchers would appear to have been reticent to use the term harmony in this way (i.e., to describe well-matched combinations of olfactory and visual stimuli) is perhaps salient.

On the Multiple Meanings of Harmony

Returning, then, to the key question that was raised at the start of this review, to what extent should all this talk of harmony beyond hearing be taken literally versus perhaps being interpreted more metaphorically (cf. Lutz, 2007; Wagner et al., 1981; Walker, 1987). When, for example, the wine expert describes a particular wine-music match as perfectly harmonious do they actually mean anything more than that the different elements of the flavor appear to pair particularly well together, or that there is a certain natural correspondence, match, or affinity between the flavors that have been deliberately combined (in a food and beverage pairing, say). And can such claims support the claim that perceptual similarity has some meaning when comparing stimuli across the senses (see von Helmholtz, 1878/1971, p. 77, for the contrary position). Much the same might be said of Stratton's early paper on the spatial harmony of vision and touch (Stratton, 1899). After all, congruent crossmodal combinations of stimuli are likely to be processed more fluently, and hence will be associated with a valence that is affectively more positive. Harmonious (combinations of) stimuli are also likely to be processed more rapidly (e.g., Kimura et al., 2012). Of course, that being said, there are many different pairing principles that one can think of that might be expected to lead to enhanced

“processing fluency” (Reber, 2012; Reber et al., 2004; Reber et al., 1998) without all of them necessarily being relevant to a discussion of consonance or harmony. For instance, consider here only how both semantic and/or crossmodal congruency give rise to increased processing fluency, without the component stimuli necessarily harmonizing (see Chen & Spence, 2017). It makes little sense (at least not to your authors) to think of the sound of a barking dog, and the outline image of a dog as harmonizing any more than the bark being paired with the picture of a cat, despite the former undoubtedly being more semantically congruent than the latter. Certainly, in the literature, one only ever appears to find people talking about semantic matching or congruency and never about semantic harmony.

On the basis of the literature reviewed here, different concepts of harmony would appear to emerge depending on the context in which the term is used. In the unisensory context, harmony is essentially conceived of as an organizational principle to order different stimuli within the same modality, but not necessarily related to any particular processing advantage. For example, musical harmony provides one way of organizing sound materials attributing different roles depending on their functions within musical language. In the multisensory context, harmony is conceived of rather as a mapping criterion allowing one to bridge sensory stimuli that pertain to different sense modalities. In this context, it is often assumed that harmoniously-combined stimuli will also be more easily and effectively processed. But, as has just been mentioned, not all combinations of fluently processed stimuli will necessarily be described as harmonious (this was the suggestion mentioned a moment ago regarding pairs of stimuli that are semantically congruent).

One open question here concerns the link between crossmodal correspondence and crossmodal harmony. To the extent that certain correspondences, both intramodal and crossmodal, are based on the component stimuli having the same affective, or connotative, meaning, then many authors would appear to answer in the affirmative. At least, that would seem to be the claim if one takes the view that affective consonance, that is aligning the affective, or connotative meaning, of pairs of stimuli (as has been mentioned already, this may be what Kandinsky, 1977, had in mind when he referred to the mysterious notion of “inner harmony”) is rightly considered as a kind of harmony (though presumably of an affective, rather than necessarily a perceptual, kind). At the same time, however, statistical correspondences, based on the co-occurrence of sensory features in the environment (Stratton, 1899), may presumably lead to a feeling that the component stimuli belong together without necessitating that they are perceived as being an especially harmonious combination (Chen & Spence, 2017). In this regard, it is interesting to note the absences in researchers’/ artists’ use of the term “harmony”. Here, for example, one might point to the very noticeable absence of the term when describing the combination of vision (color) and scent, or when different vibrotactile frequencies are combined.

Therefore, in order to know whether we should take any discussion of consonance/harmony literally or merely metaphorically when we move beyond the auditory modality (e.g., to any of the crossmodal cases discussed in the section on Crossmodal harmony) where these notions originated, we might need to resort to a consideration of the phenomenology associated with claims of perceived harmony. However, this large and complicated topic is undoubtedly best left for another occasion.

Conclusions

As this historical review of the literature has hopefully helped to make clear, thinkers, philosophers, artists, theorists, and experimental psychologists alike have long been interested in “harmony”. In the recent past, the use of the term in the context of auditory perception would appear to describe fused, coherent, and unitary percepts that, as it so happens, are also processed fluently and thus rated as pleasant. By contrast, the majority of the visual literature tends to convey the sense of a pleasing relationship of elements (such as between neighboring color patches, or between colors

and shapes). As such, the relational nature of elements is more a part of the perceptual experience in vision than it is in audition (Palmer & Griscom, 2013), where the component parts typically fuse or unite. Work on the combination of fragrances and flavors typically uses the term harmony to refer to fused, or united, percepts while at the same time talking of the balanced composition of a mixture of elements (Smith, 2010, 2013; Spence & Guilford, 1933). Another important strand of the discussion of harmony has been in terms of the shared affective meaning of sensory stimuli (Kimura et al., 2005, 2012). This is presumably what Kandinsky (1977) refers to as “inner harmony”. At the same time, however, linking the notion of crossmodal harmony to synaesthesia, as was so often done during the romantic era (Argüelles, 1972; Kandinsky, 1977; Marks, 1978; Roinard et al., 1976), feels like a misleading approach.

In the contemporary era, the term “harmony” would appear to be increasingly frequently invoked to describe an especially pleasing combination of stimuli presented simultaneously to different pairs of sensory modalities. At the same time, however, it is important to note that any talk of crossmodal harmony also raises intriguing questions about the possibility of perceiving stimuli that are presented to different sensory modalities as being similar to one another or not. Furthermore, the very possibility of crossmodal perceptual similarity has not been accepted by all authors (see von Helmholtz, 1878/1971; though see also Hartshorne, 1934). Furthermore, the structural/phenomenal similarities between the Gestalt grouping principles operating in different senses might provide another grounds on which similarity judgments might be based. However, the topic of perceiving similarity across the senses in terms of structural qualities (such as rhythm, tempo, or meter) is one that we hope to address elsewhere. As will the question of whether stimuli that are perceptually, or structurally, similar are necessarily perceived as harmonious when combined will be tackled in a subsequent article (see Spence & Di Stefano, in preparation).

Hence, while the term “harmony” itself is increasingly often to be found outside the context of audition, it can sometimes be difficult to know whether its usage should be taken literally, or else meant merely metaphorically, as a combination of stimuli that appears especially well-matched and hence that can be processed fluently, regardless of whether the component stimuli can be individuated (and hence their relation assessed by an observer). It is perhaps also worth noting here that the seemingly increasing frequency with which the term is invoked across a diverse range of areas, and its seemingly easy interpretation by those who come across the term being used outside of a purely auditory context, would seem to argue that the term is meaningful. Harmony, in other words, appears to be a concept that resonates (if you’ll excuse the pun) well beyond the confines of the auditory modality, or rather hearing, where it was first introduced in a sensory context.

At the same time, however, answering the question of whether it is possible to experience crossmodal harmony really does depend on which of the several/many meanings of harmony that one finds in the unisensory literature, one uses (see Table 1). If one adopts the suggestion that harmony is nothing more than merely a pleasant combination of stimuli (e.g., Burchett, 1991, 2002; Judd & Wysecki, 1975) then the answer would appear to be uncontroversially in the affirmative. However, it is important to stress that here such a definition has not been widely used outside the literature on color harmony, and even there it is controversial (see Palmer & Griscom, 2013; see Table 2). If instead, one uses the term harmony to refer specifically to the goodness of the relationship between parts, then one still leaves open the question of whether it is perceptual similarity (if that is even possible; see von Helmholtz, 1878/1971), structural similarity, or affective (or inner) harmony/matching that is key (see Table 2).

Future Research Questions in the Study of Harmony Beyond Audition

In the future, it will be interesting to determine whether crossmodal harmony can be experienced between any pair of sensory modalities, or whether instead it is preferentially experienced only

Table 1. Various Meanings that have been Attached to the Term “Harmony”, as a Function of the Sense(s) being Discussed.

