This is the accepted author manuscript of Motoki, K., Marks, L., & Velasco, C. (2023). Reflections on Cross-Modal Correspondences: Current Understanding and Issues for Future Research. Multisensory Research (published online ahead of print 2023). https://doi.org/10.1163/22134808-bja10114

Reflections on crossmodal correspondences: Current understanding and issues for future research

Kosuke Motoki[1], Lawrence E. Marks[2, 3], & Carlos Velasco[4]

[1]*Department of Management, The University of Tokyo, Tokyo, Japan*

[2] *Sensory Information Processing, John B. Pierce Laboratory, New Haven, CT, USA*

[3]*School of Public Health and Department of Psychology, Yale University, New Haven, CT, USA*

[4] *Centre for Multisensory Marketing, Department of Marketing, BI Norwegian Business School, Oslo, Norway*

JOURNAL: MULTISENSORY RESEARCH.

Correspondence concerning this article should be addressed to Dr. Carlos Velasco, Nydalsveien 37, 0484 Oslo. Tel: +47 464 10 000. E-mail: carlos.velasco@bi.no; Dr. Kosuke Motoki,

7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. E-mail: motoki@e.u-tokyo.ac.jp

Abstract

The past two decades have seen an explosion of research on crossmodal correspondences. Broadly speaking, this term has been used to encompass associations between and among features, dimensions, or attributes across the senses. There has been an increasing interest in this topic amongst researchers from multiple fields (psychology, neuroscience, music, art, environmental design, etc.) and, importantly, an increasing breadth of the topic's scope. Here, this narrative review aims to reflect on what crossmodal correspondences are, where they come from, and what underlies them. We suggest that crossmodal correspondences are usefully conceived as relative associations between different actual or imagined sensory stimuli, many of these correspondences being shared by most people. A taxonomy of correspondences with four major kinds of associations (physiological, semantic, statistical, and affective) characterizes crossmodal correspondences. Sensory dimensions (quantity/quality) and sensory features (lower perceptual/higher cognitive) correspond in crossmodal correspondences. Crossmodal correspondences may be understood (or measured) from two complementary perspectives: the phenomenal view (perceptual experiences of subjective matching) and the behavioural response view (observable patterns of behavioural response to multiple sensory stimuli). Importantly, we reflect on remaining questions and standing issues that need to be addressed in order to develop an explanatory framework for crossmodal correspondences. Future research needs (a) to understand better when (and why) phenomenal and behavioural measures are coincidental and when they are not, and, ideally, (b) to determine whether different kinds of crossmodal correspondence (quantity/quality, lower perceptual/higher cognitive) rely on the same or different mechanisms.

Keywords: Crossmodal correspondences, synaesthetic correspondences

Reflections on crossmodal correspondences

1. Introduction

Is a sweet taste associated with a round or an angular shape? Do loud sounds match better with dim or bright colours? Although these questions may at first seem nonsensical, it turns out that people nevertheless have strong intuitions as to how attributes or properties that pertain to perceptual experiences in different sensory modalities correspond to one other. Moreover, these intuitions are not only often strong but, importantly, people often agree about them to a surprising degree. For instance, most people consistently associate round shapes with sweet tastes (Velasco et al., 2016) and louder sounds with brighter objects (Marks, 1978). These and other, analogous phenomena, involving perceptual associations across most or all sense modalities, are broadly named *crossmodal correspondences*. A growing body of research on crossmodal correspondences has for several decades been expanding by the proverbial leaps and bounds (e.g., Marks, 1978; Spence, 2011, 2020a, 2020b; Parise, 2016; for a statistical summary, see Table 1).

Table 1. Frequency of occurrence in the scientific literature of various terms associated with crossmodal correspondence. The numbers were obtained by searching on Google Scholar for the terms within specific time frames. We chose these specific terms because they are commonly used in the literature to describe crossmodal correspondences.

Search words in ""	1960-1970	1971-1980	1981-1990	1991-2000	2001-2010	2011-2023
Crossmodal correspondences	0	0	2	3	29	2770
Crossmodal associations	0	32	16	65	86	787
Cross-sensory correspondences	0	0	2	1	19	298
Synaesthetic correspondences	4	1	9	9	34	105
Synesthetic correspondences	1	5	12	8	43	515

Note: This table is based on publication details from Google Scholar retrieved on 20 June 2023.

In fact, crossmodal correspondences have been studied for a century, albeit without using the term itself. For instance, early empirical and theoretical work was conducted by Hornbostel (1927), who seems to have coined the expression “unity of the senses“. Hornbostel argued, as did Nafe (1927) around the same time, that “brightness “ is an attribute of experience in many, perhaps all, sense modalities. Similar ideas have also appeared in philosophical discourse, for example, in the work of Hartshorne (1934). But what were these early investigators referring to, and what do contemporary researchers refer to, when they talk about crossmodal correspondences?

Note that this singular psychological construct, crossmodal correspondence, has been applied to empirical observations of a range of phenomena, from low-level perceptual associations (e.g., between pitch and size: Stevens, 1934; between pitch and spatial elevation: Parise et al., 2014) to higher-level cognitive ones (e.g., between music and paintings: Parise, 2016; Spence, 2020a; between complex soundtracks and basic tastes: Wang et al., 2015). Further, the term crossmodal correspondence has been applied to perceptual attributes that vary both qualitatively (*what kind*, e.g., hue and vowels; taste quality and auditory timbre) and quantitatively (*how much*, e.g., magnitudes of loudness and brightness, size, duration, etc.:

Stevens, 1957; Marks, 1974). Moreover, different kinds of mechanism have been proposed to explain different kinds of correspondence. For example, pitch-elevation correspondence has been explained in terms of the coding, through associative learning, of statistical regularities between sensory cues in the environment (e.g., Parise, 2016), whereas music-taste correspondences have been explained in terms of common semantic and/or affective responses in different sensory systems (e.g., emotional responses to music and tastes: Palmer et al., 2013; Motoki et al., 2020; Reinoso-Carvalho et al., 2020; Velasco et al., 2015, 2016; Wang et al., 2015).

With the aforesaid issues in mind, a more fundamental question lies at the heart of the present article: What *are* crossmodal correspondences? Whatever properties or processes characterize crossmodal correspondences should, ipso facto, also make clear what differentiates them from other forms of perceptual or conceptual association. We deem this question to be central, as answering it is necessary in order to define the boundaries of crossmodal correspondences and to build a comprehensive theory about them – a theory that should, in the long run, also serve to link their origins to their functional properties. Indeed, there has been initial progress in this direction (e.g., Parise et al., 2016; see Deroy & Spence, 2016, for a special issue on the topic; see also Spence, 2019, for recent theoretical accounts of colour-taste/flavour correspondences). However, it is still not clear whether attempts thus far to conceptualize crossmodal correspondence target an ensemble of distinct/different constructs, or whether they target a unique, singular construct, albeit perhaps a multifaceted and more than a matter of semantics one having different aspects or components (e.g., Spence, 2011).

