

Assessing the Role of Emotional Mediation in Explaining Crossmodal Correspondences Involving Musical Stimuli

Charles Spence*

Department of Experimental Psychology, Anna Watts Building, University of Oxford, Oxford, OX2 6GG, UK

Received 7 April 2019; accepted 2 July 2019

Abstract

A wide variety of crossmodal correspondences, defined as the often surprising connections that people appear to experience between simple features, attributes, or dimensions of experience, either physically present or else merely imagined, in different sensory modalities, have been demonstrated in recent years. However, a number of crossmodal correspondences have also been documented between more complex (i.e., multi-component) stimuli, such as, for example, pieces of music and paintings. In this review, the extensive evidence supporting the emotional mediation account of the crossmodal correspondences between musical stimuli (mostly pre-recorded short classical music excerpts) and visual stimuli, including colour patches through to, on occasion, paintings, is critically evaluated. According to the emotional mediation account, it is the emotional associations that people have with stimuli that constitutes one of the fundamental bases on which crossmodal associations are established. Taken together, the literature that has been published to date supports emotional mediation as one of the key factors underlying the crossmodal correspondences involving emotionally-valenced stimuli, both simple and complex.

Keywords

Crossmodal correspondence, complex, music, emotional mediation, painting

1. Introduction

It is a fact widely acknowledged that music induces emotion in the listener (e.g., Juslin and Sloboda, 2010; Juslin and Västfjäll, 2008; Marin and Bhattacharya, 2010; North and Hargreaves, 2008; though for an exception in the case of musical anhedonia, see Mas-Herrero *et al.*, 2014). Furthermore, an extensive body of research has also demonstrated how listening to music can

* E-mail: charles.spence@psy.ox.ac.uk

influence people's judgments concerning a variety of other stimuli that they may happen to be evaluating at around the same time (e.g., Logeswaran and Bhattacharya, 2009; Marin *et al.*, 2012, 2017; Tannenbaum, 1956; Zuckerman, 1949). However, beyond its effect on mood/emotion, and any crossmodal effects on stimulus evaluation, there is also a growing body of scientific research suggesting that people experience crossmodal correspondences between short musical excerpts and a range of visual stimuli. Arnheim (1986, p. 207) put it thus: "*There are... perceptually convincing correspondences between colors and sounds based on shared expressive qualities.*"

Crossmodal correspondences have been defined as the often surprising connections that people appear to experience between simple features, attributes, or dimensions of experience, either physically present or else merely imagined, in different sensory modalities (Spence, 2011, 2018). Musical correspondences have been demonstrated with everything from simple colour patches (e.g., Cutietta and Haggerty, 1987; Isbilen and Krumhansl, 2016; Karwoski and Odber, 1938; Odber *et al.*, 1942; Palmer *et al.*, 2013, 2016; Whiteford *et al.*, 2018) through to complex visual stimuli such as paintings (e.g., Albertazzi *et al.*, 2015) and even mathematical arguments (Johnson and Steinerberger, 2019). That said, the literature on crossmodal correspondences involving more complex auditory stimuli ('complex' stimuli are operationally defined here as those containing multiple individuable elements, see Eitan, 2017; Walker, 2016, for reviews) stands a little apart from the older literature on crossmodal correspondences between simple sensory stimuli, such as individual tones or visual stimuli such as outline shapes or colour patches (see Marks, 2004; Spence, 2011, 2018, for reviews). Nevertheless, it is interesting to consider the extent to which a similar explanatory framework may operate in the two cases (see Spence, in press, for a review). This, then, is one of the aims of the present review. To be absolutely clear, the specific question addressed here is the extent to which emotional mediation provides a plausible account for the crossmodal correspondences that have been documented involving complex musical stimuli (see Note 1).

One of the most relevant differences between those crossmodal correspondences involving simple and complex stimuli is that the emotional associations with pieces of composed music, which is what have been used as the complex auditory stimuli in the majority of recent empirical research in this area, are presumably likely to be more pronounced than anything elicited by the presentation of a single musical note, say (see also Valentine, 1962). As such, it should perhaps come as little surprise to find that emotional mediation tends to account for more of the variance in the empirical matching data than is the case for those crossmodal correspondences between simple, and typically less emotionally valent, sensory stimuli. That said, the literature in this area is often unclear as to whether it is the emotional associations that people have with

music or those that are evoked by listening to it that are relevant (Cespedes-Guevara and Eerola, 2018; Gabrielsson, 2001; Whiteford *et al.*, 2018; see also Kawakami *et al.*, 2013) (Note 2).

There is a long history of research assessing the feeling value, or affective tone, of lines etc. (e.g., Collier, 1996; Hevner, 1935; Lundholm, 1921; Poffenberger and Barrows, 1924). In fact, the emotional mediation account can itself perhaps be seen as emerging from, or at the very least would seem to be meaningfully related to, Osgood and colleagues' search for connotative meaning, as captured in their early work on the Semantic Differential Technique (e.g., Osgood, 1952; Osgood *et al.*, 1957). According to the literature, the meaning of any stimulus, be it simple or complex, semantic, conceptual, or sensory, can be determined by having people rate their feelings about it on a number of semantic differential scales. These seven-interval scales typically include pairs of bipolar adjectives, such as good/bad, active/passive, dominant/submissive, etc., acting as anchors at either end of the scale (see also Parrott, 1982).

In their early research, Osgood and his colleagues would often have people rate the meaning of particular terms using as many as 50 different semantic differential (bipolar adjective) scales. However, upon closer analysis using factor analytical approaches (e.g., Osgood and Suci, 1955), it often turned out that no more than three dimensions, specifically those listed above (namely valence, activity, and potency), could often explain the majority of the variance in people's responses. Note here how the first two of these dimensions, good/bad and active/passive, map neatly onto valence (pleasantness) and arousal, commonly considered to be the primary dimensions of core affect (e.g., Cespedes-Guevara and Eerola, 2018; Marin, Gingras and Bhattacharya, 2012) (see Note 3).

While the semantic differential technique was initially developed in order to help determine the meaning that people attached to words and concepts, researchers soon extended the approach in order to assess the meaning of a range of *sensory* stimuli, including everything from simple lines/shapes (e.g., Bozzi and Flores D'Arcais, 1967; Janković, 2014; and see Cowles, 1935; Karwoski *et al.*, 1942, for precursors to this approach) and colours (Adams and Osgood, 1973), through to complex pieces of music (e.g., Watson, 1942; Watt and Quinn, 2007), paintings (Albertazzi *et al.*, 2015), poetry, and architecture (Hasenfus *et al.*, 1983). In fact, a semantic differential assessment, or profile, can presumably be established for any concept, object, or stimulus. As such, this then provides one means by which the participants in laboratory-based, or increasingly online, research might presumably be able to establish a cross-modal correspondence. That is, to feel some sort of affinity between pairs of seemingly unrelated stimuli. Concretely, the suggestion is that people might match those stimuli that are more similar in terms of their meaning, or profile, in some multi-dimensional semantic differential space (Note 4).

Composed music is typically made up of a variety of different elements, including different musical notes, timbres, and vocal components, etc. As such, it is, *a priori*, uncertain as to whether people will preferentially experience, or establish, crossmodal correspondences with either the music's basic sensory features, such as its average pitch range, loudness, or timbre (Wallmark, 2019), or with more complex parameters such as, for example, the music's orchestration, melodic structure, tempo, or attack (see Cowles, 1935; Hubbard, 1996). It is an interesting empirical question, therefore, to consider which attribute(s) of a given musical excerpt may end up dominating in this regard. One might legitimately wonder, for instance, whether the mode, pitch, or timbre normally dominate over the roughness or tempo, of the music, say? Perhaps, though, the answer to this question depends on the particular range of stimuli used, and, more importantly, the contrasts, that are emphasized within a given experimental context, or musical excerpt.

That being said, the auditory dimension of pitch undoubtedly appears to have attracted more than its fair share of empirical and theoretical research. Perhaps this suggests that it is, in some sense, the dominant perceptual/conceptual sonic dimension, typically corresponding with elevation (Deroy *et al.*, 2018; Spence, 2019a). At the same time, however, it should be remembered that pitch also has a number of other connotations too: e.g., with size (Evans and Treisman, 2010; Gallace and Spence, 2006; Parise and Spence, 2012), angularity (Parise and Spence, 2012; see also Murari *et al.*, 2015), lightness (Hubbard, 1996; Marks, 1974, 1975, 1987; Wicker, 1968; cf. Simpson *et al.*, 1956), and brightness/lightness (Marks, 1987; Saysani, 2019; Walker, 2012; see Spence, 2011, 2018, for reviews).

Here, it is worth remembering that crossmodal correspondences are thought to be relative phenomena (e.g., see Brunetti *et al.*, 2018; Spence, 2011, 2019a; Walker and Walker, 2016). In other words, it is the higher-pitched of two sounds that is matched with the brighter, or higher, of two visual stimuli, rather than there being a specific match between a sound having a particular pitch and a specific visual stimulus, or elevation, say. In fact, it is the supposedly *relative* nature of the correspondences that stands in such striking contrast with early attempts to establish a more precise, or absolute, mapping between particular musical notes and specific hues; Sir Isaac Newton, for one, making a famous early suggestion along just these lines. In his *Opticks* (book III, part I, qu. 13–14), Newton (1730/1952) proposed that musical tones and colour tones shared common frequencies [see also von Goethe's book, "*Theory of Color*" (Goethe, 1810/1840) — see Caivano, 1994, for a discussion]. This idea having been revisited subsequently by a number of other researchers (e.g., see Caivano, 1994; Pridmore, 1992; see also Anikin and Johansson, 2019). In fact, as pointed out by Pridmore, there are only two physical variables as far

as sonic/radiant energy are concerned, namely amplitude (e.g., loudness or brightness) and wavelength (varying musical tone/note or hue).