Audition	<ul style="list-style-type: none"> • The study of simultaneous sounds whose sonority may be agreeable (i.e., consonant) or disagreeable (i.e., dissonant) to the ear (e.g., Bidelman & Krishnan, 2011). • Pleasant combinations of auditory stimuli that are perceived as fused, unitary, and stable (von Helmholtz, 1954; Stumpf, 1883/1890). • Stimuli (sounds) that go well together and can be processed fluently/rapidly (e.g., Tabas et al., 2019).
Vision	<ul style="list-style-type: none"> • As the pleasing proportion amongst different parts (Vitruvius, 1999). • Pleasing combination of neighboring stimuli (Burchett, 1991, 2002; Judd & Wyszecki, 1975). • Good Gestalt—good fit, good form, or Gestalt prägnanz (Griscom & Palmer, 2011); A stimulus that harmonizes especially well with many other stimuli of its type (Palmer & Griscom, 2013). • Palmer and Griscom (2013) argue that harmony (a relational property) should be distinguished from preference. • Stimulus attributes associated with similar affective meaning (Kimura et al., 2005, 2012; Oyama, 2003), sometimes referred to as “inner harmony”; see Kandinsky, 1977).
Olfaction	<ul style="list-style-type: none"> • Fused or unified combination of odorants (Spence & Guilford, 1933). • Combination of stimuli that give rise to a balanced relationship between the component parts (Ackerman, 1990).
Flavor	<ul style="list-style-type: none"> • A balanced or unified combination of elements/flavors (Smith, 2010, 2013). • How well sensations go together based on aromatic similarity (Eschevins et al., 2018, 2019).
Vision-touch	<ul style="list-style-type: none"> • Statistical crossmodal correspondence of amodal features (e.g., location), learned associatively (Stratton, 1899).
Music-visual art	<ul style="list-style-type: none"> • “Harmonious-disharmonious” as but one of the semantic differential scales used by Whiteford et al. (2018).
Music-scent	<ul style="list-style-type: none"> • Unification of sensations at the “pitch of harmony” (Holt-Hansen, 1968, 1976; though see Rudmin and Capelli, 1983). • Similar in terms of crossmodal correspondences (Spence, 2021).
Music-flavor	<ul style="list-style-type: none"> • Similar in terms of crossmodal correspondences (Spence et al., 2021; cf. Wang et al., 2016).

between certain combinations of senses—here it is already noticeable how audiovisual harmony (as in the case of color music), and music-flavor/olfactory harmony (as in the emerging field of research on sonic seasoning) appear to be the most frequently studied sensory combinations, while attempts to create harmonious combinations of auditory and vibrotactile stimuli, or visual and olfactory

Table 2. Summary of Uses of the Term Harmony in Various Sensory Domains (and the relevant section of the main text where the topic is discussed). A “Y” indicates that the definition of harmony has been used in the literature, whereas “?” indicates that no specific claims have been made in support of such a suggestion.

	Pleasurable	Going well together	Processing fluently
Audition (Harmonic sounds)	Y	Y	Y
Vision (Early history of visual harmony—Alternate definitions of harmony)	Y	Y	Y
Vibrotactile (Vibrotactile harmony)			
Olfaction (Olfactory harmony)	Y	Y	?
Flavor (Harmonious tastes/flavors)	Y	Y	?
Vision-touch (Introduction)	?	Y	?
Musical-visual art (Audiovisual harmony)	?	?	?
Music-flavor (Harmonizing music and taste/flavor)	Y	Y	?
Music-scent (Harmonizing scent and music)	?	Y	?
Olfactory-visual (Olfactory-visual (color) harmony)	?	?	?

stimuli, are both noticeable by their absence (see Spence, 2020e, 2021, for reviews). Thereafter, it will be intriguing to consider whether there is any additional value, or meaning, to a consideration of multisensory as opposed to merely crossmodal harmony (see Spence, 2015, for a review of crossmodal Gestalt phenomena).²⁵ While analogous grouping principles may well be experienced in different sensory modalities, there is far more uncertainty over the question of whether it is even possible to experience intersensory, or multisensory, Gestalten (Battey & Fischman, 2016; Zilcher, 2016; see Spence, 2015, for a review). As we have just seen, the answer to this question will likely depend upon exactly what definition of harmony one is using. Finally, it will be interesting in future research to find out whether the individual/cross-cultural differences that have been documented to exist in the case of auditory and visual harmonic preferences (Palmer & Griscorn, 2013) also extend to the crossmodal, or multisensory, case (cf. Brown et al., 2011; Eysenck, 1940).

Acknowledgements

This research was supported by the Rethinking the Senses grant from the AHRC(UK) awarded to Charles Spence (AH/L007053/1).

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Arts and Humanities Research Council (grant number AH/L007053/1).

ORCID iD

Charles Spence  <https://orcid.org/0000-0003-2111-072X>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. While consonance and dissonance are fundamental concepts when discussing harmony and melody, the terms have not been systematically applied to rhythm yet, and such discussion is primarily limited to a few musicologists. The terms “rhythmic consonance” and “rhythmic dissonance”, for example, appear in Schillinger’s *System of Musical Composition* (1946) where he refers to such combinations as “two against three,” “three against four”. A few years later, Sachs (1953) and Cooper and Meyer (1960) also used the terms consonance and dissonance with respect to meter, but only in passing.
2. As Hui (2013, pp. 65–66) notes “Helmholtz clearly appreciated not just the emotive aspects but also the intellectual aspects of harmony and melody.” Helmholtz also considered that Mozart’s great works contained “profoundly musical harmonies.”
3. Intriguingly, according to Köhler’s (1929) Minimum Principle Gestalt rule of Prägnanz, the mind organizes the world in a way that minimizes the expenditure of energy (Aksentijević et al., 2001). The latter researchers suggest a common (possibly central) mechanism guiding the perceptual grouping seen in vision and audition. They further talk of the ‘supramodal invariants of grouping’.
4. At the same time, however, it is perhaps also worth noting that auditory psychophysicists have highlighted a number of other perceptual consequences of varying the harmonic relatedness of pairs of auditory stimuli (e.g., see Plomp & Levelt, 1965; Tillmann & Bharucha, 2002), though these will not be discussed further here.
5. For, as Schloss and Palmer (2011, p. 552) note: “Because preference and harmony are so clearly different concepts in music perception, we are sceptical of claims that they are the same concept in color perception.”
6. It is therefore rather unclear to what extent the odours would have been mixed prior to hitting the olfactory receptors.
7. Of course, it is possible that the phenomenology of olfactory mixing is simply unlike that in the other senses (cf. Cooke & Myin, 2011).
8. At one point, he writes that: “It is because of the elementary stage reached by our painting that we are so little able to grasp the inner harmony of true colour and form composition.” (Kandinsky, 1977, p. 47). Note here also that Kandinsky’s book was first published in English in 1914 under the title *The art of spiritual harmony*.
9. Palmer and Griscorn (2013) argue, though ultimately unconvincingly, to your present authors, that the harmony of discrete stimuli, such as pale colour patches can also be meaningfully described as harmonious in the sense that it may be rated as more harmonious with the majority of other stimuli within the category.
10. The relational aspect is a hallmark of Leibniz’s characterizations of harmony as “similitude in dissimilar things”, “diversity compensated by identity” or as “unity in variety” (see Carlin, 2000, for a review).
11. One of the most important applications of such research may turn out to be in terms of the development of more intuitive sensory substitution solutions for those who are deaf (Alves Araujo et al., 2017; cf. Goldstein Jr. & Proctor, 1985). Another potential application, as highlighted by Yoo et al. (2014), is in terms of designing vibrotactile warning signals that can vary in terms of their perceived urgency.
12. That said, one dissimilarity is that certain successful perfumes are made up of just a single molecule (see Spence & Wang, 2018). There is presumably no equivalent in the case of popular music.
13. The term “harmony” appearing 72 times in Eschevins et al.’s (2018) article.
14. Incidentally, crossmodal correspondences research in taste-pitch matching, the participants have indeed been shown to match umami to a lower pitch range (Crisinel & Spence, 2010; Wang et al., 2016).
15. While it is sometimes suggested that the relationship between the inducers and concurrents experienced by synaesthetes are typically experienced as pleasant by synaesthetes and non-synaesthetes alike (Ward et al., 2008), this is by no means always the case.
16. According to scientific thinkers such as Kepler, the 5:4 ratio represented: “*a mathematical ideal of a divine substance, that need not even be expressed in order to exist, for it is eternal*” (quoted in James, 2012)
17. The Good Friday performance of the *Misere Mei* was also accompanied by the symbolic extinguishing of all the candles in the chapel, one by one, on, symbolizing the Last Supper, and eventually plunging the chapel into complete darkness at the point in the music symbolizing Christ’s death (Grovier, 2012).
18. Iannis Xenakis, the architect who worked alongside Le Corbusier, once highlighted this issue: “... Xenakis was not really interested in musical “translations” of architecture or vice versa. Indeed, as he stated: “we are capable of speaking two languages at the same time. One is addressed to the eyes, the other to the ears.” According to such a view, addressing the same message to the different senses runs the danger of pleonasm.

- “The opposite idea, a total dissociation between visual and aural perception, would become the key concept in the later Polytopes, large abstract sound and light installations in which Xenakis delivered his own interpretation of the total work of art.” (Sterken, 2007, pp. 32–33).
19. Arnheim (1986, p. 207) put it thus: “There are ... perceptually convincing correspondences between colors and sounds based on shared expressive qualities.”
 20. Galeyev and Vanechkina (2001) talk of the coloured visualization of Scriabin’s Luce and the tonal harmonic plan behind it.
 21. Wells (1980, p. 101) also writes that: “The author tells of events leading up to finding a possible link between musical harmony and visual color based on complementarity.” And that “The parallelism of musical harmony with visual color departs radically from current hypotheses of tonal organization.” (Wells, 1980, p. 105).
 22. The analogy between music and cooking was also highlighted in a North Indian context by McNeil (1993–1994; see also Stone, 1982).
 23. Note that in an early study Bronner et al. (2012) had already suggested a crossmodal correspondence between musical consonance and sweet-smelling aromas such as vanilla.
 24. It is, however, unclear the extent to which the composer relied on his peculiar synaesthesia, versus the crossmodal correspondences common to us all in developing his soundscapes.
 25. This, what Gilbert (1938) once referred to as ‘inter-sensory Gestalten’.