The core scientific issue here also speaks directly to an even broader question, alluded to above: How do we create or construct scientific categories (Dienes, 2008). When we observe the world, we commonly make a tacit or implicit assumption that there is a "something" out there that underlies our categories or concepts. When there are many exemplars of a particular category or concept, we typically assume that there is something common to all of them, maybe a single property, maybe a conjunction of properties (or at least we assume that the exemplars lie proximal to a singular or conjunctive prototype or category: Bruner et al., 1956). Especially difficult to categorize are those sets of objects having independent, disjunctive properties (e.g., when objects belong to “Y” if they contain “a” or “b” or “c”: e.g., Antony, 2003; Bruner et al., 1956), and this may be the case with crossmodal correspondences. In this sense, we start with three key questions: (1) What are sensory attributes and why may correspondences be considered in terms of attributes? (2) What does it mean to say that two attributes correspond or are associated? (3) How does a correspondence differ from other kinds of (unimodal and multimodal) associations?

In a nutshell, our aim here is to characterize what it is, if anything, that different examples of crossmodal correspondence share. A related question asks: How do these associations differ from other kinds of associations, for example, other sensory associations (e.g., synaesthesia)? In the present article, we first discuss the different definitions that have been given to crossmodal correspondences, summarizing different views of them and evaluating how different theories approach them. Second, we discuss how crossmodal correspondences differ from other kinds of multisensory associations (and associations at large). Finally, we set out those challenges that need to be considered and conclude by briefly discussing future directions of research.

2. Definitions and kinds of crossmodal correspondence

Table 2 provides some representative definitions that have been given to crossmodal correspondences. Note that these definitions tend to be relatively broad, with different researchers even using a variety of similar albeit not-identical terms, such as “synaesthetic correspondences”, “cross-sensory correspondences”, and “crossmodal associations” (e.g., Martino & Marks, 2000; Spence, 2011; Walker & Walker, 2016). Implicit to the multiple definitions is the notion that correspondences entail equivalences between values, either absolute or relative, of attributes in different senses, e.g., between high loudness and high brightness, and between high pitch and high visuo-spatial elevation.

Table 2. Definitions of crossmodal correspondences.

Definitions of crossmodal correspondences	
Marks (1978)	"…analogies, equivalences, translations of sensory qualities" (p. 102)
Melara & O'Brien (1987)	"Such correspondences are called synesthetic (literally, joining of the senses) meaning that they reflect a presumed connection among attributes from different sensory modalities." (p. 323)
Spence (2011)	"... crossmodal correspondence is used in this review to refer to a compatibility effect between attributes or dimensions of a stimulus[1] (i.e., an object or event) in different sensory modalities (be they redundant or not)." (p. 973)
Martino & Marks (2000)	"... describes milder forms of cross-sensory connections revealed through language and perception." (p. 62)
Walker et al., (2012)	"Cross-sensory correspondences occur when two or more sensory channels provide analogous information about basic stimulus dimensions." (p. 1186)
Deroy et al., (2013)	"Crossmodal correspondences are defined as tendencies for a certain sensory feature (or dimension) to be associated or matched with another feature (or dimension) in a distinct sensory modality…" (p. 879)
Parise & Spence (2013)	“Cross-modal correspondences can be defined as the mapping that observers expect to exist between two or more features or dimensions from different sensory modalities (such as lightness and loudness), that induce congruency effects in performance and often, but not always, also a phenomenological experience of similarity between such features.” (p. 792)

Parise (2016)	"Crossmodal correspondences refer to the systematic associations often found across seemingly unrelated sensory features from different sensory modalities." (p. 7)
Walker et al., (2016)	"The resulting systematicity in the cross-sensory associations (i.e., progressively more extreme values on one dimension are linked to progressively more extreme values on the other dimension) is what the term cross-sensory correspondence is intended to capture." (p. 773).
Jonas et al., (2017)	"Crossmodal correspondences are a feature of human perception in which two or more sensory dimensions are linked together." (p. 1104)
Dreksler & Spence (2019)	"… crossmodal correspondences research: that is, the bidirectional (i.e., transitive), nonarbitrary mappings between the attributes (or dimensions) of two sensory modalities." (p. 4)
Spence (2019)	"Crossmodal correspondences have been defined as the often-surprising crossmodal associations that people experience between features, attributes, or dimensions of experience in different sensory modalities, when either physically present, or else merely imagined." (pp. 235-236)
Spence (2020a)	"… defined as the surprising connections that the majority of people share between seemingly-unrelated stimuli presented in different sensory modalities." (p. 6)
Spence (2020b)	"Crossmodal correspondences are the sometimes-surprising associations that people experience between stimuli, attributes or perceptual dimensions… " (p. 2)
Spence & Sathian (2020)	"Crossmodal correspondences have been defined as tendency for a feature attribute, dimension, or stimulus in on sensory modality, either physically present or merely imagined, to be matched (or associated) with a feature attribute, dimension, or stimulus in another modality." (p. 239)

Note: Not all references used the term "crossmodal correspondences"; instead, some studies used similar terms (e.g., cross-sensory associations) to describe the same phenomena.

2.1. A taxonomy of correspondences

Table 3 summarizes four main types of crossmodal correspondence: structural, statistical, semantic, and affective. Where, we ask, do the four types come from? After reviewing much of the literature on audiovisual correspondences, Spence (2011) suggested a trifold classification of the sources of correspondences, namely, as physiological (or structural), statistical, or semantic (or linguistic or lexical; see Walker, 2012). As Spence and Parise (2013) later argued, their three-fold classification is not necessarily exhaustive, and the different classes of correspondences may not be mutually exclusive. Here, we add a fourth correspondence, “affective”, to Spence’s triad.

Table 3. Summary of four principal types of crossmodal correspondences. This four-fold version comes from adding “affective correspondence” to Spence’s (2011) trifold scheme.