The relative nature of so many of the crossmodal correspondences that have been documented to date also stands in stark contrast to the absolute nature of the cross-sensory mappings that are such a distinctive feature of synaesthesia proper. Deroy and Spence (2013) have reviewed the many differences between these two, often conflated, phenomena (see also Karwoski *et al.*, 1942; Lehman, 1972; Rader and Tellegen, 1987; Ward *et al.*, 2006). However, while the relative nature of so many of the correspondences undoubtedly makes sense in the case of those stimuli that can be organized prothetically (e.g., along a more than/less than dimension), it is tempting to consider the possibility that there might actually be more of an absolute mapping in the case of those stimuli that are organized metathetically instead (see Stevens, 1957, on the prothetic/metathetic distinction).

Prototypical classes of metathetic perceptual stimuli include sound timbre (Adeli *et al.*, 2014), pitch (Stevens, 1957), shape (Smith and Sera, 1992), basic taste qualities (i.e., bitter, salty, sour, sweet, and umami; Spence, 2011, 2019b), and flavours. Here, then, in the context of the debate concerning the absolute vs relative nature of crossmodal correspondences (see Spence, 2019a), it is interesting to note how often absolute matches between the timbre of different instruments are made with specific colours/tastes/flavours (e.g., Crisinel and Spence, 2010; Karwoski *et al.*, 1942; Spence *et al.*, 2015). While many such matches originate in the writings, or responses, of self-confessed synaesthetes (e.g., Kandinsky, 1925/1979), that is certainly by no means always the case (Ione and Tyler, 2004). For instance, Kandinsky (1914/1977, p. 38), in his book *Concerning the Spiritual in Art* wrote that: “a light blue is like a flute, a darker blue a cello; still darker a thunderous double bass; and the darkest blue of all — an organ.”

Pitch and hue are particularly interesting dimensions to consider given that stimuli having such properties may be organized either prothetically (in terms of frequency or wavelength) and/or metathetically (e.g., in terms of their hue). Meanwhile, hue is also a relatively unique perceptual dimension in that it is one of the few where the stimuli can be arranged in a circular manner (e.g., Gilbert *et al.*, 2016). In fact, the fact that both tones and hues can be arranged cyclically, as hue cycle and octave cycle, respectively, has, over the years, led a number of commentators (from a surprisingly diverse range of research backgrounds, it should be said) to suggest that these two dimensions might be meaningfully related/correlated (i.e., based on the perceived structure/arrangement of stimuli within each modality; see Caivano, 1994; Garner, 1978; Pridmore, 1992; Wells, 1980; though see also Davis, 1979). Wells, for instance, suggested that fruitful parallels could be drawn between the mixing of colours, and of tones, in terms of the deriving of harmonious combinations.

Notice here also how the circular organization of pitch and hue provides a natural mapping with the circumplex of affect (Russell, 1980; see also Odbert *et al.*, 1942; and see Fig. 1 below).

2. Crossmodal Correspondences Involving Complex Auditory Stimuli

2.1. Music–Colour Correspondences

As has been mentioned already, in terms of crossmodal correspondences involving complex and/or meaningful multi-element auditory stimuli, short musical excerpts have most often been used as experimental stimuli. Using such stimuli, researchers have been able to establish that people experience or, at the very least, exhibit reliable correspondences, or matches, between music and colour. So, for instance, in one influential study, Palmer and his colleagues (Palmer *et al.*, 2013) were able to demonstrate that people consistently associated different pieces of classical music with different colour patches. In their first study, the participants listened to 18 short pieces of classical orchestral music by Bach, Brahms, and Mozart. These 18–50 second musical selections varied in terms of their tempo (slow/medium/fast) and mode (major/minor). For each of the musical excerpts, the participants had to pick the five best-matching and the five worst-matching colours in order from the 37 carefully-selected Berkeley Colour Project colour patches, displayed simultaneously. The colour patches included: “*red, orange, yellow, chartreuse, green, cyan and blue and purple at four different lightness-saturation levels (saturated, light, muted, and dark), plus three greys, white, and black.*” (Palmer *et al.*, 2013, p. 8837; see also Murari *et al.*, 2014).

Intriguingly, the results revealed some surprisingly strong crossmodal correspondences between the musical selections and the colours chosen. Generally speaking, the faster tempi musical pieces, as well as those pieces played in major mode, were associated with more saturated, lighter, and yellower (i.e., warmer) colours. By contrast, those slower musical selections played in the minor mode were associated with darker, desaturated, and bluer colours instead. Furthermore, the six pieces of music from Brahms were associated with less saturated, darker, bluer colours than were the musical selections from Bach and Mozart, the music from the latter two composers not differing significantly from one another. Here, though, it is perhaps also worth pointing out that the fast Brahms pieces were significantly slower than the fast selections from Mozart and Bach.

The participants in Palmer *et al.*’s (2013) first experiment also had to rate the musical excerpts and the colour patches in terms of how strongly they associated them with each of eight emotional descriptors (happy, sad, angry, calm, strong, weak, lively, and dreary) using line scales ranging from –100 to

100. The results highlighted some very strong correlations ($0.89 < r < 0.99$) between the emotional associations of the musical excerpts and those of the colour patches that were chosen to go with them. This result supports an emotional mediation account of at least this specific crossmodal correspondence (cf. Barbieri *et al.*, 2007). According to Palmer *et al.* (2013, p. 3), the unfolding of the emotional mediation mechanism might be described thus: “*as people listen to the music, they have emotional responses... and then pick colors with similar emotional content [as the music].*” Furthermore, multi-dimensional scaling revealed a 2D solution (consistent with the dimensions of *valence* and *potency*) that was capable of accounting for 95% of the variance.

Additional support for the emotional mediation account was provided by the results of two further experiments in which Palmer *et al.*'s (2013) participants matched either the colour patches (Experiment 2) or the classical music selections (Experiment 3) to one of number of faces displaying a range of different emotional expressions. There were 14 faces in Experiment 2 and 13 in Experiment 3. Once again, a number of robust correlations were obtained. Moreover, it turned out that the music–colour matches that had been documented in Palmer *et al.*'s first experiment could be predicted on the basis of the music-emotion and colour-emotion results documented in these researchers' latter two experiments.

The fact that Palmer and his colleagues (2013) obtained similar results in groups of students from both California and Mexico hints at the possible cross-cultural generalizability of their results. Some years earlier, Walker (1987) had also looked for any cross-cultural similarities/differences in musical metaphors. Indeed, while oft-cited, Palmer *et al.*'s study is by no means unique in demonstrating such a close connection between music and visual stimuli (specifically colours). In fact, over the years, a number of other studies have also documented the existence of music–colour correspondences, often positing emotional mediation as the underlying explanation (e.g., see Barbieri *et al.*, 2007; Bresin, 2005; Cutietta and Haggerty, 1987; Isbilen and Krumhansl, 2016; Karwoski and Odbert, 1938; Lindborg and Friberg, 2015; Odbert *et al.*, 1942; Sebba, 1991). That said, many of the earlier studies in this area can be criticized for only presenting a very limited number of musical selections (see Table 1 for details). That said, research by Bresin (2005) is largely consistent with Palmer *et al.* in suggesting that music played in the major mode is typically associated with lighter colours than music in the minor mode.

The participants in Bresin's (2005) study listened to two pieces of music while viewing 24 colour patches. The first piece of music was Brahms' 1st theme of the poco allegretto 3rd movement Symphony Op. 90 No. 3, in C minor. The second was Haydn's theme from the first movement of Quartet in F major for strings, Op. 74 No. 2. These pieces were played by one of three

Table 1.

Chronologically-organized summary of research studies investigating music-colour matching, one form of crossmodal correspondence involving complex stimuli

Study	Number of musical samples	Colour stimuli	Participants	Results
Odbert <i>et al.</i> (1942)	10 short musical pieces	Name colour associated	243 participants	Grey associated with sad music; Red, yellow, green, and blue for happy music
Sebba (1991)	2 18th-century classical pieces Mozart major/Albinoni minor	Unconstrained	87 architecture students	Warmer, more saturated, lighter colours
Bresin (2005)	2 melodies × 3 instruments × 12 emotional intentions	24 colour patches	2 groups of 11 participants	Major mode associated with lighter colours than music in minor mode
Barbiere <i>et al.</i> (2007)	4 classical musical selections, 2 happy and 2 sad	Colour words	20 US UGs (UG = undergraduate)	Grey associated with sadder music; red, yellow, green, blue associated with happier music
Palmer <i>et al.</i> (2013)	18 orchestral classical pieces from Bach, Brahms and Mozart; varying in tempo (slow/medium/fast) and mode (major/minor)	37 colour patches	48 US UGs; 49 Mexican UGs	Faster music in major mode associated with more saturated, lighter and yellower colours; slower, minor music associated with desaturated, darker and bluer colours
Lindborg & Fribers (2015)	27 film music excerpts	CIE Lab colour space	22 participants	Happy music associated with yellow; red with anger; dark blue with sad music
Isbilen and Krumhansl (2016)	24 Preludes from Bach's <i>Well-Tempered Clavier</i>	8 saturated colours	45 Non-musicians, 83 musicians (19 w. absolute pitch) and music-colour synaesthetes	Colour correspondence with tempo, mode, pitch height and attack rate. Similar results for different participants
Palmer <i>et al.</i> (2016)	64 single-line piano melodies by Mozart, varying in tempo (slow/fast), note-density (sparse/dense), mode (major/minor) and pitch-height (low/high)	37 colour patches	20 US UGs	Faster music in the major mode associated with lighter, more saturated, yellower (warmer) colours than slower music in the minor mode

Table 1.
(Continued)

Study	Number of musical samples	Colour stimuli	Participants	Results
Whiteford <i>et al.</i> (2018)	34 musical excerpts, from many genres, inc. Blues, Salsa, Heavy Metal, Ska, Dubstep	37 colour patches	30 participants	Highly significant correlations reported between colour–music matches and the similarity of the emotional content

different instruments — piano, guitar, and saxophone — so as to express 12 different emotions, namely, happiness, love, contentment, pride, curiosity, indifference, sadness, fear, shame, anger, jealousy, and disgust. There were, in other words, 72 musical tracks in total. The results highlighted different colour profiles for different performances of one and the same piece of music, thus illustrating the role of expressivity in music as far as explaining the matches made to colour. In another early study, Lehman (1972) had his participants listen to a series of orchestral excerpts and rate how well each one was described by each of 18 colour words. Testing both those who could see colours in response to music and those who could not, the responses were very similar for both groups. The participants in this study also had to rate how well each of the pieces of music matched each of 18 adjectives.