References

- Abbado, A. (1988). Perceptual correspondences of abstract animation and synthetic sound. *Leonardo*, 1- (Electronic Art Supplemental Issue), 3–5. <https://doi.org/10.2307/1557901>
- Ackerman, D. (1990). *A natural history of the senses*. Random House.
- Akhtaruzzaman, M., & Shafie, A. A. (2011). Geometrical substantiation of Phi, the golden ratio and the baroque of nature, architecture, design and engineering. *International Journal of Arts*, 1(1), 1–22. <https://doi.org/10.5923/j.arts.20110101.01>
- Aksentijević, A., Elliott, M. A., & Barber, P. J. (2001). Dynamics of perceptual grouping: Similarities in the organization of visual and auditory groups. *Visual Cognition*, 8(3-5), 349–358. <https://doi.org/10.1080/13506280143000043>
- Albers, J. (1963/2013). *Interaction of color: 50th anniversary edition*. Yale University Press.
- Alberti, L. B. (1485). *De Re Aedificatoria, Firenze: Nicolò Di Lorenzo*. [English translation by James Leoni De Re Aedificatoria; Alberti’s ten books of architecture (1726) London, UK: Thomas Edlin].
- Allen, E. C., & Guilford, J. P. (1936). Factors determining the affective values of color combinations. *American Journal of Psychology*, 48(4), 643–648. <https://doi.org/10.2307/1416516>
- Alves, B. (2005). Digital harmony of sound and light. *Computer Music Journal*, 29(4), 45–54.
- Alves, B. (2012). Consonance and dissonance in visual music. *Organised Sound*, 17(2), 114–119. <https://doi.org/10.1017/S1355771812000039>
- Alves Araujo, F., Lima Brasil, F., Lima Santos, A. C., de Sousa Batista Junior, L., Fonseca Dutra, S. P., & Freire Batista, C. E. C. (2017). Auris system: Providing vibrotactile feedback for hearing impaired population. *BioMed Research International*, 2017, 2181380. <https://doi.org/10.1155/2017/2181380>
- Andermann, M., Patterson, R. D., & Rupp, A. (2020). Transient and sustained processing of musical consonance in auditory cortex and the effect of musicality. *Journal of Neurophysiology*, 123(4), 1320–1331. <https://doi.org/10.1152/jn.00876.2018>
- Argüelles, J. A. (1972). *Charles Henry and the formation of a psychophysical aesthetic*. University of Chicago Press.
- Aristotle (1908). *De sensu et sensibili [On sense and the sensible] (trans. J. I. Beare)*. Clarendon Press.
- Arnheim, R. (1986). *New essays on the psychology of art*. University of California Press.
- Athanasopoulos, G., Eerola, T., Lahdelma, I., & Kaliakatos-Papakostas, M. (2021). Harmonic organisation conveys both universal and culture-specific cues for emotional expression in music. *PLoS ONE*, 16(1), e0244964. <https://doi.org/10.1371/journal.pone.0244964>.
- Banes, S. (2001). Olfactory performances. *TDR/The Drama Review*, 45(1), 68–76. <https://doi.org/10.1162/105420401300079040>
- Bashwiner, D. (2013). Music analysis for multimedia: A perspective from music theory. In S. L. Tan, A. Cohen, S. Lipscomb, & R. Kendall (Eds.), *The psychology of music in multimedia* (pp. 89–117). Oxford University Press.

- Basquin, K. S. (2020). *Mary Ellen Bute: Pioneer animator*. John Libbey Publishing.
- Basten, F. E., Salvatore, R., & Kaufman, P. A. (1995). *Max Factor's Hollywood: Glamour, movies, make-up*. Los Angeles.
- Bathey, B., & Fischman, R. (2016). Convergence of time and space, In T. Kaduri (Ed.), *The Oxford handbook of sound and image in western art*. (pp. 61-82). Oxford, UK: Oxford University Press.
- Baudelaire, C. (1889). Curiosités Esthétiques, In *Oeuvres Complètes*. Calman Lévy.
- Beck, E. M. (1999). Marchetto da Padova and Giotto's Scrovegni Chapel frescos. *Early Music*, 27(1), 7–23. <https://doi.org/10.1093/em/27.1.7>
- Beck, E. M. (2005). *Giotto's harmony: Music and art in Padua at the crossroads of the renaissance*. European Press Academic Publishing.
- Belkin, K., Martin, R., Kemp, S. E., & Gilbert, A. N. (1997). Auditory pitch as a perceptual analogue to odor quality. *Psychological Science*, 8(4), 340–342. <https://doi.org/10.1111/j.1467-9280.1997.tb00450.x>
- Bidelman, G. M., & Heinz, M. G. (2011). Auditory-nerve responses predict pitch attributes related to musical consonance-dissonance for normal and impaired hearing. *Journal of the Acoustical Society of America*, 130(3), 1488–1502. <https://doi.org/10.1121/1.3605559>
- Bidelman, G. M., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience*, 29(42), 13165–13171. <https://doi.org/10.1523/JNEUROSCI.3900-09.2009>
- Bidelman, G. M., & Krishnan, A. (2011). Brainstem correlates of behavioral and compositional preferences of musical harmony. *Neuroreport*, 22(5), 212–216. <https://doi.org/10.1097/WNR.0b013e328344a689>
- Birren, F. (1967). *The contrast and harmony of colors*. New Hyde Park.
- Blacking, J. (1970). Tonal organization in the music of two Venda initiation schools. *Ethnomusicology*, 14(1), 1–56. <https://doi.org/10.2307/850292>
- Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, 2, 382–387. <https://doi.org/10.1038/7299>
- Bones, O., Hopkins, K., Krishnan, A., & Plack, C. J. (2014). Phase locked neural activity in the human brainstem predicts preference for musical consonance. *Neuropsychologia*, 58, 23–32. <https://doi.org/10.1016/j.neuropsychologia.2014.03.011>
- Bowling, D. L., & Purves, D. (2015). A biological rationale for musical consonance. *Proceedings of the National Academy of Sciences of the USA*, 112(36), 11155–11160. <https://doi.org/10.1073/pnas.1505768112>
- Brain, R. M. (2015). *The pulse of modernism: Physiological aesthetics in fin-du-siècle Europe*. University of Washington Press.
- Brattico, E., Pallesen, K. J., Varyagina, O., Bailey, C., Anourova, I., Järvenpää, M., Eerola, T., & Tervaniemi, M. (2009). Neural discrimination of nonprototypical chords in music experts and laymen: An MEG study. *Journal of Cognitive Neuroscience*, 21(11), 2230–2244. <https://doi.org/10.1162/jocn.2008.21144>
- Bronner, K., Bruhn, H., Hirt, R., & Piper, D. (2012). What is the sound of citrus? Research on the correspondences between the perception of sound and flavour. In Proceedings of the 12th international conference of music perception and cognition and the 8th triennial conference of the european society for the cognitive sciences of music, 23–28 July 2012, Thessaloniki, Greece. http://icmpc-escom2012.web.auth.gr/sites/default/files/papers/142_Proc.pdf.
- Brougher, K., Mattis, O., Strick, J., Wiseman, A., & Zilczer, J. (Eds.). (2005). *Visual music: Synaesthesia in art and music since 1900*. Thames and Hudson.
- Brown, S., Gao, X., Tisdelle, L., Eickhoff, S. B., & Liotti, M. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *NeuroImage*, 58(1), 250–258. <https://doi.org/10.1016/j.neuroimage.2011.06.012>
- Brown, S., & Jordania, J. (2013). Universals in the world's musics. *Psychology of Music*, 41(2), 229–248. <https://doi.org/10.1177/0305735611425896>
- Bullon, S. (1978). *Longman dictionary of the contemporary English*. Longman.
- Burchett, K. E. (1991). Color harmony attributes. *Color Research & Application*, 16(4), 275–278. <https://doi.org/10.1002/col.5080160410>