Crossmodal Correspondence	Example	Source / Explanation	Consequences
Physiological	Loudness–brightness	Possibly innate, but may also depend on maturation of neural structures for stimulus coding	Perceptual & decisional
Statistical	Pitch–elevation Pitch–size Loudness–size	Learned: Coupling priors established on the basis of experience with regularities of the environment	Perceptual & decisional
Semantic (also called linguistic and lexical) correspondence	Pitch–elevation Pitch–spatial frequency	Learned: Emerge following language development as certain terms come to be associated with more than one perceptual continuum	Primarily decisional
Affective	Taste–shape curvature	Learned: Based on common affective properties of attributes	Perceptual & decisional

Physiological correspondences may derive from similarities in the neural codes underlying the representations of sensory stimuli (e.g., between loudness and brightness; see Marks, 1989). Earlier work used “structural correspondences” to refer to these similarities (Spence, 2011), but “physiological correspondences” may be more appropriate. The term "structural correspondences" was used based on such putative neural similarities (see Spence, 2011; Spence & Di Stefano, 2022). Statistically mediated correspondences may result from internalizing the statistical regularities in the environment (e.g., pitch and spatial elevation of sound sources, see Parise et al., 2014). And, semantically mediated, or linguistic, correspondences may be metaphoric, that is, may derive from the application, explicit or implicit, of similar descriptors to sensory percepts of different modalities, as expressed in and through language (e.g., a high musical note and a high aroma description; see Deroy et al., 2013). Presumably, the metaphoric uses are in some ways not arbitrary. Importantly, some correspondences may be based on the overlap of the affective properties of different stimuli[2] (as in the case of the correspondence between taste and attributes of shape such as curvature, Velasco et al., 2016; see also Collier, 1996; Kenneth, 1923; Marks, 1996).

Figure 1 illustrates our taxonomic system of crossmodal correspondences, as based on Spence’s taxonomy (Spence, 2011; Spence, 2020b; Spence & Parise, 2013; but see also the recent system of Spence & Di Stefano, 2023). This system hypothesizes that any of four major kinds of associations (physiological, semantic, statistical, and affective) may characterize crossmodal correspondences. Specifically, it assumes that crossmodal correspondences emerge if at least one kind of association (physiological/statistical/semantic/affective) is present. It should be noted, however, that the categories of correspondences need not be restricted to these four. Others may also exist such as those embodied, which refer to the concept that the body's involvement enhances the linkage between sensory information from distinct modalities (Salgado-Montejo et al., 2016).

In other words, our physical bodies can exert a profound influence on the way we perceive and comprehend our surroundings (a concept known as embodied cognition, Glenberg, Witt, & Metcalfe, 2013). This notion highlights the close interconnection between our sensory encounters, thoughts, and cognition with our bodily sensations and actions. For example, according to Parise et al. (2014), the association between auditory pitch and elevation appears to be accounted for, at least in part, by a representation of sound that is grounded in the body. Here, it is suggested that pitch's spatial connotation is influenced by the statistics of natural auditory scenes, indicating that various related phenomena, such as the outer ear's shape, have adapted to these statistics. It should be also noted, though, that a taxonomy is only a summary of observations and not itself a theory. The taxonomy does not explain correspondences, although descriptions do or can entail constraints on theoretical explanations.

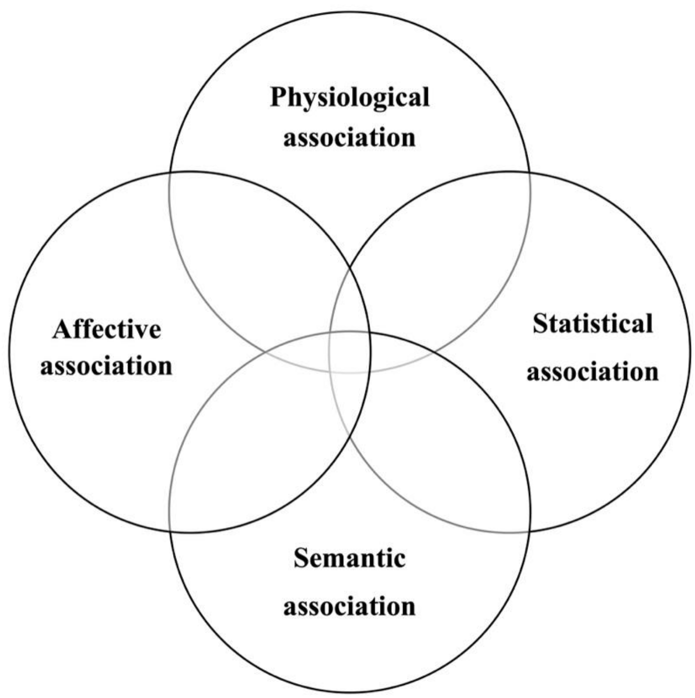

Figure 1. Taxonomy of crossmodal correspondences following Spence (2011, 2020b). This figure illustrates four major kinds of associations (physiological, semantic, statistical, and affective) that may characterize crossmodal correspondences. It assumes that crossmodal correspondences emerge if at least one kind of association (physiological/statistical/semantic/affective) is present.

A given form of crossmodal correspondence is, of course, itself a category, and need not always derive from only one kind of associations, but may, in different instantiations, reflect more than one. Different crossmodal correspondences (e.g., pitch-elevation and taste-shape) may be based on qualitatively different and various sources. For example, pitch-elevation correspondences may be based on semantic, statistical, and/or affective similarities, and possibly others too. Pitch and elevation share linguistic labels in certain languages. Pitch can have "high" quality and space "high" location. Pitch and spatial elevation can also share affective properties. Higher pitch and higher spatial location are associated with greater positive valence compared to lower pitch and lower location (Belyk & Brown, 2014; Meier & Robinson, 2004).

3. What key elements are needed to characterize and explain crossmodal correspondences?

3.1. On the meaning of "correspondence"

Do correspondences reside in people's subjective feelings of similarity (e.g., individuals' personal judgments of how two things are alike) or analogy? And/or are correspondences better considered to reside wholly in the measures of observable patterns of behaviours?

Here, we see two complementary perspectives on crossmodal correspondences. One, the phenomenal view, defines crossmodal correspondences in terms of perceptual experiences, as properties of the conscious mind, of which people are, or can be, generally aware. We take the subjective reports directly as evidence of crossmodal correspondences. Consider, for instance, the way that high pitch is (more) similar to white than to black (e.g., Melara, 1989). One inference from such phenomenal reports is that there is a conscious, crossmodal correspondence between auditory pitch and visual brightness. Early studies by Nafe (1927), and Hornbostel (1927) suggested that there are fundamental perceptual qualities that are common to experiences across different senses, such as brightness, which these investigators thought to constitute a universal attribute of sensory perception, a property of experiences in all sense modalities.