The 87 first-year architecture students who took part in Sebba's (1991) study had to render a piece of 18th-century music using colour. That is, they were encouraged to transfer their impression of the music in a free-style manner. Those participants ($n = 34$) who chose to listen to Mozart's Serenade in C-Major — "*Eine Kleine Nachtmusik*" — incorporated warmer, more saturated, and lighter colours (i.e., red, orange, yellow) with higher contrast than did those participants ($n = 53$) who chose to paint while listening to Albinoni's Adagio in C-minor instead: blue, black, and purple were found to be the dominant colours for the latter selection.

The 20 North American psychology student volunteers who took part in a study by Barbieri *et al.* (2007) had to assign five points to 11 colour names in response to four 90-s clips from classical musical selections. The participants could either assign all of the points to one colour or else distribute them as they saw fit across the different colours. The music selections were *Morning Mood* by Grieg (track 4), *Pictures at an Exhibition* Mussorgsky/Ravel (track 5), *Ase's Death* by Grieg (track 5), and *Adagio for Strings* by Samuel Barber (track 7). In this case, the results demonstrated that grey was associated with the sadder musical selections (the latter two pieces; with grey being assigned more

than 25% of the points). Barbieri *et al.*'s participants associated the colours yellow/red/blue/green with the happier music selections instead, comprising the first pair of tracks listed above.

The sad–grey association reported by Barbieri *et al.* (2007) essentially replicated early results from Odbert *et al.* (1942). In the latter study, participants listened to 10 short song clips and picked the colour, or colours, that they thought most appropriate for each piece of music. The participants also picked a selection of emotion-related adjectives that best matched each piece of music. Finally, they picked the colours that matched those adjectives. The results provided some of the earliest evidence in support of the emotional mediation account of crossmodal correspondences between music and colour. In particular, grey was chosen for sad songs, while red, yellow, green, and blue were chosen for those songs described as gay or happy.

In a more recent study reported by Lindborg and Friberg (2015), 22 participants recruited from a school of art, design, and media listened to 27 film music excerpts while continuously manipulating the size and colour of a colour patch on screen using a tablet interface that allowed the participants to navigate the CIE Lab colour space. The results once again confirmed previous findings in showing that participants associated happy music with the colour yellow. Meanwhile, music that expressed anger was associated with the colour red, and sad music was associated with smaller patches toward dark blue. Subsequent data analysis revealed that models including emotion were able to explain 60–75% of the variation in each of the colour patch parameters. Such results can be taken as suggesting emotional mediation as the dominant mechanism underpinning people's choice of colours when matching to music. Furthermore, interviews with the participants also revealed that they tended to report matching the colour associated with the perceived emotion of each piece of music when asked: "*How would you describe the way you chose a colour for a sound?*" (see Fig. 1).

In 2018, Whiteford *et al.* presented 30 participants with 34 15-s musical excerpts from a variety of different genres including blues, salsa, heavy metal, jazz, country & western, hip-hop, Arabic music, etc. Once again, the participants were instructed to match each of the musical selections presented in a random order to the 37 Berkeley colour patches (mentioned earlier). In particular, they were instructed to pick the three best-matching colours followed by the three worst-matching colours. Thereafter, the participants rated the music clips against ten emotion dimensions. Highly significant correlations were reported between colour–music matches and the similarity of the emotional content for all but one of the emotion dimensions evaluated. Specifically, the dimensions were anchored by appealing vs disgusting; calm vs agitated; complex vs simple; happy vs sad; harmonious vs disharmonious; loud vs quiet; spicy vs bland; warm vs cool; whimsical vs serious. By contrast, there was

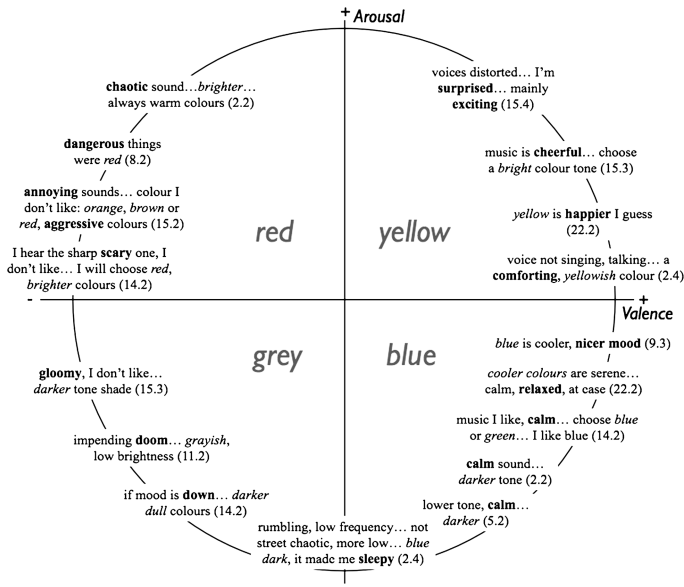


Figure 1. Fragments of spoken interview responses indicating an association between colour and emotion words, distributed in the Valence–Arousal circumplex. Numbers in parentheses refer to the participant number ($n = 22$) and the interview question (2) “How would you describe the way you chose a colour for a sound?”; (3) “What aspects of the colour did you focus on?”; and (4) “What aspects of the sound or music did you focus on?”. [Figure reprinted from Lindborg and Friberg, 2015, Figure 5; doi:10.1371/journal.pone.0144013.g005.]

only a very weak correlation when considering the participants’ preference for the various colour patches and musical selections. So, for example, “*loud, punchy, distorted music was generally associated with darker, redder, more saturated colours*” (Whiteford *et al.*, 2018, p. 1). Analysis of the results of this study (involving parallel factor analysis) were consistent with the view that factors that were consistent with the dimensions of arousal and valence were key emotional attributes mediating the crossmodal association, or correspondence, between music and colour. Bear in mind here also that arousal and valence are two of the three key dimensions highlighted by work on the semantic differential technique (see Osgood *et al.*, 1957).

Elsewhere, Isbilen and Krumhansl (2016) had 89 participants match one of eight saturated colours taken from the Berkeley Colour Project to excerpts from 24 of the Preludes from Bach’s *Well-Tempered Clavier*. In this study, the participants’ colour choices grouped the Preludes together according to their tempo, mode, pitch height, and attack rate. In two follow-up experiments, the participants rated the colour patches and music excerpts using nine 11-point emotion scales labelled: lively, happy, positive, weak, calm, sad, negative, strong, and angry. Just as in Palmer *et al.*’s (2013) study, discussed earlier,

the participants' music–colour matches could once again be predicted on the basis of their colour-emotion and music-emotion ratings. Interestingly, there turned out to be little difference between the performance of the various groups of participants who took part in this particular study. These included a mix of musicians, some of whom possessed absolute pitch, non-musicians, and music-colour synaesthetes — both musicians and non-musicians.

While the crossmodal matching results documented so far are undoubtedly interesting, it should be acknowledged that the selection of short musical excerpts from pre-composed music, be it classical or any other genre is, in some sense, unconstrained/uncontrolled in terms of stimulus selection. That is, it is difficult to isolate and vary the specific sound parameter that one is interested in without inadvertently also varying some others at the same time. For instance, Crawshaw (2012) illustrates this potential concern by pointing to the fact that Queen's Bohemian Rhapsody changes from major to minor mode (https://en.wikipedia.org/wiki/Bohemian_Rhapsody), while Mozart's Piano Sonata No. 12 in F. Major (K332) also changes quite substantially.

In order to address any such potential concern, Palmer *et al.* (2016) presented their participants with 64 precisely controlled single-line melodies instead. The Mozart pieces used in this study were systematically manipulated in terms of their tempo (fast/slow), note density (sparse/dense), mode (major/minor), and pitch height (low/high). Nevertheless, these researchers were able to demonstrate much the same pattern of emotionally-mediated crossmodal matches (to one of the 37 Berkeley Colour Project colour patches), as documented in previous research by the same group. So, for instance, Palmer *et al.* (2016, p. 157) report that: “*The cross-modal choices showed that faster music in the major mode was associated with lighter, more saturated, yellower (warmer) colors than slower music in the minor mode.*” ... “*the emotional ratings of the melodies were very highly correlated with the emotional associations of the colors chosen as going best/worst with the melodies ($r = 0.92, 0.85, 0.82$ and 0.70 for happy/sad, strong/weak, angry/not-angry and agitated/calm, respectively).*” Intriguingly, Palmer *et al.* (2016, pp. 184–185) also note that additional unpublished data from their group support the view that aesthetic preference does not underpin the colour–music matching responses made by their participants.