- Burchett, K. E. (2002). Color harmony. *Color Research & Application*, 27(1), 28–31. <https://doi.org/10.1002/col.10004>
- Burnham, R. W., Hanes, R. M., & Bartleson, J. C. (1963). *Color: A guide to basic facts and concepts*. Wiley.
- Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (pp. 215–264). Academic Press.
- Caivano, J. L. (1994). Color and sound: Physical and psychophysical relations. *Color Research and Application*, 19(2), 126–133. <https://doi.org/10.1111/j.1520-6378.1994.tb00072.x>
- Cardullo, R. J. (2017). Gesamtkunstwerk, synesthesia, and the Avant-garde: Wassily Kandinsky's the yellow sound as a work of art. *Hermeneia*, 18, 5–21.
- Carlin, L. (2000). On the very concept of harmony in Leibniz. *The Review of Metaphysics*, 54(1), 99–125.
- Caswell, A. B. (1980). The Pythagoreanism of Arcimboldo. *The Journal of Aesthetics and Art Criticism*, 39(2), 155–161. https://doi.org/10.1111/1540_6245.jaac39.2.0155
- Cavazos Quero, L., Lee, C.-H., & Cho, J.-D. (2021). Multi-sensory color code based on sound and scent for visual art appreciation. *Electronics*, 10, 1696. <https://doi.org/10.3390/electronics10141696>
- Cerretani, L., Biasini, G., Bonoli-Carbognin, M., & Bendini, A. (2007). Harmony of virgin olive oil and food pairing: A methodological proposal. *Journal of Sensory Studies*, 22(4), 403–416. <https://doi.org/10.1111/j.1745-459X.2007.00115.x>
- Chartier, F. (2012). *Taste buds and molecules: The art and science of food, wine, and flavor* (translated by Levi Reiss). John Wiley and Sons.
- Chartier, F., & Kitano, H. (2019). The aromatic theory of molecular harmonies and sommelry. In Paper presented at the science & cooking world congress, Barcelona, February 4–6.
- Chen, Y.-C., & Spence, C. (2017). Assessing the role of the “unity assumption” on multisensory integration: A review. *Frontiers in Psychology*, 8, 445. <https://doi.org/10.3389/fpsyg.2017.00445>
- Chester, N. (2017). I spoke to an “aroma jockey”, who DJs smells for deaf audiences. *Vice*, March 15th. <https://www.vice.com/en/article/9amnq7/i-spoke-to-an-aroma-jockey-who-djs-smells-for-deaf-audiences>.
- Chevreul, M. E. (1967). *The principles of harmony and contrast of colors* (C. Martel, Trans.). Reinhold. Original French edition published 1839.
- Chiandetti, C., & Vallortigara, G. (2011). Chicks like consonant music. *Psychological Science*, 22(10), 1270–1273. <https://doi.org/10.1177/0956797611418244>
- Cho, J. D., Jeong, J., Kim, J. H., & Lee, H. (2020). Sound coding color to improve artwork appreciation by people with visual impairments. *Electronics*, 9(11):1981. <https://doi.org/10.3390/electronics9111981>
- Cichelli, A., Cerretani, L., Di Lecce, G., & Piochi, M. (2020). Exploring harmony in extra virgin olive oils and vegetables pairings. *Grasas y Aceites*, 71(2):353. <https://doi.org/10.3989/gya.0117191>
- Collier, G. L. (1996). Affective synesthesia: Extracting emotion space from simple perceptual stimuli. *Motivation and Emotion*, 20(1), 1–32. <https://doi.org/10.1007/BF02251005>
- Comanini, G. (1590). *The Figino, or on the purpose of painting: Art theory in the late renaissance*, [English translation by A. Doyle-Anderson and G. Maiorino, with Introduction and notes, (2016), Toronto: University of Toronto Press, 2016. <https://doi.org/10.3138/9781442681279>
- Conard, N. J., Malina, M., & Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, 460(7256), 737–740. <https://doi.org/10.1038/nature08169>
- Cooke, E., & Myin, E. (2011). Is trilled smell possible? How the structure of olfaction determines the phenomenology of smell. *Journal of Consciousness Studies*, 18(11–12), 59–95.
- Cooper, G., & Meyer, L. (1960). *The rhythmic structure of music*. University of Chicago Press.
- Coucquyt, P., Lahousse, B., & Langenbick, J. (2020). *The art and science of Foodpairing: 10,000 flavour matches that will transform the way you eat*. Mitchell Beazley.
- Cousineau, M., McDermott, J. H., & Peretz, I. (2012). The basis of musical consonance as revealed by congenital amusia. *Proceedings of the National Academy of Sciences of the USA*, 109(48), 19858–19863. <https://doi.org/10.1073/pnas.1207989109>
- Crespo-Bojorque, P., Monte-Ordoño, J., & Toro, J. M. (2018). Early neural responses underlie advantages for consonance over dissonance. *Neuropsychologia*, 117, 188–198. <https://doi.org/10.1016/j.neuropsychologia.2018.06.005>

- 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(7), 1994–2002. <https://doi.org/10.3758/APP.72.7.1994>
- Crisinel, A.-S., & Spence, C. (2012). A fruity note: Crossmodal associations between odors and musical notes. *Chemical Senses*, 37(2), 151–158. <https://doi.org/10.1093/chemse/bjr085>
- Csibra, G., Davis, G., Spratling, M. W., & Johnson, M. H. (2000). Gamma oscillations and object processing in the infant brain. *Science*, 290(5496), 1582–1585. <https://doi.org/10.1126/science.290.5496.1582>
- Dalton, P., Maute, C., Oshida, A., Hikichi, S., & Izumi, Y. (2008). The use of semantic differential scaling to define the multi-dimensional representation of odors. *Journal of Sensory Studies*, 23(4), 485–497. <https://doi.org/10.1111/j.1745-459X.2008.00167.x>
- da Vinci, L. (1802). *A treatise on painting*. London, UK: Taylor.
- Di Dio, C., Macaluso, E., & Rizzolatti, G. (2007). The golden beauty: Brain response to classical and renaissance sculptures. *PLoS ONE*, 2(11), e1201. <https://doi.org/10.1371/journal.pone.0001201>.
- Di Stefano, N., Murari, M., & Spence, C. (2022). Crossmodal correspondences in art and science: Odours, poetry, and music. In N. Di Stefano & M. T. Russo (Eds.), *Olfaction: An interdisciplinary perspective from philosophy to life sciences, human perspectives in health sciences and technology* 4 (pp. 155–189). Springer. https://doi.org/10.1007/978-3-030-75205-7_10
- Eschevins, A., Giboreau, A., Allard, T., & Dacremont, C. (2018). The role of aromatic similarity in food and beverage pairing. *Food Quality and Preference*, 65, 18–27. <https://doi.org/10.1016/j.foodqual.2017.12.005>
- Eschevins, A., Giboreau, A., Julien, P., & Dacremont, C. (2019). From expert knowledge and sensory science to a general model of food and beverage pairing with wine and beer. *International Journal of Gastronomy & Food Science*, 17, 100144. <https://doi.org/10.1016/j.ijgfs.2019.100144>
- Euler, L. (1739). *Tentamen novae theoriae musicae ex certissimis harmoniae principiis dilucide expositae* [A new music theory clearly exposed on the basis of the most reliable principles of harmony]. St. Petersburg Academy.
- Eysenck, H. J. (1940). The general factor in aesthetic judgements. *British Journal of Psychology*, 31(1), 94–102. <https://doi.org/10.1111/j.2044-8295.1940.tb00977.x>
- Field, G. (1835). *Chromatics: Or the analogy, harmony, and philosophy of colours*. David Bogue.
- Field, G. (1845). *Chromatography; or, a treatise on colours and pigments, and of their powers in painting*, London, UK: Winsor and Newton.
- Fishman, Y. I., Volkov, I. O., Noh, M. D., Garell, P. C., Bakken, H., Arezzo, J. C., Howard, M. A., & Steinschneider, M. (2001). Consonance and dissonance of musical chords: Neural correlates in auditory cortex of monkeys and humans. *Journal of Neurophysiology*, 86, 2761–2788. <https://doi.org/10.1152/jn.2001.86.6.2761>
- Fleischer, M. (2012). Incense & decadents: Symbolist theatre's use of scent. In S. Banes & A. Lepecki (Eds.), *The senses in performance* (pp. 105–114). Routledge.
- Foss, A. H., Altschuler, E. L., & James, K. H. (2007). Neural correlates of the Pythagorean ratio rules. *Neuroreport*, 18, 1521–1525. <https://doi.org/10.1097/WNR.0b013e3282ef6b51>
- Galeyev, B. M., & Vanechkina, I. L. (2001). Was Scriabin a synesthete? *Leonardo*, 34(4), 357–361. <https://doi.org/10.1162/00240940152549357>
- Garner, W. R. (1974). *The processing of information and structure*. Lawrence Erlbaum Associates.
- Gawboy, A. M., & Townsend, J. (2012). Scriabin and the possible. *MTO*, 18(2), 1–21. <https://doi.org/10.30535/mto.18.2.2>
- Gilbert, G. M. (1938). A study in inter-sensory Gestalten. *Psychological Bulletin*, 35(1), 698. <https://doi.org/10.1037/h0061128>
- Gill, K. Z., & Purves, D. (2009). A biological rationale for musical scales. *PLoS One*, 4(12), e8144. <https://doi.org/10.1371/journal.pone.0008144>
- Goeltzenleuchter, B. (2017). Odophonics: A performance for scent and chamber ensemble. <https://www.bgprojects.com/home/2017/9/4/odophonics>.
- Goldstein, M. H.Jr., & Proctor, A. (1985). Tactile aids for profoundly deaf children. *Journal of the Acoustical Society of America*, 77(1), 258–265. <https://doi.org/10.1121/1.392267>
- Granger, G. W. (1953). Area balance in color harmony: An experimental study. *Science* 117, 59–61. <https://doi.org/10.1126/science.117.3029.59>