In another complementary view, crossmodal correspondences refer to observable patterns of behavioural response that do not necessarily require phenomenal awareness of the correspondences themselves, or in any case do not require any attempt to assess phenomenal awareness (Marks, 2004). By this view, a crossmodal correspondence may refer, for example, to a pattern of behavioural responses to multiple sensory stimuli. Noteworthy are "congruency effects," which arise from manipulating different (congruent vs. incongruent) components of multiple sensory stimuli and are measured in terms of speed and/or accuracy of behavioural responses (Parise & Spence, 2013). In this view, crossmodal correspondences may be defined as the congruency effects themselves, rather than (or in addition to) being expressions of, say, mental equivalences between experiences of different sensory modalities.

There is a substantial difference, however, between saying that congruency effects *are* crossmodal correspondences and saying that they *are expressions of* crossmodal correspondences. The former is simply a definition, the latter an assumption bearing deeper theoretical implications but requiring correspondingly greater justification: What's added by the assumption that crossmodal correspondences underlie crossmodal congruence effects? The answer, presumably, comes with a theoretical account of correspondence(s). In this sense, one may think of crossmodal correspondences as a theoretical construct (see Hyland, 1981) that underlies the patterns of responses.

3.2. What corresponds in "crossmodal correspondences"?

When researchers study correspondences, they describe and analyse a number of different potentially corresponding attributes, whose characteristics we summarize in Table 4. We acknowledge the other fined grained kinds of dimensions (e.g., Polar, Circular; see Spence & Di Stefano, 2022), but our table provides only a descriptive account of the major different types.

Table 4. Sensory dimensions and sensory features that may correspond.

Sensory dimension	Definition	Examples
Quantity	How much (from not at all to very much)	brightness, temperature, size
Quality	What kind (qualitive differences of sensory stimuli)	taste qualities (e.g., sweet, bitter, sour), colour hue (e.g., red, green), consonants (e.g., voiced, voiceless)
Sensory feature		
Lower-level	Single sensory attributes that cannot be separated	pitch, loudness, temperature
Higher-level	Multiple sensory elements and/or attributes	music, painting, flavour, foods, drinks

Note: What corresponds in "crossmodal correspondences" can be described in terms of sensory dimensions and sensory features. Sensory dimensions include quantity (e.g., brightness, temperature) and quality (e.g., taste qualities, hue). For example, in temperature-hue correspondences, a quantitative dimension (temperature) and a qualitative dimension (hue) correspond. Sensory features refer to lower-level (e.g., pitch, loudness) and higher-level (e.g., musical piece) properties. For example, in colour saturation - music correspondences, a lower-level dimension (saturation) and a higher-level dimension (musical piece) correspond (Palmer et al., 2013).

The broad definition of crossmodal correspondence given at the beginning of this article (i.e., people's associations between attributes or properties that pertain to perceptual experiences in different sensory modalities) may take on multiple meanings (Table 2), reflected, in turn, in the different elements of correspondence (Table 3). As suggested by Turoman et al. (2018), a theory of crossmodal correspondences should apply to both metathetic and prothetic attributes/dimensions. Both terms typically refer to low-level attributes of sense perception, in particular, to those which vary, respectively, in quality (metathetic = "what kind", as in taste qualities, sweet, bitter, sour) and intensity (prothetic = "how much", as in loudness) (see also Marks, 1974; Stevens, 1957).

As suggested by Parise (2016), the term crossmodal correspondence has been applied to everything from associations between simple sensory dimensions (e.g., brightness and loudness, with attributes "bright", vs "dark" or "dim", and "soft" vs. "loud") to more complex perceptual ones (e.g., major/minor mode and colour hue, Palmer et al. 2013); note, however, that in the latter example, a crossmodal correspondence between music and colour might reflect a relation between relatively "lower-level" attributes, such as the relation between dark vs. light and, say, minor vs. major mode or music played in low vs. high register. In any case, it is worth reflecting on whether all of these examples exemplify a singular phenomenon. Indeed, if one follows strictly the definitions of crossmodal correspondence (see Table 2), one would not speak of correspondence between music and colour in general, but instead about correspondences between, say, specific dimensions or features of colours (e.g., hue) and specific dimensions or features of musical sounds (e.g., tempo, pitch).

In addition, relevant here is the distinction between lower- and higher-level attributes. Lower-level attributes (e.g., brightness) are likely to be closely related to effects of stimulating sensory receptors (e.g., in the retina or olfactory bulb), whilst higher-level attributes are likely to depend on the of integration of lower-level attributes into more meaningful, experiential, aspects of our mental life (music, visual scenes) (see Rouw et al., 1997; Spence, 2020a). It should be noted that the distinction between lower- and higher-level attributes can sometimes be ambiguous. A sensory attribute (e.g., brightness) can be treated as both lower- and higher-level. Brightness is sometimes appropriately treated as lower-level by manipulating the luminance of the visual stimuli (e.g., Marks, 1987); but brightness can also depend on higher-level processes, for example, when the word "bright" serves as a stimulus (e.g., Marks, 1982). Acknowledging that the claim may be reductionistic, Parise (2016) nevertheless suggests that crossmodal correspondences may have, at their core, lower-level attributes. Parise (also notes that most previous research on crossmodal correspondences focused on low-level stimulus properties but could be also applied to higher-level stimulus properties (i.e., cognitive associations between concepts and complex stimuli).

3.3. How do attributes correspond in crossmodal correspondences?

Some kinds of crossmodal correspondences presumably depend on *inferential* processes, as, presumably, does multisensory perception itself (Parise, 2015). Multisensory perception can be described as the result of a process akin to logical inference (e.g., Helmholtz, 1909; Parise, 2015). Humans, and presumably some other species, infer the state of the world (e.g., weight of an object) from multisensory cues (e.g., visual and haptic information). In this case, sensory cues from different modalities provide redundant (or complementary) information (e.g., colour, shape, roughness, hardness) about the inferred stimulus property (e.g., weight). By analogy, crossmodal correspondences may also rely on implicit inferential processes (e.g., inferences about physical properties, identities, etc.).

In order to understand what it means to say that two or more sensory attributes correspond, it may be worthwhile thinking of associations in terms of a continuum of associative redundancy[3] that indicates the extent to which signals from different senses provide the same or overlapping information about objects or events in our environment (Parise, 2016). Parise situated crossmodal correspondences within the broader framework of sensory cue integration. Sensory perception can be described as the outcome of a process of inferencing (Helmholtz, 1909). The brain can infer the most likely state of the world by combining noisy sensory cues and prior knowledge (see Parise, 2016). For example, people integrate information from different sensory modalities (i.e., vision and touch) to estimate the width of a grasped object by implicitly considering the relative reliability of the sensory cues (e.g., Ernst & Banks, 2002).