The musical stimuli used in pretty much all of the crossmodal correspondences research that has been published to date have been professionally composed in order, one presumes, to elicit some sort of emotional response in the listener. The potential downside when using music created specifically for experimental purposes is presumably that it may well lack the emotional impact of such professionally composed pieces. Note here only how people

seem to be remarkably good at determining the emotional meaning of professionally composed musical compositions (e.g., see Rigg, 1937). As such, it is perhaps not so surprising to find that the emotional mediation account should prove to have so much explanatory validity, at least when compared to its role in explaining the crossmodal correspondences that have been documented between the presumably much less emotionally-valenced simple auditory and visual stimuli used in much of the research in this area. That being said, emotional mediation, what is sometimes termed affective correspondence, has also been reported to account for a significant proportion of the variance by those researchers working with a range of simpler sensory stimuli too (e.g., Collier, 1996; Collier and Hubbard, 2001; Levitan *et al.*, 2015; Schifferstein and Tanudjaja, 2004; Seo *et al.*, 2010; Wang *et al.*, 2016). Here, of course, it should be borne in mind that the fact that emotional mediation does such a good job of accounting for music–colour matches certainly does not mean that it is the only factor underlying crossmodal correspondences (see also Barbieri *et al.*, 2007, on this theme), especially for those correspondences that involve less emotionally valenced stimuli. In particular, some part of the variance in people's crossmodal matches might be explained by statistical correspondences (just think of how heavy metal so often tends to be associated with black iconography), and/or perhaps semantic correspondences (e.g., see Holm *et al.*, 2009; cf. Sievers *et al.*, 2017).

The emotional mediation account of crossmodal correspondences has undoubtedly become an increasingly common theme, or explanatory approach, in the literature in this area in recent years (e.g., see Bhattacharya and Lindsen, 2016). However, that said, one thing that has often remained unclear on the basis of much of the previous research is whether emotional mediation is based on the emotion experienced in response to the music, or rather, on a more cognitive assessment of the emotion that one might be tempted to associate with the musical stimuli themselves (Gabrielsson, 2001). When discussing this issue, Marin and Bhattacharya (2010, p. 1) helpfully distinguish between what they call *emotion perception* and *emotion induction*. Others, meanwhile, question whether we should really be talking about *affect* rather than *emotion* (Cespedes-Guevara and Eerola, 2018). Indeed, it has to be said, and it is, in fact, often commented upon, that emotion remains poorly defined. What is more, one often finds differing definitions (or exemplars) being used in different studies. That being said, and having now summarized the empirical literature on the matching of colour to music, the next section takes a closer look at the matching of paintings with music. Paintings, note, constituting a more complex and semantically rich, visual stimulus than simple colour patches.

2.2. Music–Painting Correspondences

In an early series of studies reported by John T. Cowles (1935), separate groups of participants were shown as many as eight landscapes and scenes with simple content and listened to as many as eight pieces of instrumental classical music. The paintings were by various well-known artists while the music consisted of well-known compositions that included Holst's Mars from *The Planets* as well as various pieces from Beethoven and Wagner (Note 5). As the music was played over a phonograph in Cowles' study, the pictures were shown one after the other. The various groups of participants were told to pick the painting that corresponded with each of the pieces of music.

The results revealed partial agreement amongst the participants in terms of the pictures that were chosen to correspond with each of the scenes. In fact, the same picture was chosen for one piece of music by as many as 86% of the participants in one study. Equally interesting, however, was the observation that there were certain of the pictures that were never chosen as matching a particular piece of music [14 out of 49 possible combinations of 7 pictures \times 7 pieces of music in one of Cowles' (1935) studies with 14 musical participants]. Similar results were obtained no matter whether the participants were encouraged to match the media on the basis of their affective mood, or when instead left to match based on their own introspection being told to 'select the picture which best fits the music' (Cowles, 1935, p. 467). Cowles (1935, p. 461) summarized these results thus: *"Among the combinations most frequently chosen, pictures with represented content capable of motor activity were selected with musical selections of prominent dynamic changes; and likewise, pictures of slight content were selected with music of weak dynamic qualities. The introspective reports suggested this relationship. The less dynamic music and the pictures with less express content were most often said to correspond on a basis of mood or abstract elements. Formal elements of the pictures were rarely noted. Rhythm, tempo, and changes in loudness were most frequently noted. There was no mention of hedonic response. No substantial difference was found either in the choices or reports of musical and unmusical observers."*

More recently, Albertazzi *et al.* (2015) had 63 Italian students associate 15 pieces of Spanish guitar music with 15 highly-saturated materic (described as *"painting realized with a great quantity of pictorial material, and characterized by a thick and tendentially 3D pictorial surface"* (Albertazzi *et al.*, 2015, p. 3) and expressionist paintings by Matteo Boato, a relatively unknown painter. Twenty-two unipolar semantic scales, assessing the following adjectives (slow, quick, agitated, calm, happy, sad, warm, cold, heavy, light, continuous, rhythmic, strong, weak, dark, bright, hard, soft, impression of horizontality, impression of verticality, adagio, and presto) were used. Initially, the participants saw each image for 10 s and listened to each piece of

music for 60 s. The stimuli, presented in a random order, were rated on the semantic scales. Thereafter, the participants listened to each piece of music and picked, in ranked order, up to three paintings that they associated with the music. Note that all 15 paintings were shown simultaneously.

Despite the fact that the auditory and visual stimuli were unfamiliar to the participants, the results revealed a significantly non-random pattern of matching responses. In particular, while certain of the paintings were never chosen as matching a particular piece of music, others were chosen by more than 30% of the participants. Overall, 21 of the music–painting combinations were chosen significantly more often than would be expected by chance, while a further 18 were picked significantly less often (from a total of 225 possible matches). The crossmodal matching results were consistent with a matching of meaning as assessed by the semantic differential. In particular, the attributes that played the most important role in terms of crossmodal association being: ‘quick’, ‘agitated’, ‘strong’, and their antonyms ‘slow’, ‘calm’, and ‘weak’. Albertazzi *et al.* (2015, p. 11) explained their results in terms of: “*an association due to patterns of qualitative similarity present in stimuli of different sensory modalities.*” The authors further suggest that timbre and tempo were the dominant attributes of the music, finding little role for mode (i.e., major/minor). What is more, similar results were obtained in music experts, painting experts, and non-experts.

Meanwhile, a separate line of empirical research that is potentially relevant here involves the assessment of people’s sensitivity to so-called cross-media artistic styles (see Hasenfus *et al.*, 1983; see also Chmiel and Schubert, 2019; Lessing, 1766/1962; though see Wellek and Warren, 1948). So, for example, Hasenfus *et al.* presented aesthetically naïve observers with a range of reproductions of unfamiliar paintings, examples of architecture, poetry, and music (15 in each category) from several different historical epochs/styles. These included baroque, neoclassical, and romantic (Ruth and Kolehmainen, 1974). Across a series of experiments, the participants were instructed to sort the stimuli as they saw fit in what is known as a free-sorting task (i.e., grouping the stimuli based on their similarity). Intriguingly, the participants were significantly more likely to group the stimuli from the same stylistic period together, seemingly regardless of media format, than expected by chance.

Interestingly, when the participants in Hasenfus *et al.*’s (1983) Experiment 1 were asked on what basis they made their sorting judgments, nine out of 16 mentioned the emotion, or feeling, evoked, 13/16 mentioned content (either real, or, in the case of music, imagined), and 9/16 mentioned stylistic variables (e.g., complex, flowery, chaotic, pretty). A similar pattern of results was obtained in other experiments in which the participants were instructed to classify works of painting, poetry, architecture, and music selected from the years 1600–1840, arranged into six successive 40-year epochs (i.e., period styles).

Hasenpus *et al.* suggested that the bases for the sensitivity to cross-media and period styles were the dimensions of realistic versus unrealistic and the arousal potential of the stimuli. Once again, therefore, emotional mediation may play a significant role in people's matching behaviour.

However, impressive though such results are, it turns out that they can sometimes potentially be explained more parsimoniously on the basis of basic sensory/perceptual correspondences, rather than necessarily needing to postulate the extraction of common underlying stylistic qualities in the cross-media stimuli used (see Duthie, 2013; Duthie and Duthie, 2015), this presumably being more likely in those cases where the participants are unfamiliar with the complex stimuli that they happen to be rating. What is also perhaps worth noting here is how the same semantic descriptors are sometimes classed by researchers as an emotion, while, at other times, they may be classed as a stylistic quality instead. Such a lack of clarity hinting at the uncertainty surrounding the definition of basic emotions, a point to which I will return at the end of this review.

Meanwhile, in her Dissertation work, Amanda Catherine Duthie (2013; see also Duthie and Duthie, 2015), assessed, and then analysed, the fundamental sensory properties/characteristics of 80 French and 80 Russian oil paintings from a 50-year period (1870–1920) and music (77 French; 76 Russian) written for piano, organ, harmonium, choir (a cappella), brass, and string quartet. Extraction of the mean and standard deviation of the stimulus properties revealed that the French oil paintings were significantly lighter (greyscale lightness, not the saturation of the chroma) than the Russian works. By contrast, there was no significant difference between the Russian and French musical selections in terms of their fundamental pitch. Such results hint at the possibility that cross-modal correspondences between music and painting might sometimes rely, at least in part, on fundamental correspondences/sensory differences rather than necessarily providing support for the notion of a sensitivity exclusively to higher-level cross-media artistic styles. That said, further empirical data is needed in order to further resolve this issue.