- Granger, G. W. (1955). An experimental study of colour harmony. *The Journal of General Psychology*, 52(1), 21–35. <https://doi.org/10.1080/00221309.1955.9918341>
- Griscom, W. S., & Palmer, S. E. (2011). Cross-modal relations between emotional content and preference for harmony. *Journal of Vision*, 11(11), 390. <https://doi.org/10.1167/11.11.390>
- Grovier, K. (2012). Secret harmony: Mozart, Michelangelo, and the *Miserere Mei*. *The Times Literary Supplement*, June 8th.
- Gurwitsch, A. (1979). *Phenomenology and theory of science*. Northwestern University Press.
- Halperin, J. U. (1988). *Félix Fénéon: Aesthete and anarchist in fin-de-siècle Paris*. Yale University Press.
- Handel, S. (1988a). No one analogy is sufficient: Rejoinder to Kubovy. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 321. <https://doi.org/10.1037/h0092865>
- Handel, S. (1988b). Space is to time as vision is to audition: Seductive but misleading. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 315–317. <https://doi.org/10.1037/0096-1523.14.2.315>
- Handel, S., & Buffardi, L. (1968). Pattern perception: Integrating information presented in two modalities. *Science (New York, N.Y.)*, 162, 1026–1028. <https://doi.org/10.1126/science.162.3857.1026>
- Handel, S., & Buffardi, L. (1969). Using several modalities to perceive one temporal pattern. *Quarterly Journal of Experimental Psychology*, 21(3), 256–266. <https://doi.org/10.1080/14640746908400220>
- Hartshorne, C. (1934). *The philosophy and psychology of sensation*. University of Chicago Press.
- Hay, D. R. (1836). *The laws of harmonious colouring*. William and Robert Chambers.
- Hogg, J. (1969). The prediction of semantic differential ratings of color combinations. *The Journal of General Psychology*, 80(1), 141–152. <https://doi.org/10.1080/00221309.1969.9711280>
- Holt-Hansen, K. (1968). Taste and pitch. *Perceptual and Motor Skills*, 27(1), 59–68. <https://doi.org/10.2466/pms.1968.27.1.59>
- Holt-Hansen, K. (1976). Extraordinary experiences during cross-modal perception. *Perceptual and Motor Skills*, 43(3_Suppl), 1023–1027. <https://doi.org/10.2466/pms.1976.43.3f.1023>
- Huang, J., Gamble, D., Sarnlertsophon, K., Wang, X., & Hsiao, S. (2012). Feeling music: Integration of auditory and tactile inputs in musical meter perception. *PLoS ONE*, 7(10), e48496. <https://doi.org/10.1371/journal.pone.0048496>
- Hui, A. (2013). *The psychophysical ear: Musical experiments, experimental sounds, 1840–1910*. MIT Press.
- Itoh, K., Suwazono, S., & Nakada, T. (2003). Cortical processing of musical consonance: An evoked potential study. *Neuroreport*, 14(18), 2303–2306. <https://doi.org/10.1097/00001756-200312190-00003>
- Itten, J. (1970). *The elements of color (Ernst van Haagen, Trans.)*. Wiley.
- Izumi, A. (2000). Japanese monkeys perceive sensory consonance of chords. *Journal of the Acoustical Society of America*, 108(6), 3073–3078. <https://doi.org/10.1121/1.1323461>
- Jacobsen, T. (2002). Kandinsky's questionnaire revisited: Fundamental correspondence of basic colors and forms? *Perceptual & Motor Skills*, 95(3), 93–913. <https://doi.org/10.2466/pms.2002.95.3.903>
- Jacobsen, T. (2004). Kandinsky's color-form correspondence and the Bauhaus colors: An empirical view. *Leonardo*, 37(2), 134–136. <https://doi.org/10.1162/0024094041139193>
- Jacobson, E., Granville, W. C., & Foss, C. E. (1948). *Color harmony manual*. Container Corporation of America.
- James, J. (2012). *The music of the spheres: Music, science, and the natural order of the universe*. Springer.
- Johnson, H. (2005). *Wine: A life uncorked*. Weidenfeld & Nicolson.
- Judd, D. B., & Wyszecki, G. (1975). *Color in business, science and industry* (3rd ed.). Wiley.
- Julesz, B., & Hirsh, I. J. (1972). Visual and auditory perception—An essay of comparison. In E. E. David Jr. & P. B. Denes (Eds.), *Human communication: A unified view* (pp. 283–340). McGraw-Hill.
- Kadia, S. C., & Wang, X. (2003). Spectral integration in A1 of awake primates: Neurons with single- and multi-peaked tuning characteristics. *Journal of Neurophysiology*, 89(3), 1603–1622. <https://doi.org/10.1152/jn.00271.2001>
- Kaiser, J., Bühler, M., & Lutzenberger, W. (2004). Magnetoencephalographic gamma-band responses to illusory triangles in humans. *NeuroImage*, 23(2), 551–560. <https://doi.org/10.1016/j.neuroimage.2004.06.033>
- Kandinsky, W. (1914). *The art of spiritual harmony*. Constable and Co.
- Kandinsky, W. (1925). *Point and line to plane (transl. H. Dearstyne & H. Rebay, 1979)*. Dover Publications.
- Kandinsky, W. (1977). *Concerning the spiritual in art, especially in painting (trans. M. T. H. Sadler)*. Dover Publications.

- Kenneth, J. H. (1923). Mental reactions to smell stimuli. *Psychological Review*, 30(1), 77–79. <https://doi.org/10.1037/h0068405>
- Kepler, J. (1619). *Harmonices Mundi Libri V*. Johann Planck.
- Kiltinavičiūtė, A. (2020). Sensation (un)bound: Literary synesthesia and cross-sensory perception in Dante's "Purgatorio" 24. *Bibliotheca Dantesca: Journal of Dante Studies*, 3, 4. <https://repository.upenn.edu/bibdant/vol3/iss1/4>.
- Kim, S., & Lecat, B. (2017). An exploratory study to develop Korean food and wine pairing criteria. *Foods (Basel, Switzerland)*, 3(3), 40. <https://doi.org/10.3390/beverages3030040>
- Kimchi, R., & Razpurker-Apfeld, I. (2004). Perceptual grouping and attention: Not all groupings are equal. *Psychonomic Bulletin & Review*, 11(4), 687–696. <https://doi.org/10.3758/BF03196621>
- Kimura, A., Dan, I., Watanabe, N., Yamada, H., & Wada, Y. (2012). Reaction time as an indicator of stimulus-response binding in affective judgment of visual stimuli. *Japanese Psychological Research*, 54(4), 335–347. <https://doi.org/10.1111/j.1468-5884.2012.00511.x>
- Kimura, A., Wada, Y., & Noguchi, K. (2005). The effect of the similarity of affective meanings on harmony between shape and color. *Bulletin of Japanese Society for the Science of Design*, 52, 1–8.
- Knapton, S. (2015). Why sparkling wine sounds like beans falling on a plastic tray. *The Daily Telegraph*, May 2nd, 7. <http://www.telegraph.co.uk/news/science/science-news/11577381/What-does-wine-sound-like-Youre-about-to-find-out.html>.
- Koffka, K. (1935). *Principles of Gestalt psychology*. Harcourt, Brace, & World.
- Köhler, W. (1929). *Gestalt psychology*. Liveright.
- Krebs, H. (1987). Some extensions of the concepts of metrical consonance and dissonance. *Journal of Music Theory*, 31(1), 99–120. <https://doi.org/10.2307/843547>
- Kubovy, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 55–98). Erlbaum.
- Kubovy, M. (1988). Should we resist the seductiveness of the space: Time::vision:Audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, 14(2), 318–320. <https://doi.org/10.1037/0096-1523.14.2.318>
- Kubovy, M., & Pomerantz, J. J. (Eds.). (1981). *Perceptual organization*. Erlbaum.
- Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, 80(1-2), 97–126. [https://doi.org/10.1016/S0010-0277\(00\)00155-4](https://doi.org/10.1016/S0010-0277(00)00155-4)
- Laudamiel, C. (2008). Haute perfumery and haute cuisine. In H. Blumenthal (Ed.), *The big fat duck cookbook* (pp. 478–482). Bloomsbury.
- Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *Journal of Neuroscience*, 29(18), 5832–5840. <https://doi.org/10.1523/JNEUROSCI.6133-08.2009>
- Leow, H. C. (2015). Never heard of Sensploration? Time to study up on Epicure's biggest high-end pattern. *The Veox*, December 22nd. <http://www.thevoex.com/never-heard-of-sensploration-time-to-study-up-on-epicures-biggest-high-end-pattern/>.
- Lippman, E. A. (1963). Hellenic conceptions of harmony. *Journal of the American Musicological Society*, 16(1), 3–35. <https://doi.org/10.2307/829917>
- Lippman, E. A. (1964). *Musical thought in ancient Greece*. Columbia University Press.
- Lippman, E. A. (1990). *Musical aesthetics: A historical reader*. Pendragon Press.
- Livio, M. (2008). *The golden ratio: The story of phi, the world's most astonishing number*. Crown.
- Lutz, J. (2007). Transpositions: Architecture as instrument/instrument as architecture. In M. W. Muecke & M. S. Zach (Eds.), *Resonance: Essays on the intersection of music and architecture* (pp. 169–190). Culicidae Press.
- Mahvash, K. (2007). Site + sound: Space. In M. W. Muecke & M. S. Zach (Eds.), *Resonance: Essays on the intersection of music and architecture* (pp. 53–75). Culicidae Press.
- Marks, L. (1978). *The unity of the senses: Interrelations among the modalities*. Academic Press.
- Masataka, N. (2006). Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Developmental Science*, 9(1), 46–50. <https://doi.org/10.1111/j.1467-7687.2005.00462.x>
- McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature*, 535, 547–550. <https://doi.org/10.1038/nature18635>