3.4. How are crossmodal correspondences measured?

Crossmodal correspondences have been assessed through several measures, including subjective matches and response times. Subjective matches can be obtained by various psychophysical methods, including direct matching or stimulus-adjustments, as well as by forced choice methods, Likert-scales, visual analogue scales, etc: e.g., Ngo et al., 2013; Motoki et al., 2020; Turoman et al., 2018; Velasco et al., 2014). The so-called “Bouba–Kiki”/“Maluma–Takete” effects, well-known examples of shape-sound correspondences, have typically been measured using a binary forced-choice task, where participants select the appropriate name that “matches” each of two possible shapes (e.g., Bremmer et al., 2013; Köhler 1947; Ramachandran & Hubbard, 2001). Participants tend to match “Bouba” and “Maluma” to rounded shapes and “Kiki” and “Takete” to angular shapes.

Moreover, by using response times as a measure, numerous studies have shown congruency effects in a variety of tasks involving both perceptual and "cognitive" (e.g., linguistic) stimuli. Following earlier research using speeded classification tasks, the domain of congruency effects has been expanded to include enhanced perception of a given attribute (e.g., sweetness: Velasco et al., 2018). Similar procedures have been used to investigate various types of correspondences, including pitch-brightness (e.g., Marks, 1974), loudness-brightness (e.g., Root & Ross, 1965), pitch-hue (e.g., Simpson et al., 1956), sound-taste (e.g., Crisinel & Spence, 2010; Wang et al., 2016) and shape-taste (e.g., Velasco et al., 2014).

The study by Simpson et al. (1956) is a noteworthy early investigation into crossmodal correspondences in children. The children tested in that study were in grades 3-6 (although the ages are not given, in typical American schools in the 1950s, the ages would range from 9-12 years). A later developmental study (Marks, Hammeal, & Bornstein, 1987) investigated three crossmodal correspondences (pitch-brightness, loudness-brightness, and pitch-size) in a population of nearly 500 school-aged children (3.5-13.5 years) and more than 100 adults. Children of all ages and adults showed reliable pitch-brightness and loudness-brightness correspondences in tests using both sensory and verbal stimuli; but only the oldest children (11.5-13-5 years) and adults showed the normative (inverse) pitch-size correspondence. That study also found evidence of visual-auditory correspondence using verbal stimuli when participants rated the implied denotations of an ensemble of words, e.g., the implied loudness (literal) and brightness (crossmodal/metaphorical) of "whisper" and "shout".

3.5. Differences between crossmodal correspondence and related terms

There are several phenomena that appear similar on the surface to, but differ from, crossmodal correspondences. These phenomena include synaesthesia and semantic congruency. We provide some characteristics of phenomena that border on crossmodal correspondences. The characteristics of crossmodal correspondences differ in terms of incidence, time, space, instantiations, and/or relative/absolute aspects from those of several similar but distinct phenomena (Table 5; see also Marks & Mulvenna, 2013). Although crossmodal correspondences are often shared by most people, synaesthesia is not. Nevertheless, note that synesthetes appear to share and experience the culture-based, learned crossmodal associations of the world around them. Semantic congruency refers to similar responses made to different sensory signals having a common *identity* or *meaning* (Chen & Spence, 2017), whereas crossmodal correspondences need not derive from a common identity or meaning. Although synaesthesia and semantic congruency both lie at the conceptual borders of crossmodal correspondences, the three phenomena differ in several ways, as summarized in Table 5.

Table 5. Differences among crossmodal correspondence, synaesthesia, and semantic congruence.

	Crossmodal correspondence	**Synaesthesia**	**Semantic congruence**
Incidence	Shared by most people	Uncommon	Shared by most people
Time and space	Not confined	Not confined	Not confined
Instantiations	Subjective matching (forced choice, rating) Facilitation in information processing (e.g.,	Concurrent experiences (stimulation to one sensory stimulus gives rise to an experience of another sensory	Subjective matching of meanings/identity (forced choice, rating), Facilitation in information processing (e.g., speeded

	speeded classification, visual search)	stimulus: e.g., alphabetical letters induce colour)*.	classification, visual search)
Relative or absolute	Relative	Absolute	Absolute?

Note: * Synaesthesia does not entail the substitution of one sensory experience for another. Instead, the stimulation of one sense leads to involuntary experiences in another, hence concurrent experiences (stimulation by one sensory stimulus gives rise to an additional experience typically produced by another sensory stimulus). It should be noted that the most common form of synaesthesia appears to be intramodal rather than crossmodal.

Relativity of crossmodal correspondences

In general, crossmodal correspondences seem to depend more on relative than absolute values (see Spence, 2020a, for a review). Evidence for the relativity of crossmodal correspondences appears in many findings, for example those of Brunetti et al. (2017), who showed the relativity of pitch-size correspondences using a sequential speeded-classification task. Participants classified the size of visual stimuli (large, small) while hearing concurrent, task-irrelevant sounds (high, low, intermediate). The intermediate sounds were interpreted as “lower” in pitch following a high-pitched tone, but “higher” in pitch following a low-pitched tone. The results showed faster classification responses on sequence-congruent trials (e.g., a small visual stimulus paired with the intermediate-pitched tone preceded by a low-pitched tone) than on sequence-incongruent trials (e.g., a small visual stimulus paired with the intermediate-pitched tone preceded by a high-pitched tone). Walker and Walker (2016) found size-brightness correspondences to be relative rather than absolute: In a speeded-classification task, participants were faster at classifying a visual stimulus as brighter or darker when incidental stimuli were relatively smaller or larger, respectively. That is, regardless of the absolute values of pitch or size, the larger of two objects is matched with the lower-pitched of two sounds. Note that it is also worth considering whether the relativity of crossmodal correspondences is confined to stimuli that can be organized along quantitative dimensions such as temperature. For example, relativity might not appear to characterize crossmodal correspondences between stimuli producing different sensory qualities (e.g., taste qualities).