The results reported in this section, while undoubtedly fewer in number than for the music–colour matching documented in the preceding section, nevertheless still do highlight how crossmodal correspondences may be established between complex auditory stimuli (music) and complex visual stimuli (paintings, etc.). However, it is important to remember when using such rich and meaningful stimuli that it can sometimes be difficult to determine the fundamental basis on which such correspondences are established. While some researchers have wanted to support the controversial notion of cross-media artistic styles (e.g., Hasenpus *et al.*, 1983), others have suggested instead that a more 'primitive' structural (in the sense of the structure of the stimuli themselves), or sensory correspondence might, in some cases at least, underlie the

results obtained (see Duthie, 2013; Duthie and Duthie, 2015; Lakens *et al.*, 2013). Future research would presumably be able to eliminate these low-level sensory alternative accounts, if the statistical image properties (e.g., in terms of lightness, colour properties) were to be equated by modifying the images in some way.

2.3. Addressing the File Drawer Problem

Before closing, it is worth noting that when reviewing the evidence in any area, there is a danger of the oft-observed publication bias colouring the results. In particular, those studies documenting significant results are more likely to be published, and hence make it into reviews like this, than those studies where null results were obtained. This is the so-called ‘file drawer’ problem (Rosenthal, 1979). However, beyond acknowledging that such a bias may exist in this area of scientific endeavour, as in any other, the author can do little more than report the he is not aware of any null results. And given the fact that research in this area stretches back nearly a century, many of the prominent early researchers are now presumably either retired or dead, thus making correspondence decidedly difficult. At the same time, however, one thing that is perhaps worth noting in this context is that given the sheer number of statistical comparisons typically made in the studies in this area, it is presumably less likely that a study generates no significant results whatsoever. Instead, a more salient concern in the research in this area, in particular, is whether the appropriate statistical tests have been conducted, and whether any of the assumptions underlying those tests were violated (see Dreksler and Spence (2018) for a detailed discussion of this issue in the context of assessing Kandinsky’s colour–shape correspondences).

3. Conclusions

In this review, the literature on those crossmodal correspondences involving complex auditory stimuli has been critically evaluated. As it turns out, research in this area has nearly always used short pre-corded musical clips as stimuli, mostly taken from the classical music repertoire, the sole exceptions here being Whiteford *et al.*’s (2018) cross-musical-genre study, and Lindborg and Friberg’s (2015) study using film music excerpts. As has hopefully become clear, robust crossmodal correspondences have been demonstrated between music and both colour patches and paintings. Furthermore, according to the results of a growing number of such studies, these crossmodal connections would seem to be largely mediated by ‘emotion’, however the latter is defined, or exemplified. Problematically though, as was noted earlier, definitions of emotion seems to differ from one study to the next. Nevertheless, the suggestion made here is that emotional mediation may come to account

for an increasing proportion of the variance in people's matching responses as the emotional valence of the stimuli themselves increases. However, while plausible-sounding, this suggestion awaits robust empirical support.

3.1. Do We Need the Semantic Differential and/or the Emotional Mediation Accounts?

One of the outstanding issues in this area concerns the relative explanatory power of the semantic differential theory of matching versus that provided by the emotional mediation account. One might provocatively ask whether emotional mediation is, in fact, a necessary construct here, or whether instead the semantic differential technique can explain why it is that people tend to match complex stimuli (such as music and paintings) in quite the way that they do, the suggestion being that perhaps people match those stimuli that share their valence, arousal, and/or dominance (Osgood *et al.*, 1957; cf. Walker *et al.*, 2012). Alternatively, however, and this is the point made by Palmer *et al.* (2016) in response to just such a suggestion, the emotional effects that were documented in their study turned out to be highly specific to particular 'real' emotions (e.g., happiness, sadness, anger, agitation, and calmness) rather than to the generalized affective dimensions that have been postulated to underlie all emotions: valence (*good/bad* or *pleasant/unpleasant*) and potency (*strong/weak* or *active/passive*). Palmer and his colleagues also make the point that almost everything can presumably be mapped onto valence and potency (e.g., Mehrabian and Russell, 1974; Osgood *et al.*, 1957), but that not everything maps onto specific emotions.

Potentially relevant here, Marin *et al.* (2012) reported that while the arousal potential associated with listening to music (romantic piano solo music) transfers to ratings of subsequently presented visual stimuli (IAPS pictures), pleasantness does not. Meanwhile, Whiteford *et al.* (2018) found that the valence (i.e., aesthetic value) associated with pieces of music (selected from a wide range of styles) and with colours fails to help predict the crossmodal correspondences that they observed in their recent study (see also Palmer *et al.*, 2016, for a similar observation). And, going back to one of the very first studies in this area, the participants tested in Cowles' (1935) study also failed to mention their hedonic response when justifying their choices of music played to match to colour reproductions of paintings. All-in-all, then, at the present time, the evidence would seem to favour the emotional mediation account in terms of doing a better job of explaining certain crossmodal matches than the semantic differential account. That said, however, more research is undoubtedly needed in order to determine how exactly these two accounts of crossmodal correspondences involving complex musical stimuli differ, especially on the assumption that the emotion is associated with the stimuli themselves rather than being induced in the observer.

At the same time, one does need to think reasonably carefully about whether to accept all of the exemplars of emotions that one finds mentioned in the various studies that have been published in this area. For instance, just take the following so-called ‘emotional’ dimensions from Whiteford *et al.*’s (2018) study to get a sense of the broad use of the term ‘emotion’ by researchers nowadays: complex vs simple; harmonious vs disharmonious; loud vs quiet; spicy vs bland; warm vs cool; whimsical vs serious). Accepting such a seemingly unconstrained collection of emotional terms risks reducing the explanatory validity of the emotional account. Ultimately, it is worth bearing in mind that both theoretical accounts may have some explanatory validity, even beyond whatever they may share in common. Note here, for instance, that associations with happy and sad might appear both in the emotional mediation account, and as a part of the core affective dimension of the semantic differential approach. Remember also that using multidimensional scaling, Palmer *et al.* (2013) found that a 2D solution (consistent with *valence* and *potency* dimensions, i.e., two of the three key dimensions according to the semantic differential approach) was capable of accounting for 95% of the variance. Similarly, Whiteford *et al.* (2018) reported that *arousal* and *valence* were key emotional attributes mediating the crossmodal association, or correspondence, between music and colour (see also Fig. 1).

3.2. *Bidirectional Crossmodal Correspondences/Influences*

It is interesting to note that the majority of the crossmodal matching studies that have been published to date have had the participants pick the best/worst matching pictures to go with a given piece of music. As far as I am aware, there have been no studies looking for crossmodal matches going in the opposite direction. This presumably reflects nothing more than the practical constraint that musical stimuli are time-varying and cannot easily be presented simultaneously in the way that colour patches or pictures can. This should probably not matter too much, at least not if the claim from various researchers that crossmodal correspondences are bidirectional is to be believed (see Deroy and Spence, 2013, for a review; see also Parise, 2016; Walker *et al.*, 2012). However, that said, it is probably prudent to assess whether the crossmodal correspondences in this area (i.e., involving musical stimuli) are as bidirectional as has been suggested by researchers typically theorizing about correspondences between simple stimuli (see also Parrott, 1982). Then, beyond any crossmodal correspondences that might be observed, it will be interesting in future research to determine whether such crossmodal correspondences also give rise to any crossmodal influences or not (see Spence, 2019b, in press, on this theme).

3.3. Individual Differences in Crossmodal Correspondences Involving Complex Stimuli

Looking to the future, it will be interesting to establish the cross-cultural generalizability of the crossmodal correspondences involving complex stimuli that have been documented thus far (see Adams and Osgood, 1973; Bremner *et al.*, 2013; Knoeferle *et al.*, 2015; Osgood, 1960; Walker, 1987, for prior cross-cultural studies of the crossmodal correspondences) (Note 6). While cultural differences have, on occasion, been documented in the case of sound symbolism (e.g., Parkinson *et al.*, 2012; Styles and Gawne, 2017), speech stimuli may be importantly different from non-linguistic stimuli in this regard. That said, Bremner *et al.* (2013) did find some striking cross-cultural differences in the correspondences between shape and taste/oral-somatosensory food properties. Nevertheless, the majority of the research has been conducted on a very limited range of people (see Henrich *et al.*, 2010). Will the same be true in the case of musical stimuli?

In some of the few studies so far to have been published in this space, Knoeferle *et al.* (2015) demonstrated that short musical excerpts that had been especially composed in order to convey each of the basic tastes, were more-or-less equally often matched with the intended taste in a group of participants from North America as in a group from the Indian sub-continent, the latter participants coming from a part of the world with a very different musical repertoire. Potentially relevant here, with only a few exceptions (e.g., Lindborg and Friberg, 2015; Whiteford *et al.*, 2018), it is noticeable how the majority of the research in this area to date has focused solely on instrumental classical music. Given how broad a range of complex auditory stimuli there is, it would be nice, in future research, to explore the nature of any crossmodal associations with a wider range of both vocal and non-vocal auditory stimuli. One of the few studies to have been conducted along such lines, Holm *et al.* (2009) assessed colour associations (from 12 colour patches) with 18 specific musical genres, presented as a list of 18 word labels, in an online study of 104 participants. The colour black was found to be associated with ‘metal rock’, blue with ‘blues’, pink with ‘pop’, etc.