- McNeil, A. (1993–1994). Why Hindustani musicians are good cooks: Analogies between music and food in North India. *Asian Music*, 25(1/2), 69–80. <https://doi.org/10.2307/834191>
- McNeill Whistler, J. A. (1978). Letter to the world. May 22nd. [Reprinted in the *Gentle Art of Making Enemies*. New York, NY: Dover, 1967, pp. 127–128].
- Meillon, S., Viala, D., Medel, M., Urbano, C., Guillot, G., & Schlich, P. (2010). Impact of partial alcohol reduction in Syrah wine on perceived complexity and temporality of sensations and link with preference. *Food Quality and Preference*, 21(7), 732–740. <https://doi.org/10.1016/j.foodqual.2010.06.005>
- Minati, L., Rosazza, C., D'Incerti, L., Pietrocini, E., Valentini, L., Scaioli, V., Loveday, C., & Bruzzone, M. G. (2009). Functional MRI/event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *NeuroReport*, 20(1), 87–92. <https://doi.org/10.1097/WNR.0b013e32831af235>
- Muecke, M. W., & Zach, M. S. (Eds.). (2007). *Resonance: Essays on the intersection of music and architecture*. Culicidae Press.
- Müller, M. M., Junghöfer, M., Elbert, T., & Roehstroh, B. (1997). Visually induced gamma-band responses to coherent and incoherent motion: A replication study. *NeuroReport*, 8(11), 2575–2579. <https://doi.org/10.1097/00001756-199707280-00031>
- Nagamachi, M. (1995). Kansei engineering: A new ergonomic consumer-oriented technology for product development. *International Journal of Industrial Ergonomics*, 15(1), 3–11. [https://doi.org/10.1016/0169-8141\(94\)00052-5](https://doi.org/10.1016/0169-8141(94)00052-5)
- Newton, I. (1704). *Opticks*. Smith and Walford.
- Ninomiya, K. (2015). Science of umami taste: Adaptation to gastronomic culture. *Flavour*, 4(1), 13. <https://doi.org/10.1186/2044-7248-4-13>
- Ninomiya, K., Kitamura, S., Saiga-Egusa, A., Ozawa, S., Hirose, Y., Kagemori, T., Moriki, A., Tanaka, T., & Nishimura, T. (2010). Changes in free amino acids during heating bouillon prepared at different temperatures. *Journal of Home Economics Japan*, 61(12), 765–773. <https://doi.org/10.11428/jhej.61.765>
- Osgood, C. E. (1962). Studies on the generality of affective meaning systems. *American Psychologist*, 17(1), 10–28. <https://doi.org/10.1037/h0045146>
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois Press.
- Ou, L.-C. (2015). Color emotion and color harmony. In A. J. Elliott, M. D. Fairchild, & A. Franklin (Eds.), *Handbook of color psychology* (pp. 401–418). Cambridge University Press.
- Ou, L., & Luo, M. R. (2006). A colour harmony model for two-colour combinations. *Color Research and Application*, 31(3), 191–204. <https://doi.org/10.1002/col.20208>
- Ou, L.-C., Luo, M. R., Woodcock, A., & Wright, A. (2004). A study of color emotion and color preference. Part II: Color emotions for two-color combinations. *Color Research and Application*, 29(4), 292–298. <https://doi.org/10.1002/col.20024>
- Oyama, T. (2003). Affective and symbolic meanings of color and form; Experimental psychological approaches. *Empirical Studies of the Arts*, 21(2), 137–142. <https://doi.org/10.2190/EYFT-B5UE-APPW-10B6>
- Oyama, T., Yamada, H., & Iwasawa, T. (1998). Synesthetic tendencies as the basis of sensory symbolism; A review of a series of experiments by means of semantic differential. *Psychologia*, 41, 203–215.
- Palmer, S. E. (1991). Goodness, gestalt, groups, and garner: Local symmetry sub-groups as a theory of figural goodness. In G. Lockhead & J. Pomerantz (Eds.), *The perception of structure* (pp. 23–40). American Psychological Association.
- Palmer, S. E., & Griscorn, W. (2013). Accounting for taste: Individual differences in preference for harmony. *Psychonomic Bulletin & Review*, 20(3), 453–461. <https://doi.org/10.3758/s13423-012-0355-2>
- Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual aesthetics and human preference. *Annual Review of Psychology*, 64(1), 77–107. <https://doi.org/10.1146/annurev-psych-120710-100504>
- Park, J. Y., Park, H., Kim, J. I., & Park, H. J. (2011). Consonant chords stimulate higher EEG gamma activity than dissonant chords. *Neuroscience Letters*, 488(1), 101–105. <https://doi.org/10.1016/j.neulet.2010.11.011>
- Parncutt, R. (2014). The emotional connotations of major versus minor tonality: One or more origins? *Musicae Scientiae*, 18(3), 324–353. <https://doi.org/10.1177/1029864914542842>

- Passynkova, N., Neubauer, H., & Scheich, H. (2007). Spatial organization of EEG coherence during listening to consonant and dissonant chords. *Neuroscience Letters*, 412(1), 6–11. <https://doi.org/10.1016/j.neulet.2006.09.029>
- Paulsen, M. T., Rognså, G. H., & Hersleth, M. (2015). Consumer perception of food–beverage pairings: The influence of unity in variety and balance. *International Journal of Gastronomy and Food Science*, 2(2), 83–92. <https://doi.org/10.1016/j.ijgfs.2014.12.003>.
- Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., & Rovelli, R., ... S. Koelsch (2010). Functional specializations for music processing in the human newborn brain. *Proceedings of the National Academy of Sciences of the USA*, 107(10), 4758–4763. <https://doi.org/10.1073/pnas.0909074107>
- Pesic, P. (2014). *Music and the making of modern science*. MIT Press.
- Picabia, F. (n.d.). *Undated manuscript, on hotel Brevourt stationery, Alfred Stieglitz archives, Yale collection of American literature, Beinecke rare book and manuscript library*. Yale University.
- Piesse, G. W. (1891). *Piesse's art of perfumery* (5th ed.). Piesse and Lubin. <http://www.gutenberg.org/files/16378/16378-h/16378-h.htm>.
- Piesse, G. W. S. (1867). *The art of perfumery and the methods of obtaining the odors of plants: With instructions for the manufacture of perfumes for the handkerchief, scented powders, odorous vinegars, dentifrices, Pomatums, cosmetics, perfumed soap, etc., to which is added an appendix on preparing artificial fruit-essences, Etc*. Lindsay & Blakiston.
- Plantinga, J., & Trehub, S. E. (2014). Revisiting the innate preference for consonance. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 40–49. <https://doi.org/10.1037/a0033471>
- Plomp, R., & Levelt, W. J. M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, 38(4), 548–560. <https://doi.org/10.1121/1.1909741>
- Plummer, H. C. (1915). Color music—A new art created with the aid of science. The color organ used in Scriabine's symphony "Prometheus". *Scientific American*, 112(15), 343, 350–351. <https://doi.org/10.1038/scientificamerican04101915-343>
- Prete, G., Bondi, D., Verratti, V., Aloisi, A. M., Rai, P., & Tommasi, L. (2020). Universality vs experience: A cross-cultural pilot study on the consonance effect in music at different altitudes. *PeerJ*, 8, e9344. <https://doi.org/10.7717/peerj.9344>
- Pridmore, R. W. (1992). Music and color: Relations in the psychophysical perspective. *Color Research & Application*, 17(1), 57–61. <https://doi.org/10.1002/col.5080170110>
- Pursey, T., & Lomas, D. (2018). Tate Sensorium: An experiment in multisensory immersive design. *The Senses and Society*, 13(3), 354–366. <https://doi.org/10.1080/17458927.2018.1516026>
- Qiao, X., Wang, P., Li, Y., & Hu, Z. (2014). Study on a correlation model between the kansei image and the texture harmony. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 7(4), 73–84. <https://doi.org/10.14257/ijcip.2014.7.4.07>
- Rameau, J. P. (1722). *Traité de l'harmonie reduite à ses principes naturelles [Treatise on harmony, trans by Philip Gosset, New York: Dover Publications, 1971]*. Ballard.
- Reber, R. (2012). Processing fluency, aesthetic pleasure, and culturally shared taste. In A. P. Shimamura & S. E. Palmer (Eds.), *Aesthetic science: Connecting minds, brains, and experience* (pp. 223–249). Oxford University Press.
- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8(4), 364–382. https://doi.org/10.1207/s15327957pspr0804_3
- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, 9(1), 45–48. <https://doi.org/10.1111/1467-9280.00008>
- Rehding, A. (2019). Consonance and dissonance. In A. Rehding & S. Rings (Eds.), *The Oxford handbook of critical concepts in music theory* (pp. 437–466). Oxford University Press.
- Riedweg, C. (2005). *Pythagoras: His life, teaching, and influence*. Cornell University Press.
- Rimington, A. W. (1895). *A new art: Colour-music*. A paper read at St. James's Hall on June 6, 1895, published in pamphlet form by Messrs. Spottiswoode & Co., New St. Square. June 13, 1895. [Reprinted in "Colour Music, the Art of Light", by A. B. Klein, Lockwood, London, 1930, pp. 256–261].
- Rimington, A. W. (1911). *Colour music: The art of mobile colour*. Frederick A. Stokes.
- Rimington, A. W. (1915). Color music. *Scientific American*, 113(4), 79. <https://doi.org/10.1038/scientificamerican07241915-79>