Physically present and imagined stimuli

Crossmodal correspondences can emerge both when sensory stimuli are physically present and when they are just imagined (Spence, 2020a). Earlier findings have revealed evidence of crossmodal correspondences with imagined stimuli (e.g., linguistic stimuli or words, though described in other terms (e.g., “synesthetic tendencies”, “synesthetic metaphors”: Karwoski, Odbert, & Osgood, 1942; Marks, 1982; Osgood, 1960). For example, non-synesthetic participants matched words describing visual attributes (e.g., large) to words denoting auditory attributes (e.g., loud) and emotional experiences (e.g., bad; Karwoski et al., 1942). More recently, Woods et al. (2013) reported a range of crossmodal correspondences (e.g., colour-weight, shape-taste) when participants imagined sensory stimuli after seeing just their names. In an online word-matching task, participants saw descriptors (e.g., ”heavier”, ”boulder”) and had choose the better match between two response options (e.g., ”red” or ”yellow”, ”sour” or ”sweet”). In this example, the participants reliably matched “heavier” with “red” and “boulder” with “sour”. Recent evidence has also shown that early blind people, who had some early visual experience (but did not recall colours), showed colour-weight correspondences (Barilari et al., 2018). When early blind people were asked whether “red” or “yellow” is heavier, “red” was chosen more frequently than “yellow” (Barilari et al., 2018)[4]. And physically present and imagined stimuli show similar crossmodal correspondences. A sweet tastant, like the word "sweet", is associated with round versus angular shapes (Velasco et al., 2015, 2016). Furthermore, both imagining drinking cold water and actually drinking cold water are associated with higher pitch and faster tempo of auditory stimuli (i.e., short melody; Wang & Spence, 2017). Together, these findings suggest, perhaps not surprisingly, that crossmodal correspondences do not require the physical presence of sensory stimuli; the mere imagining of sensory stimuli triggered by linguistic signals can be matched with attributes in different senses[1].

4. Issues for future research

Crossmodal correspondences refer to the equivalence of values on perceptual dimensions in different senses, such as brightness and loudness or pitch and visuo-spatial elevation. These correspondences can be classified into four main types: physiological, statistical, semantic, and affective, and can arise from various sources such as similarities in neural codes, statistical regularities in the environment, and metaphoric uses of language. The concept of crossmodal correspondence has been used in the context of both lower-level (i.e., single sensory attributes that cannot be separated) and higher-level attributes (i.e., multiple sensory elements and/or attributes) of sense perception. Crossmodal correspondences can be understood from two complementary perspectives: the phenomenal view and the behavioural response view. The former defines crossmodal correspondences in terms of perceptual experiences, while the latter refers to observable patterns of behavioural response to multiple sensory stimuli. The distinction between these views highlights the need for a theoretical account of correspondences as a construct that underlies the patterns of responses. Crossmodal correspondences appear to be relative rather than absolute. Moreover, crossmodal correspondences can also emerge both when sensory stimuli are physically present and when they are just imagined, suggesting that the mere imagining of sensory stimuli triggered by linguistic signals can be matched with attributes in different senses.

Defining and conceptualizing crossmodal correspondences is not an easy task. Although most researchers would agree that crossmodal correspondences, broadly speaking, refer to associations between attributes across the senses, a number of important points need to be considered in this definition. For example, how does one define "correspondence"? What constitutes an association? What are attributes? Moreover, it is not clear whether all correspondences are of the same kind. Through Spence's (2011) taxonomy, we can classify a number of observations. But a taxonomy does not tell us much about the mechanisms underlying crossmodal correspondences, nor explain how we access them.

We need a more nuanced understanding of the two forms of correspondence: phenomenological similarity and congruence effects. What do these two measures imply for characterizing correspondences? Another alternative is that there are instances where individuals lack conscious awareness of correspondences, yet researchers may establish a correspondence through consensus among respondents, regardless of observable behavioural effects. The question here would be whether this still qualifies as a crossmodal correspondence per se. Merely obtaining statistically reliable responses does not necessarily indicate the presence of a statistically reliable behavioural effect. Instead, it might be considered an association that belongs to a broader association architecture (e.g., Alvarado et al., 2023).

With the aforesaid points in mind, it is important to note that there are also other important standing challenges in research on crossmodal correspondences. Are crossmodal correspondences bidirectional or symmetrical? Do various crossmodal correspondences (quantity/quality, lower perceptual/higher cognitive) operate through the same mechanisms? Although we presented Spence's (2011) taxonomy of crossmodal correspondences in Figure 1, above, we note that the strength of several kinds of associations (physiological/semantic/affective/statistical) may vary among the different kinds of crossmodal correspondence (quantity/quality, lower perceptual/higher cognitive).

Last but not least, it is worth noting that another question remains: How do crossmodal correspondences differ from other types of associations? Is there a specific architecture associated with these correspondences, or do they belong to a broader sense of connection between dimensions of stimuli? If the latter were the case, it is unclear whether they would differ from other processes, such as semantic knowledge (e.g., Humphreys & Forde, 2001).

Do different kinds of crossmodal correspondence (quantity/quality, lower perceptual/higher cognitive) rely on the same or different mechanisms?

Spence (2020) suggests that affect contributes more to correspondences of higher cognitive (complex) stimuli (e.g., music) than of lower perceptual (simple) stimuli (e.g., pure tones). Recently, crossmodal correspondences involving complex (higher cognitive) stimuli have been observed (Albertazzi et al., 2015, 2016; Levitan et al., 2015). Higher cognitive stimuli (e.g., a musical composition and a painting) have multiple elements or attributes (Spence, 2020a). For example, musical pieces contain multiple elements of sounds (e.g., tonal complexes varying in pitch, tempo, timbre, etc.) and paintings contain multiple attributes of visual stimuli (e.g., elements or forms varying in hue, shape, brightness, etc,). Crossmodal correspondences involving higher cognitive stimuli might be less based than lower cognitive stimuli on perceptual similarity. This would presumably arise because complex sensory attributes in a pair of associated higher cognitive stimuli (e.g., a musical composition and a painting) are less likely than less complex attributes to be integrated into a single multisensory representation (Spence, 2020a). Complex crossmodal correspondences involving higher cognitive stimuli might be processed in terms of affective meaning (valence/arousal). Thus, it is imperative in future research to investigate whether various forms of crossmodal correspondence are governed by shared or distinct mechanisms. Neuroimaging techniques, for example, may be useful in examining whether different kinds of crossmodal correspondence (quantity/quality, lower perceptual/higher cognitive) rely on the same or different mechanisms (cf. Bien et al., 2012).

5. Conclusions

Overall, gaining a clear understanding of what correspondences are is an important step in developing a comprehensive theory of crossmodal correspondences. Future research needs to understand better when (and why) phenomenal measures (perceptual experiences of subjective matching) and behavioural measures (e.g., observed congruency effects) coincide and when they do not. Future research also needs to examine whether different kinds of crossmodal correspondence (quantity/quality, lower perceptual/higher cognitive) rely on the same or different psychological mechanisms.

Notes

1. Note that it is not always be clear what the stimulus and the percepts are. In experimental settings, the stimulus and percepts are often manipulated while keeping everything else

constant, but this might not always hold true at the perceptual level, especially when proper experimental controls are not in place or when individuals are exposed to natural environments. For instance, consider the flavor of gin – is the stimulus the flavour itself, the taste of alcohol, the aroma, or the mouthfeel? What is more, the perception of these elements can trigger further experiences, effectively acting as internal stimuli.