It will perhaps also be interesting in future research to follow-up on Moon *et al.*’s (2014) suggestions that people’s musical preferences can influence their mood associations with music too. It should, though, be noted that were any such individual differences to be obtained, they would stand in contrast with the repeated failure to find any such differences when the crossmodal correspondences involving short musical excerpts have been tested with different groups of individuals. These have included groups of art experts, music experts/musicians vs non-experts, music–colour synaesthetes vs non-synaesthetes, and musicians with absolute pitch (e.g., Albertazzi *et al.*, 2015;

Cowles, 1935; Isbilen and Krumhansl, 2016; see also Lehman, 1972). Indeed, one might have imagined that musical training would have been another relevant individual difference that might affect crossmodal matching in this area (i.e., for those correspondences involving auditory stimuli). However, while this factor has been incorporated in a few of the studies reviewed here (e.g., see Albertazzi *et al.*, 2015; Isbilen and Krumhansl, 2016), it has not proved to be especially relevant.

3.4. Closing Comments

In closing, it is worth noting that the emotional mediation account of music's link with visual stimuli, be they simple (i.e., colour patches) or complex (e.g., paintings) means that any granular crossmodal correspondence that has proved so appealing to artists (not to mention those working in the crossmodal translation of sensation; Haverkamp, 2014) may override the low-level crossmodal correspondence of loudness with brightness, perceived location, and timbre with shape (or colour) that so many have searched for (e.g., see Abbado, 1988; Adeli *et al.*, 2014; Barbieri *et al.*, 2007; Collopy, 2000; Galeyev, 2003, 2007; Gerstner, 1986; Margounakis and Politis, 2006; Tan *et al.*, 2013).

The majority of the visual stimuli that have been studied by researchers working on crossmodal correspondences with music to date have been static, while auditory stimuli are inherently time-varying. It is interesting to ask, therefore, what would happen in the case of crossmodal correspondences between music and dynamic visual stimuli. A priori, though, it is easy to imagine how additional correspondences might be uncovered under such conditions. At the same time, however, the synchrony/temporal correlation of unisensory inputs might prove to be a dominant cue influencing the matching of stimuli across the senses in this case (see Parise *et al.*, 2012). Due to space constraints, though, we will have to leave this intriguing topic for future research. Another area that has not been covered by the present review concerns the role that emotional mediation plays in explaining the crossmodal correspondences that have been documented between musical stimuli and olfactory (see Deroy *et al.*, 2013) or flavourful stimuli, as in the case of wine-music matching (e.g., Burzynska *et al.*, 2019; Spence *et al.*, 2013; see Spence and Wang, 2015a, b, c, for reviews). Once again, though, given space constraints, the interested reader is directed to the above-mentioned reviews for further discussion of this topic.

Notes

1. The term 'mediation' is one that has been used by a number of researchers working in the field. However, note that the term is not used in the sense of causal mediation (i.e., in terms of predictor, mediator, and outcome variable).

2. Indeed, one could ask a number of different questions concerning a specific piece of music, such as: “How was Beethoven feeling when writing the sonata you just heard?”, “How do you feel when listening to it?”, or “What emotion is Beethoven trying to convey in this sonata?” It is currently an open question as to whether these different questions would index different emotional values, or, under the strong correspondence hypothesis, perhaps they would all result in people giving the same answer.
3. Distinguishing *affect* from *emotion* is undoubtedly difficult (e.g., Cespedes-Guevara and Eerola, 2018). Nevertheless, in terms of operational definitions, definitions of affect talk of ‘touch the feelings of, or to move emotionally’, while emotion has been defined as ‘a strong feeling deriving from one’s circumstances, mood, or relationships with others.’
4. Intriguingly here, it has been suggested that the degree to which a particular stimulus gives rise to differing responses on a range of descriptors across a group of observers may provide an index of the ‘sensory intricacy’ of the stimulus (see Snitz *et al.*, 2016). That said, according to Snitz *et al.*’s research, ‘sensory intricacy’ is a qualitatively different notion than rated, or perceived, complexity. One might, though, wonder where one would get to if one were to replace the sensory descriptors used by Snitz *et al.* with a selection of prototypical semantic differential scales instead.
5. And see Walker (1927, pp. 83, 91–96), for an even earlier study of the crossmodal matching of music and pictures in young children.
6. It is worth noting here that if there were to be non-linear changes between the colour/emotion mappings as a function of culture, then some of the linear relationships between colour space and music (if mediated through emotion) might be expected to break down.

References

- Abbado, A. (1988). Perceptual correspondences of abstract animation and synthetic sound, *Leonardo* **1**(*Electronic Art Supplemental Issue*), 3–5.
- Adams, F. M. and Osgood, C. E. (1973). A cross-cultural study of the affective meanings of color, *J. Cross-Cult. Psychol.* **4**, 135–156.
- Adeli, M., Rouat, J. and Molotchnikoff, S. (2014). Audiovisual correspondence between musical timbre and visual shapes, *Front. Hum. Neurosci.* **8**, 352. DOI:10.3389/fnhum.2014.00352.
- Albertazzi, L., Canal, L. and Micciolo, R. (2015). Cross-modal association between materic painting and classical Spanish music, *Front. Psychol.* **6**, 424. DOI:10.3389/fpsyg.2015.00424.
- Anikin, A. and Johansson, N. (2019). Implicit associations between individual properties of color and sound, *Atten. Percept. Psychophys.* **81**, 764–777.

- Arnheim, R. (1986). *New Essays on the Psychology of Art*. University of California Press, Berkeley, CA, USA.
- Barbieri, J. M., Vidal, A. and Zellner, D. A. (2007). The color of music: correspondence through emotion, *Empir. Stud. Arts* **25**, 193–208.
- Bhattacharya, J. and Lindsen, J. P. (2016). Music for a brighter world: brightness judgment bias by musical emotion, *PLoS One* **11**, e0148959. DOI:10.1371/journal.pone.0148959.
- Bozzi, P. and Flores D'Arcais, G. (1967). Ricerca sperimentale sui rapporti intermodali fra qualità espressive [Experimental research on the intermodal relationships between expressive qualities], *Arch. Psicol. Neurol. Psichiatr.* **28**, 377–420.
- Bremner, A. J., Caparos, S., Davidoff, J., de Fockert, J., Linnell, K. J. and Spence, C. (2013). “Bouba” and “Kiki” in Namibia? A remote culture make similar shape–sound matches, but different shape–taste matches to Westerners, *Cognition* **126**, 165–172.
- Bresin, R. (2005). What is the color of that music performance?, in: *Proceedings of the International Computer Music Conference*, pp. 367–370. Barcelona, Spain.
- Brunetti, R., Indraccolo, A., Del Gatto, C., Spence, C. and Santangelo, V. (2018). Are cross-modal correspondences relative or absolute? Sequential effects on speeded classification, *Atten. Percept. Psychophys.* **80**, 527–534.
- Burzynska, J., Wang, Q. J., Spence, C. and Bastian, S. E. P. (2019). Taste the bass: low frequencies increase the perception of body and aromatic intensity in red wine, *Multisens. Res.* **32**, 429–454.
- Caivano, J. L. (1994). Color and sound: physical and psychophysical relations, *Color Res. Appl.* **19**, 126–133.
- Cespedes-Guevara, J. and Eerola, T. (2018). Music communicates affects, not basic emotions — a constructionist account of attribution of emotional meanings to music, *Front. Psychol.* **9**, 215. DOI:10.3389/fpsyg.2018.00215.
- Chmiel, A. and Schubert, E. (2019). Psycho-historical contextualization for music and visual works: a literature review and comparison between artistic mediums, *Front. Psychol.* **10**, 182. DOI:10.3389/fpsyg.2019.00182.
- Collier, G. L. (1996). Affective synaesthesia: extracting emotion space from simple perceptual stimuli, *Motiv. Emot.* **20**, 1–32.
- Collier, W. G. and Hubbard, T. L. (2001). Judgements of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo, *Psychomusicology* **17**, 36–55.
- Collopy, F. (2000). Colour, form, and motion: dimensions of a musical art of light, *Leonardo* **33**, 355–360.
- Cowles, J. T. (1935). An experimental study of the pairing of certain auditory and visual stimuli, *J. Exp. Psychol.* **18**, 461–469.
- Crawshaw, A. (2012). *How Musical Emotion May Provide Clues for Understanding the Observed Impact of Music on Gustatory and Olfactory Perception in the Context of Wine-Tasting*, Unpublished manuscript.
- Crisinel, A.-S. and Spence, C. (2010). As bitter as a trombone: synesthetic correspondences in non-synesthetes between tastes/flavors and musical instruments and notes, *Atten. Percept. Psychophys.* **72**, 1994–2002. DOI:10.3758/APP.72.7.1994.
- Cutiitta, R. A. and Haggerty, K. J. (1987). A comparative study of colour association with music at various age levels, *J. Res. Music Educ.* **35**, 78–91.
- Davis, J. W. (1979). A response to W. Garner's observation on the relationship between colour and music, *Leonardo* **12**, 218–219.