- Roinard, P., de Labrély, F., Champagne, L., & Deák, N. J. (1976). The song of songs of Solomon. *TDR*, 20(3), 130–135.
- Rozin, P., & Rozin, A. (2018). Advancing understanding of the aesthetics of temporal sequences by combining some principles and practices in music and cuisine with psychology. *Perspectives in Psychological Science*, 13(5), 598–617. <https://doi.org/10.1177/1745691618762339>.
- Rudmin, F., & Cappelli, M. (1983). Tone-taste synesthesia: A replication. *Perceptual & Motor Skills*, 56(1), 118. <https://doi.org/10.2466/pms.1983.56.1.118>
- Russell, M. (1913). "Introduction," *Les Synchronistes*. In S. Macdonald-Wright & M. Russell (Eds.). Paris: Bernheim-Jeune & Cie, October 27th–November 8th. [Translated from the French by M. Laing, and republished in G. Levin, *Synchronism and American Color Abstraction, 1910–1925*, New York, NY, 1978, p. 130].
- Sachs, C. (1953). *Rhythm and tempo: A study in music history*. W. W. Norton.
- Sands, J. (2007). Transgressing boundaries: Considering a societal function of music and architecture through Markus Pernthaler's Helmut-List-Halle. In M. W. Muecke & M. S. Zach (Eds.), *Resonance: Essays on the intersection of music and architecture* (pp. 213–227). Culicidae Press.
- Sato, M., Basirat, A., & Schwartz, J. (2007). Visual contribution to the multistable perception of speech. *Perception & Psychophysics*, 69(8), 1360–1372. <https://doi.org/10.3758/BF03192952>
- Sato, S., Six, J., Pfordresher, P., Fujii, S., & Savage, P. E. (2019). Automatic comparison of global children's and adult songs supports a sensorimotor hypothesis for the origin of musical scales. In Proceedings of the 9th international workshop on folk music analysis (FMA2019), <https://doi.org/10.31234/osf.io/kt7py>
- Schellenberg, E. G., & Trehub, S. E. (1996). Natural musical intervals: Evidence from infant listeners. *Psychological Science*, 7(5), 272–277. <https://doi.org/10.1111/j.1467-9280.1996.tb00373.x>
- Schillinger, J. (1946). *The Schillinger system of musical composition*. Fischer.
- Schloss, K. B., & Palmer, S. E. (2011). Aesthetic response to color combinations: Preference, harmony, and similarity. *Attention, Perception, & Psychophysics*, 73(2), 551–571. <https://doi.org/10.3758/s13414-010-0027-0>
- Schneider, M. (1955). *Singende Steine. Rhythmus-Studien an drei katalanischen Kreuzgängen romanischen Stils [Singing stones. Rhythm studies on three Catalan Romanesque cloisters]*. Bärenreiter.
- Schoenberg, A. (1983). *Theory of harmony*. University of California Press.
- Schöffner, N. (1985). Sonic and visual structures. *Leonardo*, 18(2), 59–68. <https://doi.org/10.2307/1577872>
- Scholes, P. A. (1980). *The Oxford companion to music* (10th ed.) (Edited by J. O. Ward). Oxford University Press.
- Schütte, S. T., Eklund, J., Axelsson, J. R., & Nagamachi, M. (2004). Concepts, methods and tools in Kansei engineering. *Theoretical Issues in Ergonomics Science*, 5(3), 214–231. <https://doi.org/10.1080/1463922021000049980>
- Sebba, R. (1991). Structural correspondence between music and color. *Color Research & Application*, 16(2), 81–88. <https://doi.org/10.1002/col.5080160206>
- Secor, T. M. (2013). *Mystic chord harmonic and light transformations in Alexander Scriabin's "Prometheus"* [MA Thesis]. University of Oregon School of Music.
- Segnit, N. (2010). *The flavour thesaurus: Pairings, recipes and ideas for the creative cook*. Bloomsbury.
- Shen, Y.-C., Chen, Y.-S., & Hsu, W. H. (1996). Quantitative evaluation of color harmony via linguistic-based image scale for interior design. *Color Research and Application*, 21(5), 353–374. [https://doi.org/10.1002/\(SICI\)1520-6378\(199610\)21:5<353::AID-COL5>3.0.CO;2-X](https://doi.org/10.1002/(SICI)1520-6378(199610)21:5<353::AID-COL5>3.0.CO;2-X)
- Shen, Y.-C., Yuan, W.-H., Hsu, W.-H., & Chen, Y.-S. (2000). Color selection in the consideration of color harmony for interior design. *Color Research and Application*, 25(1), 20–31. [https://doi.org/10.1002/\(SICI\)1520-6378\(200002\)25:1<20::AID-COL4>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1520-6378(200002)25:1<20::AID-COL4>3.0.CO;2-5)
- Shepherd-Barr, K. (1999). "Mise en scent": The théâtre d'Art's *Cantique des cantiques* and the use of smell as a theatrical device. *Theatre Research International*, 24(2), 152–159. <https://doi.org/10.1017/S0307883300020770>
- Siefkes, M., & Arielli, E. (2015). An experimental approach to multimodality: How musical and architectural styles interact in aesthetic perception. In J. Wildfeuer (Ed.), *Building bridges for multimodal research: International perspectives on theories and practices of multimodal analysis* (pp. 247–265). Peter Lang.

- Smith, C. (2010). Astringency and harmony in tannins. *Wines & Vines*, September. <https://winesvinesanalytics.com/columns/section/92/article/78078/Astringency-And-Harmony-In-Tannins> (accessed September 26, 2021).
- Smith, C. (2013). *Postmodern winemaking: Rethinking the modern science of an ancient craft*. University of California Press.
- Smith, M. W. (2007). *The total work of art: From Bayreuth to cyberspace*. Routledge.
- Snider, J. G., & Osgood, C. E. (1969). *Semantic differential technique: A sourcebook*. Aldine Publishing.
- South, W. (2001). *Color, myth, and music: Stanton Macdonald-Wright and Synchronism*. Raleigh: North Carolina Museum of Art.
- Spence, C. (2012). Synaesthetic marketing: Cross sensory selling that exploits unusual neural cues is finally coming of age. *The Wired World in 2013*, November, 104–107.
- Spence, C. (2015). Cross-modal perceptual organization. In J. Wagemans (Ed.), *The Oxford handbook of perceptual organization* (pp. 649–664). Oxford University Press.
- Spence, C. (2020a). Assessing the role of emotional mediation in explaining crossmodal correspondences involving musical stimuli. *Multisensory Research*, 33(1), 1–29. <https://doi.org/10.1163/22134808-20191469>
- Spence, C. (2020b). Designing scented colours: On the art & science of olfactory-colour crossmodal correspondences. *The Baltic International Yearbook of Cognition, Logic and Communication*, 14(1), 2. <https://doi.org/10.4148/1944-3676.1125>.
- Spence, C. (2020c). Flavour pairing: A critical review of the literature on food and beverage pairing. *Food Research International*, 133, 109124. <https://doi.org/10.1016/j.foodres.2020.109124>
- Spence, C. (2020d). Multisensory flavour perception: Blending, mixing, fusion, and pairing within and between the senses. *Foods (Basel, Switzerland)*, 9(4), 407. <https://doi.org/10.3390/foods9040407>.
- Spence, C. (2020e). Olfactory-colour crossmodal correspondences in art, science, & design. *Cognitive Research: Principles & Implications (CRPI)*, 5, 52. <https://doi.org/10.1186/s41235-020-00246-1>. <https://rdcu.be/b9oDJ>
- Spence, C. (2021). Musical scents: On the surprising absence of scented musical/auditory events entertainments, and experiences. *i-Perception*, 12(4), 1–31. <https://doi.org/10.1177/20416695211038747>
- Spence, C. (2022). Gastrophysics: Getting creative with pairing flavours. *International Journal of Gastronomy & Food Science*, 27, 100433. <https://doi.org/10.1016/j.ijgfs.2021.100433>
- Spence, C., Deroy, O., & Bremner, A. (2013). Questioning the utility of the concept of amodality: Towards a revised framework for understanding crossmodal relations. *Multisensory Research*, 26(Suppl.), 57. <https://doi.org/10.1163/22134808-000S0036>
- Spence, C., & Wang, Q. J. (2018). On the meaning(s) of complexity in the chemical senses. *Chemical Senses*, 43(7), 451–461. <https://doi.org/10.1093/chemse/bjy047>
- Spence, C., Wang, Q. J., Reinoso-Carvalho, F., & Keller, S. (2021). Commercializing sonic seasoning in multisensory offline experiential events and online tasting experiences. *Frontiers in Psychology*, 12, 740354. <https://doi.org/10.3389/fpsyg.2021.740354>
- Spence, W., & Guilford, J. P. (1933). The affective value of combinations of odors. *American Journal of Psychology*, 45(3), 495–501. <https://doi.org/10.2307/1415049>
- Sterken, S. (2007). Music as an art of space: Interactions between music and architecture in the work of Iannis Xenakis. In M. W. Muecke & M. S. Zach (Eds.), *Resonance: Essays on the intersection of music and architecture* (pp. 21–51). Culicidae Press.
- Stokes, J. (1972). *Resistible theatres: Enterprise and experiment in the late nineteenth century*. Paul Elek Books.
- Stone, R. M. (1982). *Let the inside be sweet: The interpretation of music event among the Kpelle of Liberia*. Indiana University Press.
- Stratton, G. M. (1899). The spatial harmony of touch and sight. *Mind: A Quarterly Review of Psychology and Philosophy*, 8(32), 492–505. <https://doi.org/10.1093/mind/VIII.4.492>
- Stumpf, C. (1883/1890). *Tonpsychologie [Psychology of tones]* (Vols 1 & 2). S. Hirzel.
- Sullivan, J. W. N. (1914). An organ on which color compositions are played. The new art of color music and its mechanism. *Scientific American*, 110(8), 163, 170. <https://doi.org/10.1038/scientificamerican02211914-163>
- Tabas, A., Andermann, M., Schuberth, V., Riedel, H., Balaguer-Ballester, E., & Rupp, A. (2019). Modeling and MEG evidence of early consonance processing in auditory cortex. *PLoS Computational Biology*, 15(2), e1006820. <https://doi.org/10.1371/journal.pcbi.1006820>.