2. Note, however, that affect is often considered to have not only “quality” but also “magnitude”, that is, the extent to which sensory stimuli feel, say, more or less positive or negative. Indeed, affect is treated in many instances as a bipolar dimension, varying in magnitude in two directions from neutral (zero affective magnitude). Considerable research effort has gone into developing valid and accurate quantitative measures of affect (hedonic value), using bipolar rating scales (e.g., Lim, Wood, & Green, 2009).
3. It is important to mention, however, that it is not always clear what the basis of redundancy is in all crossmodal correspondences (e.g., redundancy in terms of a sensory attribute or affect), nor is it clear whether redundancy might be specific to a subset of crossmodal correspondences.
4. The act of imagining visual attributes (e.g., round shapes) and crossmodal correspondences involving vision seem to differ in sighted individuals and in those who are congenitally blind (Deroy et al., 2016; but see Hamilton-Fletcher et al., 2018). Congenitally blind individuals could conceive of, or understand visual stimuli, based on descriptions provided by others, even though they lack direct prior experience. Semantic knowledge through such indirect experiences may play a role in the development of crossmodal correspondences related to visual stimuli in congenitally blind individuals.

ACKNOWLEDGEMENTS

We thank the anonymous reviewers and, especially, the action editor, Charles Spence, for their thoughtful and thorough comments on the manuscript. We also thank Andy T. Woods for his valuable contributions to earlier versions of this manuscript.

REFERENCES

Albertazzi, L., Bacci, F., Canal, L., & Micciolo, R. (2016). The tactile dimensions of abstract paintings: A cross-modal study. *Perception*, *45*(7), 805-822.

Albertazzi, L., Canal, L., & Micciolo, R. (2015). Cross-modal associations between materic painting and classical Spanish music. *Frontiers in Psychology*, *6*:424.

Alvarado, J., Velasco, C., & Salgado, A. (2023). The organization of semantic associations between senses in language. *PsyArXiv*, 27 March. 2023. https://doi.org/10.31234/osf.io/9jh7k

Antony, L. (2003). Who's afraid of disjunctive properties? *Philosophical Issues*, *13*(1), 1-21. doi: 10.1111/1533-6077.00001.

Barilari, M., De Heering, A., Crollen, V., Collignon, O., & Bottini, R. (2018). Is red heavier than yellow even for blind? *i-Perception*, *9*(1):2041669518759123.

Belyk, M., & Brown, S. (2014). The acoustic correlates of valence depend on emotion family. *Journal of Voice*, *28*(4), 523.e9-523.e18.

Bien, N., Ten Oever, S., Goebel, R., & Sack, A. T. (2012). The sound of size: crossmodal binding in pitch-size synesthesia: a combined TMS, EEG and psychophysics study. *NeuroImage*, *59*(1), 663-672.

Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: John Wiley.

Brunetti, R., Indraccolo, A., Mastroberardino, S., Spence, C., & Santangelo, V. (2017). The impact of cross-modal correspondences on working memory performance. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(4), 819-831.

Chen, Y.-C., & Spence, C. (2017). Assessing the role of the 'unity assumption' on multisensory integration: A review. *Frontiers in Psychology*, *8*:445.

Collier, G. L. (1996). Affective synesthesia: Extracting emotion space from simple perceptual stimuli. *Motivation and Emotion*, *20*(1), 1-32.

Crisinel, A.-S., & Spence, C. (2010). As bitter as a trombone: Synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Attention, Perception, & Psychophysics*, *72*(7), 1994-2002.

Deroy, O., Crisinel, A.-S., & Spence, C. (2013). Crossmodal correspondences between odors and contingent attributes: Odors, musical notes, and geometrical shapes. *Psychonomic Bulletin & Review*, *20*(5), 878-896.

Deroy, O., Fasiello, I., Hayward, V., & Auvray, M. (2016). Differentiated audio-tactile correspondences in sighted and blind individuals. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1204-1214.

Di Stefano, N., & Spence, C. (2023). Perceptual similarity: Insights from crossmodal correspondences. *Review of Philosophy and Psychology*, *14*, https://doi.org/10.1007/s13164-023-00692-y

Dienes, Z. (2008). *Understanding psychology as a science: An introduction to scientific and statistical inference*. Hampshire: Palgrave Macmillan.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429-433.

Glenberg, A. M., Witt, J. K., & Metcalfe, J. (2013). From the revolution to embodiment: 25 years of cognitive psychology. *Perspectives on Psychological Science, 8*(5), 573-585.

Hamilton-Fletcher, G., Pisanski, K., Reby, D., Stefańczyk, M., Ward, J., & Sorokowska, A. (2018). The role of visual experience in the emergence of cross-modal correspondences. *Cognition*, *175*, 114-121.

Hartshorne, C. (1934). *The philosophy and psychology of sensation.* Chicago, IL: University of Chicago Press.

Helmholtz, H. V. (1909). Handbuch der Physiologischen Optik (Handbook of Physiological Optics). Voss, Hamburg, Germany.

Hornbostel, E. M. (1927). The unity of the senses. *Psyche*, *7,* 83-89.

Humphreys, G. W., & Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral and Brain Sciences*, *24*, 453-476.

Hyland, M. (1981). *Introduction to theoretical psychology*. London: Palgrave.

Jonas, C., Spiller, M. J., & Hibbard, P. (2017). Summation of visual attributes in auditory–visual crossmodal correspondences. *Psychonomic Bulletin & Review*, *24*(4), 1104-1112.

Karwoski, T. F., Odbert H. S., & Osgood, C. E. (1942). Studies in synesthetic thinking. II. The role of form in visual responses to music. *Journal of General Psychology.*, *26*, 199-222.

Kenneth, J. H. (1923). Mental reactions to smell stimuli. *Psychological Review*, *30*(1), 77-79.

Köhler, W. (1947). Gestalt psychology (2nd ed.). New York: Liveright.

Levitan, C. A., Charney, S., Schloss, K. B., & Palmer, S. E. (2015). The smell of jazz: Crossmodal correspondences between music, odor, and emotion. Proceedings of the *37th Annual Conference of the Cognitive Science Society*, Pasadena, CA (pp. 1326-1331).

Lim, J., Wood, A., & Green, B. G. (2009). Derivation and evaluation of a labeled hedonic scale. *Chemical Senses*, *34*, 739-751.

Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *The American Journal of Psychology*, *87*(1/2), 173-188.

Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities*. New York: Academic Press.

Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(2), 177-193.

Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, *13*(3), 384-394.