- Deroy, O. and Spence, C. (2013). Weakening the case for ‘weak synaesthesia’: why crossmodal correspondences are not synaesthetic, *Psychon Bull. Rev.* **20**, 643–664.
- Deroy, O., Crisinel, A.-S. and Spence, C. (2013). Crossmodal correspondences between odors and contingent features: odors, musical notes, and geometrical shapes, *Psychon. Bull. Rev.* **20**, 878–896.
- Deroy, O., Fernandez-Prieto, I., Navarra, J. and Spence, C. (2018). Unravelling the paradox of spatial pitch, in: *Spatial Biases in Perception and Cognition*, T. L. Hubbard (Ed.), pp. 77–93. Cambridge University Press, Cambridge, UK.
- Dreksler, N. and Spence, C. (2018). *Colour–shape correspondences: a historical primer and Empirical Literature Review*, Unpublished manuscript.
- Duthie, A. C. (2013). *Do Music and Art Influence one another? Measuring Cross-Modal Similarities in Music and Art*, Graduate Theses and Dissertations. Paper 13163, Iowa State University, Ames, IA, USA.
- Duthie, A. C. and Duthie, A. B. (2015). Do music and art influence one another? Measuring cross-modal similarities in music and art, *Polymath Interdiscip. Arts Sci. J.* **5**, 1–22.
- Eitan, Z. (2017). Musical connections: cross-modal connections, in: *The Routledge Companion to Music Cognition*, R. Ashley and R. Timmers (Eds), pp. 213–224. Taylor and Francis, London, UK.
- Evans, K. K. and Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features, *J. Vis.* **10**, 6. DOI:10.1167/10.1.6.
- Gabrielsson, A. (2001). Emotion perceived and emotion felt: same or different?, *Mus. Sci.* **5**, 123–147.
- Galeyev, B. M. (2003). Evolution of gravitational synesthesia in music: to color and light! *Leonardo* **36**, 129–134.
- Galeyev, B. M. (2007). The nature and functions of synesthesia in music, *Leonardo* **40**, 285–288.
- Gallace, A. and Spence, C. (2006). Multisensory synesthetic interactions in the speeded classification of visual size, *Percept. Psychophys.* **68**, 1191–1203.
- Garner, W. (1978). The relationship between colour and music, *Leonardo* **11**, 225–226.
- Gerstner, K. (1986). *The Forms of Color: the Interaction of Visual Elements*. MIT Press, Cambridge, MA, USA.
- Gilbert, A. N., Fridlund, A. J. and Lucchina, L. A. (2016). The color of emotion: a metric for implicit color associations, *Food Qual. Pref.* **52**, 203–210.
- Hasenpus, N., Martindale, C. and Birnbaum, D. (1983). Psychological reality of cross-media artistic styles, *J. Exp. Psychol. Hum. Percept. Perform.* **9**, 841–863.
- Haverkamp, M. (2014). *Synesthetic Design: Handbook for a Multisensory Approach*. Birkhäuser, Basel, Switzerland.
- Henrich, J., Heine, S. J. and Norenzayan, A. (2010). The weirdest people in the world?, *Behav. Brain Sci.* **33**, 61–83.
- Hevner, K. (1935). Experimental studies of the affective value of colours and lines, *J. Appl. Psychol.* **19**, 385–398.
- Holm, J., Aaltonen, A. and Siirtola, H. (2009). Associating colours with musical genres, *J. New Music Res.* **38**(1), 87–100.
- Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval, *Am. J. Psychol.* **109**, 219–238.

- Ione, A. and Tyler, C. (2004). Neuroscience, history and the arts synesthesia: is F-sharp colored violet?, *J. Hist. Neurosci.* **13**, 58–65.
- Isbilen, E. S. and Krumhansl, C. L. (2016). The color of music: emotion-mediated associations to Bach's well-tempered clavier, *Psychomusicol. Music Mind Brain* **26**, 149–161.
- Janković, D. (2014). Cross-modal nature of evaluative meaning, in: *Different Psychological Perspectives on Cognitive Processes: Current Research Trends in Alps-Adria Region*, A. Galmonete and R. Actis-Grosso (Eds), pp. 58–75. Cambridge Scholars Publishing, Newcastle-upon-Tyne, UK.
- Johnson, S. G. B. and Steinerberger, S. (2019). Intuitions about mathematical beauty: A case study in the aesthetic experience of ideas, *Cognition* **189**, 242–259.
- Juslin, P. N. and Sloboda, J. A. (2010). *Handbook of Music and Emotion: Theory, Research, Applications*. Oxford University Press, Oxford, UK.
- Juslin, P. N. and Västfjäll, D. (2008). Emotional responses to music: the need to consider underlying mechanisms, *Behav. Brain Sci.* **31**, 559–575.
- Kandinsky, W. (1977). *Concerning the Spiritual in Art, Especially in Painting* (transl. M. T. H. Sadler). Dover Publications, New York, NY, USA. (Original work published 1914.)
- Kandinsky, W. (1979). *Point and Line to Plane* (transl. H. Dearstyne and H. Rebay). Dover Publications, New York, NY, USA. (Original work published 1925.)
- Karwoski, T. F. and Odbert, H. S. (1938). Color–music, *Psychol. Monogr.* **50**, 1–60.
- Karwoski, T. F., Odbert, H. S. and Osgood, C. E. (1942). Studies in synesthetic thinking. II. The rôle of form in visual responses to music, *J. Gen. Psychol.* **26**, 199–222.
- Kawakami, A., Furukawa, K., Katahira, K. and Okanoya, K. (2013). Sad music induces pleasant emotion, *Front. Psychol.* **4**, 311. DOI:10.3389/fpsyg.2013.00311.
- Knoeflerle, K. M., Woods, A., K  ppler, F. and Spence, C. (2015). That sounds sweet: using crossmodal correspondences to communicate gustatory attributes, *Psychol. Mark.* **32**, 107–120.
- Lakens, D., Fockenberg, D., Lemmens, K., Ham, J. and Midden, C. (2013). Brightness differences influence the evaluation of affective pictures, *Cogn. Emot.* **27**, 1225–1246.
- Lehman, R. S. (1972). A multivariate model of synesthesia, *Multivar. Behav. Res.* **7**, 403–439.
- Lessing, G. E. (1962). *Laoco  n: an Essay on the Limits of Painting and Poetry*. Bobbs-Merrill, Indianapolis, IN, USA. (Originally published, 1766.)
- Levitan, C. A., Charney, S. A., Schloss, K. B. and Palmer, S. E. (2015). The smell of jazz: crossmodal correspondences between music, odor, and emotion, in: *Cogsci 2015 Proceedings*. Pasadena, CA, USA, pp. 1326–1331.
- Lindborg, P. M. and Friberg, A. K. (2015). Colour association with music is mediated by emotion: evidence from an experiment using a CIE Lab interface and interviews, *PLoS One* **10**, e0144013. DOI:10.1371/journal.pone.0144013.
- Logeswaran, N. and Bhattacharya, J. (2009). Crossmodal transfer of emotion by music, *Neurosci. Lett.* **455**, 129–133.
- Lundholm, H. (1921). The affective tone of lines: experimental researches, *Psychol. Rev.* **28**, 43–60.
- Margounakis, D. and Politis, D. (2006). Converting images to music using their colour properties, in: *Proceedings of the 12th International Conference on Auditory Display*, pp. 198–205, London, UK, June 20–23.

- Marin, M. M. and Bhattacharya, J. (2010). Music induced emotions: some current issues and cross-modal comparisons, in: *Music Education*, J. Hermida and M. Ferrero (Eds), pp. 1–38. Nova Science Publishers, Hauppauge, NY, USA.
- Marin, M. M., Gingras, B. and Bhattacharya, J. (2012). Crossmodal transfer of arousal, but not pleasantness, from the musical to the visual domain, *Emotion* **12**, 618–631.
- Marin, M. M., Schober, R., Gingras, B. and Leder, H. (2017). Misattribution of musical arousal increases sexual attraction towards opposite-sex faces in females, *PLoS One* **12**, e0183531. DOI:10.1371/journal.pone.0183531.
- Marks, L. E. (1974). On associations of light and sound: the mediation of brightness, pitch, and loudness, *Am. J. Psychol.* **87**, 173–188.
- Marks, L. E. (1975). On colored-hearing synesthesia: cross-modal translations of sensory dimensions, *Psychol. Bull.* **82**, 303–331.
- Marks, L. E. (1987). On cross-modal similarity: auditory–visual interactions in speeded discrimination, *J. Exp. Psychol. Hum. Percept. Perform.* **13**, 384–394.
- Marks, L. E. (2004). Cross-modal interactions in speeded classification, in: *The Handbook of Multisensory Processes*, G. A. Calvert, C. Spence and B. E. Stein (Eds), pp. 85–105. MIT Press, Cambridge, MA, USA.
- Mas-Herrero, E., Zatorre, R. J., Rodriguez-Fornells, A. and Marco-Pallarés, J. (2014). Dissociation between musical and monetary reward responses in specific musical anhedonia, *Curr. Biol.* **24**, 699–704.
- Mehrabian, A. R. and Russell, J. A. (1974). *An Approach to Environmental Psychology*. MIT Press, Cambridge, MA, USA.
- Moon, C. B., Kim, H., Lee, H. A. and Kim, B. M. (2014). Analysis of relations between mood and color for different musical preferences, *Color Res. Appl.* **39**, 413–423.
- Murari, M., Rodà, A., Da Pos, O., Canazza, S., De Poli, G. and Sandri, M. (2014). How blue is Mozart? Nonverbal sensory scales for describing music qualities, in: *Proceedings of the Combined 40th Verbal and Icon Ratings of Music International Computer Music Conference and 11th Sound and Music Computing Conference*, A. Georgaki and G. Kouroupetroglou (Eds). 14–20 September, Athens, Greece.
- Murari, M., Rodà, A., Canazza, S., De Poli, G. and Da Pos, O. (2015). Is Vivaldi smooth and takete? Nonverbal sensory scales for describing music qualities, *J. New Music Res.* **44**, 359–372.
- Newton, I. (1952). *Opticks or a Treatise on the Reflections, Refractions, Inflections and Colours of Light*. Dover Publications, Mineola, NY, USA. (Original work published 1730.)
- North, A. and Hargreaves, D. (2008). *The Social and Applied Psychology of Music*. Oxford University Press, Oxford, UK.
- Odbert, H. S., Karwoski, T. F. and Eckerson, A. B. (1942). Studies in synesthetic thinking. I. Musical and verbal associations of colour and mood, *J. Gen. Psychol.* **26**, 153–173.
- Osgood, C. E. (1952). The nature and measurement of meaning, *Psychol. Bull.* **49**, 197–237.
- Osgood, C. E. (1960). The cross-cultural generality of visual-verbal synesthetic tendencies, *Behav. Sci.* **5**, 146–169.
- Osgood, C. E. and Suci, G. J. (1955). Factor analysis of meaning, *J. Exp. Psychol.* **50**, 325–338.
- Osgood, C. E., Suci, G. J. and Tannenbaum, P. H. (1957). *The Measurement of Meaning*. University of Illinois Press, Urbana, IL, USA.
- Palmer, S. E., Schloss, K. B., Xu, Z. and Prado-León, L. R. (2013). Music–color associations are mediated by emotion, *Proc. Natl Acad. Sci. USA* **110**, 8836–8841.