- Tan, S. L., Cohen, A., Lipscomb, S., & Kendall, R. (Eds.). (2013). *The psychology of music in multimedia*. Oxford University Press.
- Tatarkiewicz, W. (1980). *A history of six ideas: An essay in aesthetics*. Njijhoff.
- Tervaniemi, M., Sannemann, C., Noyranen, M., Salonen, J., & Pihko, E. (2011). Importance of the left auditory areas in chord discrimination in music experts as demonstrated by MEG. *European Journal of Neuroscience*, *34*(3), 517–523.
- Tillmann, B., & Bharucha, J. J. (2002). Effect of harmonic relatedness on the detection of temporal asynchronies. *Perception & Psychophysics*, *64*(4), 640–649. <https://doi.org/10.3758/BF03194732>
- Trainor, L. J., & Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior and Development*, *21*(1), 77–88. [https://doi.org/10.1016/S0163-6383\(98\)90055-8](https://doi.org/10.1016/S0163-6383(98)90055-8)
- Trainor, L. J., Tsang, C. D., & Cheung, V. H. W. (2002). Preference for sensory consonance in 2- and 4-month-old infants. *Music Perception*, *20*(2), 187–194. <https://doi.org/10.1525/mp.2002.20.2.187>
- Tramo, M. J., Cariani, P. A., Delgutte, B., & Braida, L. D. (2001). Neurobiological foundations for the theory of harmony in Western tonal music. *Annals of the New York Academy of Sciences*, *930*(1), 92–116. <https://doi.org/10.1111/j.1749-6632.2001.tb05727.x>
- Turek, R. (1976). *The elements of music: Concepts and applications* (Vol. 1). McGraw-Hill.
- Uhlhaas, P. J., Pipa, G., Neuenschwander, S., Wibral, M., & Singer, W. (2011). A new look at gamma? High (>60 Hz) γ -band activity in cortical networks: Function, mechanisms and impairment. *Progress in Biophysics and Molecular Biology*, *105*(1–2), 14–28. <https://doi.org/10.1016/j.pbiomolbio.2010.10.004>
- Van de Geer, J. P., & Levelt, W. J. M. (1962). The connotations of musical consonance. *Acta Psychologica*, *20*, 308–319. [https://doi.org/10.1016/0001-6918\(62\)90028-8](https://doi.org/10.1016/0001-6918(62)90028-8)
- 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. <https://doi.org/10.3389/fpsyg.2014.01352>
- Velasco, C., & Spence, C. (in press). Capitalizing on the crossmodal correspondences between audition and olfaction in the design of multisensory experiences. In N. Pomirleanu, B. J. Mariadoss, & J. Schibrowsky (Eds.), *Experiential marketing in an era of hyper-connectivity*. Cambridge Scholars Publishing.
- Virtala, P., Huottilainen, M., Partanen, E., Fellman, V., & Tervaniemi, M. (2013). Newborn infants' auditory system is sensitive to Western music chord categories. *Frontiers in Psychology*, *4*, 492. <https://doi.org/10.3389/fpsyg.2013.00492>
- Vitruvius (1999). *Ten books on architecture*. Cambridge University Press.
- von Békésy, G. (1959a). Similarities between hearing and skin sensations. *Psychological Review*, *66*(1), 1–22. <https://doi.org/10.1037/h0046967>
- von Békésy, G. (1959b). Synchronism of neural discharges and their demultiplication in pitch perception on the skin and in hearing. *Journal of the Acoustical Society of America*, *31*(3), 338–349. <https://doi.org/10.1121/1.1907722>
- von Erhardt-Siebold, E. (1932). Harmony of the senses in English, German, and French romanticism. *PMLA*, *47*, 577–592.
- von Helmholtz, H. (1863). *On the sensations of tone as a physiological basis for the theory of music*. Dover.
- von Helmholtz, H. (1878/1971). *Treatise on physiological optics* (Vol. II). Dover Publications.
- von Helmholtz, H. (1954). *On the sensations of tone* (A. J. Ellis, Ed.). Dover Publications.
- Wagner, B., Bowling, D. L., & Hoeschele, M. (2020). Is consonance attractive to budgerigars? No evidence from a place preference study. *Animal Cognition*, *23*(5), 973–987. <https://doi.org/10.1007/s10071-020-01404-0>
- Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). “Metaphorical” mapping in human infants. *Child Development*, *52*, 728–731. <https://doi.org/10.2307/1129200>
- Walker, D. P. (1967). Kepler's celestial music. *Journal of the Warburg and Courtauld Institutes*, *30*, 228–250. <https://doi.org/10.2307/750744>
- Walker, P. (2012). Cross-sensory correspondences and cross talk between dimensions of connotative meaning: Visual angularity is hard, high-pitched, and bright. *Attention, Perception, & Psychophysics*, *74*(8), 1792–1809. <https://doi.org/10.3758/s13414-012-0341-9>

- Walker, P., & Walker, L. (2012). Size-brightness correspondence: Crosstalk and congruity among dimensions of connotative meaning. *Attention, Perception, & Psychophysics*, 74(6), 1226–1240. <https://doi.org/10.3758/s13414-012-0297-9>
- Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound. *Perception & Psychophysics*, 42(5), 491–502. <https://doi.org/10.3758/BF03209757>
- Wang, Q. (J.), & C. Spence (2016). “Striking a sour note”: Assessing the influence of consonant and dissonant music on taste perception. *Multisensory Research*, 29(1-3), 195–208. <https://doi.org/10.1163/22134808-00002505>
- Wang, Q. (J.), S. Wang & C. Spence (2016). “Turn up the taste”: Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and pitch. *Chemical Senses*, 41(4), 345–356. <https://doi.org/10.1093/chemse/bjw007>
- Ward, J., Moore, S., Thompson-Lake, D., Salih, S., & Beck, B. (2008). The aesthetic appeal of auditory-visual synaesthetic perceptions in people without synaesthesia. *Perception*, 37(8), 1285–1297. <https://doi.org/10.1068/p5815>
- Weiss, M. W., Cirelli, L. K., McDermott, J. H., & Trehub, S. E. (2020). Development of consonance preferences in western listeners. *Journal of Experimental Psychology: General*, 149(4), 634–649. <https://doi.org/10.1037/xge0000680>
- Wells, A. (1980). Music and visual color: A proposed correlation. *Leonardo*, 13, 101–107. <https://doi.org/10.2307/1577978>
- Westland, S., Laycock, K., Cheung, V., Henry, P., & Mahyar, F. (2007). Colour harmony. *Colour: Design & Creativity*, 1(1), 1–15.
- Whiteford, K. L., Schloss, K. B., Helwig, N. E., & Palmer, S. E. (2018). Color, music, and emotion: Bach to the blues. *i-Perception*, 9(6), 1–27. <https://doi.org/10.1177/2041669518808535>
- Whitney, J. (1980). *Digital harmony: On the complementarity of music and visual art*. Peterborough, NH: Byte Books.
- Yau, J. M., Olenczak, J. B., Dammann, J. F., & Bensmaia, S. J. (2009). Temporal frequency channels are linked across audition and touch. *Current Biology*, 19(7), 561–566. <https://doi.org/10.1016/j.cub.2009.02.013>
- Yoo, Y., Hwang, I., & Choi, S. (2014). Consonance of vibrotactile chords. *IEEE Transactions on Haptics*, 7(1), 3–13. <https://doi.org/10.1109/TOH.2013.57>
- Zarlino, G. (1558). *Le Istitutioni Harmoniche [Elements of harmony]*. Pietro da Fino.
- Zentner, M. R., & Kagan, J. (1998). Infants’ perception of consonance and dissonance in music. *Infant Behavior & Development*, 21(3), 483–492. [https://doi.org/10.1016/S0163-6383\(98\)90021-2](https://doi.org/10.1016/S0163-6383(98)90021-2)
- Zhang, J., Harbottle, G., Wang, C., & Kong, Z. (1999). Oldest playable musical instruments found at Jiahu early Neolithic site in China. *Nature*, 401(6751), 366–368. <https://doi.org/10.1038/43865>
- Zhang, Y. (2007). Altar and studio: Musical design in 18th-century Chinese architecture. In M. W. Muecke & M. S. Zach (Eds.), *Resonance: Essays on the intersection of music and architecture* (pp. 191–211). Culicidae Press.
- Zilczer, J. (1987). “Color music”: Synaesthesia and nineteenth-century sources for abstract art. *Artibus et Historiae*, 8(16), 101–126. <https://doi.org/10.2307/1483303>
- Zilczer, J. (2016). American Rhapsody: From Modern to Postmodern in visual music. In T. Kaduri (Ed.), *The Oxford handbook of sound and image in western art* (pp. 17–35). Oxford, UK: Oxford University Press.

How to cite this article

Spence, C., & Di Stefano, N. (2022). Crossmodal harmony: Looking for the meaning of harmony beyond hearing. *i-Perception*, 13(1), 1–40. <https://doi.org/10.1177/20416695211073817>