Marks, L. E. (1989). On cross-modal similarity: The perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 586-602.

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

Marks, L. E. (2004). Cross-modal interactions in speeded classification. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.). *The Handbook of Multisensory Processes* (pp. 85-105). Cambridge, MA: MIT Press.

Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor. *Monographs of the Society for Research in Child Development*, 52, No.1 (Serial No. 215).

Marks, L. E., & Mulvenna, C. M. (2013). Synesthesia, at and near its borders. *Frontiers in Psychology*, *4:*651.

Martino, G., & Marks, L. E. (2000). Cross-modal interaction between vision and touch: the role of synesthetic correspondence. *Perception*, *29*(6), 745-754.

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: Associations between affect and vertical position. *Psychological Science*, *15*(4), 243-247.

Melara, R. D. (1989). Dimensional interaction between color and pitch. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(1), 69-79.

Melara, R. D., & O'Brien, T. P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General*, *116*(4), 323-336.

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

Nafe, J. P. (1927). The psychology of felt experience. *American Journal of Psychology*, *39*,367-389.

Ngo, M. K., Velasco, C., Salgado, A., Boehm, E., O'Neill, D., & Spence, C. (2013). Assessing crossmodal correspondences in exotic fruit juices: The case of shape and sound symbolism. *Food Quality and Preference, 28*(1), 361-369.

Osgood, C. E. (1960). The cross-cultural generality of visual-verbal synesthetic tendencies. *Behavioral Science, 5*(2), 146-169.

Palmer, S. E., Schloss, K. B., Xu, Z., & Prado-León, L. R. (2013). Music–color associations are mediated by emotion. *Proceedings of the National Academy of Sciences, 110*(22), 8836-8841.

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

Parise, C. V., Knorre, K., & Ernst, M. O. (2014). Natural auditory scene statistics shapes human spatial hearing. *Proceedings of the National Academy of Sciences, 111*(16), 6104-6108.

Parise, C. V., & Spence, C. (2012). Audiovisual crossmodal correspondences and sound symbolism: A study using the implicit association test. *Experimental Brain Research*, *220*(3-4), 319-333.

Parise, C. V., & Spence, C. (2013). Audiovisual cross-modal correspondences in the general population. In J. Simner & E. Hubbard (Eds.), *The Oxford handbook of synaesthesia* (pp. 790-815). Oxford, UK: Oxford University Press.

Parise, C. V., Spence, C., & Deroy, O. (2016). Understanding the correspondences: Introduction to the special issue on crossmodal correspondences. *Multisensory Research, 29*(1-3), 1-6.

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

Root, R. T., & Ross, S. (1965). Further validation of subjective scales for loudness and brightness by means of cross-modality matching. *American Journal of Psychology*, *78*(2), 285-289.

Rouw, R., Kosslyn, S. M., & Hamel, R. (1997). Detecting high-level and low-level properties in visual images and visual percepts. *Cognition*, *63*(2), 209-226.

Salgado-Montejo, A., Marmolejo-Ramos, F., Alvarado, J. A., Arboleda, J. C., Suarez, D. R., & Spence, C. (2016). Drawing sounds: representing tones and chords spatially. *Experimental Brain Research*, *234*, 3509-3522.

Simpson, R. H., Quinn, M., & Ausubel, D. P. (1956). Synesthesia in children: Association of colors with pure tone frequencies. *Journal of Genetic Psychology*, *89*(1), 95-103.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, *73*(4), 971-995.

Spence, C. (2019). On the relationship (s) between color and taste/flavor. *Experimental Psychology*, *66*(2), 99-111.

Spence, C. (2020a). Simple and complex crossmodal correspondences involving audition. *Acoustical Science and Technology*, *41*(1), 6-12.

Spence, C. (2020b). Temperature-based crossmodal correspondences: Causes and consequences. *Multisensory Research*, *33*(6), 645-682.

Spence, C. (2022). Exploring group differences in the crossmodal correspondences. *Multisensory Research*, *35*, 495-536.

Spence, C., & Deroy, O. (2013). How automatic are crossmodal correspondences? *Consciousness and Cognition*, *22*(1), 245-260.

Spence, C., & Di Stefano, N. (2022). Coloured hearing, colour music, colour organs, and the search for perceptually meaningful correspondences between colour and sound. *i-Perception*, *13*(3):20416695221092802.

Spence, C., & Di Stefano, N. (2023). Sensory translation between audition and vision. *Psychonomic Bulletin & Review*, doi.org/10.3758/s13423-023-02343-w

Spence, C., & Sathian, K. (2020). Audiovisual crossmodal correspondences: Behavioral consequences and neural underpinnings. In K. Sathian and V.S. Ramachandran (Eds.), Multisensory perception from laboratory to clinic (pp. 239-258). Academic Press.

Stevens, S. S. (1934). Are tones spatial? *American Journal of Psychology*, *46*, 145-147.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, *64*, 153-181.

Turoman, N., Velasco, C., Chen, Y. C., Huang, P. C., & Spence, C. (2018). Symmetry and its role in the crossmodal correspondence between shape and taste. *Attention, Perception, & Psychophysics*, *80*(3), 738-751.

Velasco, C., Hyndman, S., & Spence, C. (2018). The role of typeface curvilinearity on taste expectations and perception. *International Journal of Gastronomy and Food Science*, *11*, 63-74.

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

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

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

Walker, L., & Walker, P. (2016). Cross-sensory mapping of feature values in the size–brightness correspondence can be more relative than absolute. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(1), 138-150.

Walker, L., Walker, P., & Francis, B. (2012). A common scheme for cross-sensory correspondences across stimulus domains. *Perception*, *41*(10), 1186-1192.

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.

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

Walker, P., Scallon, G., & Francis, B. (2017). Cross-sensory correspondences: Heaviness is dark and low-pitched. *Perception*, *46*(7), 772-792.

Wang, Q. J., & Spence, C. (2017). The role of pitch and tempo in sound-temperature crossmodal correspondences. *Multisensory Research*, *30*(3-5), 307-320.

Wang, Q. J., Wang, S., & Spence, C. (2016). "Turn up the taste": assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and pitch. *Chemical Senses*, *41*(4), 345-356.

Wang, Q., Woods, A. T., & Spence, C. (2015). "What's your taste in music?" A comparison of the effectiveness of various soundscapes in evoking specific tastes. *i-Perception*, *6*(6): 2041669515622001.

Woods, A. T., Spence, C., Butcher, N., & Deroy, O. (2013). Fast lemons and sour boulders: Testing crossmodal correspondences using an internet-based testing methodology. *i-Perception*, *4*(6), 365-379.