- Palmer, S. E., Langlois, T. A. and Schloss, K. B. (2016). Music-to-color associations of single line piano melodies in non-synesthetes, *Multisens. Res.* **29**, 157–193.
- Parise, C. V. (2016). Crossmodal correspondences: standing issues and experimental guidelines, *Multisens. Res.* **29**, 7–28.
- Parise, C. V. and Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: a study using the implicit association test, *Exp. Brain Res.* **220**, 319–333.
- Parise, C. V., Spence, C. and Ernst, M. (2012). When correlation implies causation in multisensory integration, *Curr. Biol.* **22**, 46–49.
- Parkinson, C., Kohler, P. J., Sievers, B. and Wheatley, T. (2012). Associations between auditory pitch and visual elevation do not depend on language: evidence from a remote population, *Perception* **41**, 854–861.
- Parrott, A. C. (1982). Effects of paintings and music, both alone and in combination, on emotional judgment, *Percept. Mot. Skills* **54**, 635–641.
- Poffenberger, A. T. and Barrows, B. E. (1924). The feeling value of lines, *J. Appl. Psychol.* **8**, 187–205.
- Pridmore, R. W. (1992). Music and color: relations in the psychophysical perspective, *Color Res. Appl.* **17**, 57–61.
- Rader, C. M. and Tellegen, A. (1987). An investigation of synesthesia, *J. Pers. Soc. Psychol.* **52**, 981–987.
- Rigg, M. (1937). An experiment to determine how accurately college students can interpret the intended meanings of musical compositions, *J. Exp. Psychol.* **21**, 223–229.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results, *Psychol. Bull.* **86**, 638–641.
- Russell, J. A. (1980). Circumplex model of affect, *J. Pers. Soc. Psychol.* **39**, 1161–1178.
- Ruth, J.-E. and Kolehmainen, K. (1974). Classification of art into style periods: a factor-analytical approach, *Scand. J. Psychol.* **15**, 322–327.
- Saysani, A. (2019). How the blind hear colour, *Perception* **48**, 237–241.
- Schifferstein, H. N. J. and Tanudjaja, I. (2004). Visualizing fragrances through colors: the mediating role of emotions, *Perception* **33**, 1249–1266.
- Sebba, R. (1991). Structural correspondence between music and color, *Color Res. Appl.* **16**, 81–88. DOI:10.1002/col.5080160206.
- Seo, H.-S., Arshamian, A., Schemmer, K., Scheer, I., Sander, T., Ritter, G. and Hummel, T. (2010). Cross-modal integration between odors and abstract symbols, *Neurosci. Lett.* **478**, 175–178.
- Sievers, B., Lee, C., Haslett, W. and Wheatley, T. (2017). *A Multi-Sensory Code for Arousal*, Retrieved from <https://psyarxiv.com/wucs4/>.
- Simpson, R. H., Quinn, M. and Ausubel, D. P. (1956). Synaesthesia in children: association of colors with pure tone frequencies, *J. Gen. Psychol.* **89**, 95–103.
- Smith, L. B. and Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions, *Cogn. Psychol.* **24**, 99–142.
- Snitz, K., Arzi, A., Jacobson, M., Secundo, L., Weissler, K. and Yablonka, A. (2016). A cross modal performance-based measure of sensory stimuli intricacy, *PLoS One* **11**, e0147449. DOI:10.1371/journal.pone.0147449.
- Spence, C. (2011). Crossmodal correspondences: a tutorial review, *Atten. Percept. Psychophys.* **73**, 971–995.

- Spence, C. (2018). Crossmodal correspondences: a synopsis, in: *Senses and Sensation: Critical and Primary Sources, Vol. III*, D. Howes (Ed.), pp. 91–125. Bloomsbury Academic, London, UK. <https://www.bloomsbury.com/uk/senses-and-sensation-9781474274050/>.
- Spence, C. (2019a). On the relative nature of (pitch-based) crossmodal correspondences, *Multisens. Res.* **32**, 235–265.
- Spence, C. (2019b). On the relationship(s) between colour and taste, *Exp. Psychol.* **66**, 99–111.
- Spence, C. (in press). Simple and complex crossmodal correspondences involving audition, *Acoust. Sci. Technol.*
- Spence, C., Richards, L., Kjellin, E., Huhnt, A.-M., Daskal, V., Scheybeler, A., Velasco, C. and Deroy, O. (2013). Looking for crossmodal correspondences between classical music & fine wine, *Flavour* **2**, 29. DOI:10.1186/2044-7248-2-29.
- Spence, C. and Wang, Q. (J.) (2015a). Wine and music (I): on the crossmodal matching of wine and music, *Flavour* **4**, 34. DOI:10.1186/s13411-015-0045-x.
- Spence, C. and Wang, Q. (J.) (2015b). Wine and music (II): can you taste the music? Modulating the experience of wine through music and sound, *Flavour* **4**, 33. DOI:10.1186/s13411-015-0043-z.
- Spence, C. and Wang, Q. (J.) (2015c). Wine and music (III): so what if music influences taste?, *Flavour* **4**, 36. DOI:10.1186/s13411-015-0046-9.
- Spence, C., Wan, X., Woods, A., Velasco, C., Deng, J., Youssef, J. and Deroy, O. (2015). On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes, *Flavour* **4**, 23. DOI:10.1186/s13411-015-0033-1.
- Stevens, S. S. (1957). On the psychophysical law, *Psychol. Rev.* **64**, 153–181.
- Styles, S. J. and Gawne, L. (2017). When does maluma/takete fail? Two key failures and a meta-analysis suggest that phonology and phonotactics matter, *i-Perception* **8**, 1–17. DOI:10.1177/2041669517724807.
- Tan, S.-L., Cohen, A. J., Lipscomb, S. D. and Kendall, R. A. (Eds) (2013). *The Psychology of Music in Multimedia*. Oxford University Press, Oxford, UK.
- Tannenbaum, P. H. (1956). Music background in the judgment of stage and television drama, *Av. Commun. Rev.* **4**, 92–101.
- Valentine, C. W. (1962). *The Experimental Psychology of Beauty*. Routledge, London, UK.
- von Goethe, J. W. (1840). *Theory of Colours* (transl. Charles Locke Eastlake). John Murray, London, UK. (Original work published 1810.)
- Walker, E. (1927). *Das musikalische Erlebnis und seine Entwicklung [The musical experience and its evolution]*. Vandenhoeck & Ruprecht, Gottingen, Germany.
- Walker, L. and Walker, P. (2016). Cross-sensory mapping of feature values in the size–brightness correspondence can be more relative than absolute, *J. Exp. Psychol. Hum. Percept. Perform.* **42**, 138–150.
- Walker, L., Walker, P. and Francis, B. (2012). A common scheme for cross-sensory correspondences across stimulus dimensions, *Perception* **41**, 1186–1192.
- Walker, P. (2012). Cross-sensory correspondences and naïve conceptions of natural phenomena, *Perception* **41**, 620–622.
- Walker, P. (2016). Cross-sensory correspondences: a theoretical framework and their relevance to music, *Psychomusicol. Music Mind Brain* **26**, 103–116.
- Walker, R. (1987). The effects of culture, environment, age, and musical training on choices of visual metaphors for sound, *Percept. Psychophys.* **42**, 491–502.

- Wallmark, Z. (2019). Semantic crosstalk in timbre perception, *Music & Science* **2**, 1–18.
- Wang, Q. J., Wang, S. and Spence, C. (2016). “Turn up the taste”: assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and pitch, *Chem. Senses* **41**, 345–356.
- Ward, J., Huckstep, B. and Tsakanikos, E. (2006). Sound–colour synaesthesia: to what extent does it use cross-modal mechanisms common to us all?, *Cortex* **42**, 264–280.
- Watson, K. B. (1942). The nature and measurement of musical meanings, *Psychol. Monogr.* **54**, 1–43.
- Watt, R. and Quinn, S. (2007). Some robust higher-level percepts for music, *Perception* **36**, 1834–1848.
- Wellek, R. and Warren, A. (1948). *Theory of Literature*. Harcourt, Brace and Company, New York, NY, USA.
- Wells, A. (1980). Music and visual color: a proposed correlation, *Leonardo* **13**, 101–107.
- Whiteford, K. L., Schloss, K. B., Helwig, N. E. and Palmer, S. E. (2018). Color, music, and emotion: Bach to the blues, *i-Perception* **9**, 1–27. DOI:10.1177/2041669518808535.
- Wicker, F. W. (1968). Mapping the intersensory regions of perceptual space, *Am. J. Psychol.* **81**, 178–188.
- Zuckerman, J. V. (1949). *Music In Motion Pictures: review of literature with implications for Instructional Films*, Technical Report SDC 269-7-2, Department of the Navy, Office of Naval Research, Special Devices Center, Port Washington, NY, USA. (Mimeo.)