Title	Understanding personalised auditory-visual associations in multi-modal interactions
Authors	O'Toole, Patrick
Publication date	2021-10-18
Original Citation	O'Toole, P. (2021) 'Understanding personalised auditory-visual associations in multi-modal interactions', ICMI '21: Proceedings of the 2021 International Conference on Multimodal Interaction, Montréal QC, Canada, 18-22 October, pp. 812-816. doi: 10.1145/3462244.3481277
Type of publication	Conference item
Link to publisher's version	10.1145/3462244.3481277
Rights	© 2021, the Author. This work is licensed under a Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/) - https://creativecommons.org/licenses/by/4.0/
Download date	2024-04-23 07:02:45
Item downloaded from	https://hdl.handle.net/10468/12272

Understanding Personalised Auditory-Visual Associations in Multi-Modal Interactions

Patrick O'Toole
patrick.otoole@umail.ucc.ie
University College Cork
Cork, Ireland

ABSTRACT

Can we sharpen our auditory and visual senses and better understand the relationship between these modalities to benefit our interactions in human-computer interfaces? This research paper proposes a framework to understand auditory-visual associations and explore the impact of emotion, personality, age and gender in understanding information from both modalities. Studies into the areas of emotion and personality as well as their association with the auditory and visual senses have increased within the fields of psychology, neuroscience, affective computing and human-computer interaction (HCI). From a HCI perspective, advances in technologies and machine learning techniques provide a new way to understand people and to develop systems where computers work along side people to help develop efficient interactions and clearer perceptions of our environment. The proposed framework will be developed along side a personalised auditory-visual interface that can be used to provide intelligent interactions with users that can help them learn from efficient associations between their senses. This research can be used to create personalised auditory-visual-emotion-personality profiles that can be use in adaptive musical teaching platforms, as well as mental health and wellness applications for more personalised care programs.

CCS CONCEPTS

• **Human-centered computing** → **Sound-based input / output**; *Interaction design process and methods*; • **Applied computing** → Sound and music computing; • **Social and professional topics** → **User characteristics**.

KEYWORDS

auditory-visual associations, multi-modal interactions, music, synaesthesia, machine learning

1 INTRODUCTION

There is a quote from W.B. Yeats that states, *"The world is full of magic things, patiently waiting for our senses to grow sharper"*. With our advancement in technology and computing, the tools are at our fingertips to understand and enhance our senses and how they can be combined to intelligently interact in the world. The field of human-computer interaction (HCI) can assist people in developing their sensory perception and striking the words, *"patiently waiting"* from W.B. Yeats statement above.

Interdisciplinary research between the disciplines of HCI and psychology, philosophy and neuroscience will be investigated as part of my PhD, with extensive research into auditory-visual associations already present in the aforementioned disciplines. Different approaches to auditory-visual association research has been conducted in the past, with investigations into basic and complex cross-modal associations between the auditory and visual senses [2, 15, 21, 26], connections between synaesthesia and associations between senses in non-synaesthetes [5, 8], and the role of the emotion mediated hypothesis in auditory-visual associations [19, 24]. New research is needed, with advancements in technology capable on adding new knowledge that previously could not be obtained. To bring a new understanding to auditory-visual association research, a more in-depth investigation into cross-modal associations, synaesthesia, emotion, mood, personality, age and gender is required. A framework using knowledge regarding the aforementioned variables and their relation to our auditory and visual senses can be implemented in human-computing interfaces to assist in sharpening our senses and creating pleasant personalised experiences.

Synaesthesia is a rare condition that, typically, arises as one sense (an inducer) receives stimulation that triggers a reaction in a second sense (a concurrent) that is unstimulated [8]. Research into synaesthesia and its subcategory, Chromaesthesia, which is sound-colour synaesthesia, has helped further our understanding of associations between these senses in the general population [2].

Psychology and neuroscience disciplines are at the foreground of this research. Technological advancements and new intelligent computing processes like machine learning, computing, and more directly human-computer interaction can provide new insights into auditory-visual associations. Machine learning and, more broadly, AI are viewed in popular culture as domains that will bring the end of humanity, with automation taking over the majority of work previously done by people, eventually leave us without any purpose or direction. The proposal of hybrid intelligence (HI), that views a need for humans and machines to work together to achieve goals that are unreachable by either one independently, looks to counter the negative effects that AI can have on our society [1]. The research

proposed in this paper seeks to develop a framework for sharpening auditory-visual associations with the principles of HI in mind.

Affective states such as emotion, mood and personality traits can have an impact on our auditory and visual perception. Studies into musical sounds have provided insights into the impact of emotion in auditory-visual associations. These studies showed that faster music in the major modes produced more saturated and lighter colours, with the opposite colour types for slower music in the minor mode. Strong correlations have been found between the emotional associations of music and those of colours, supporting the emotion mediation hypothesis [19, 24]. Long-term affective states, such as personality traits, have produced interesting associations with musical stimuli. A study that investigated the relationship of different modes and tempos in music to personality traits revealed that the emotional stability and optimism traits were significant predictors of preference for music in fast tempos and a major key, with the opposite music preference showing to be significantly predicted by the traits of openness to experience, introversion and also gender [10]. These studies involving emotion, personality and gender, give insight into the complexity of our senses and how we can understand the different variables that can impact auditory-visual associations. A new approach is presented here that will explore a personalised approach to auditory-visual associations using a human-computer interface to assist in gaining a new understanding.

The research questions I seek to answer are as follows:

(1) Can auditory-visual associations be sharpened with technological assistance using a human-computer interface?
(2) Is a personalised association between auditory and visual senses more beneficial than the previously-researched general associations?
(3) If more personalised auditory-visual associations are beneficial to the overall goal, what is the importance of emotion, personality, age and gender in creating more accurate and personalised interactions between the auditory and visual senses?

The rest of this paper is laid out as follows, Section 2: Background, Section 3: Research Methodology, Section 4: Results and Section 5: Future Work.

2 BACKGROUND & RELATED WORK

Presently, computer-centred design is still at the foreground of how we interact with computing environments. It is predicted that human-centred design will be adopted more prominently in the design of future systems, blending computing into our everyday lives with the human user at the forefront [33]. The study of affective computing, in tandem with HCI, is important in designing future human-centred systems. Gratch states that, *"Technology is less efficient if it perturbs human emotions; more efficient if it engages with them productively; more attractive if it appeals to human emotions; and often it is primarily concerned with enabling humans to experience particular emotions (notably happiness)"* [11]. Where affective computing will provide the understanding of users affective states, it is HCI that will use this information to create an accessible and clear interaction between the computer and user.

A recent article outlining a new perspective on the future of AI presents the term *"Hybrid Intelligence"*, which can be defined as the

"combination of human and machine intelligence, augmenting human intellect and capabilities instead of replacing them and achieving goals that were unreachable by either humans or machines" [1]. The research presented here sets out to understand auditory-visual associations, with HCI, affective computing and the principles of HI as a base.

2.1 Auditory-visual Associations

The associations between the auditory and visual senses are well researched, with different auditory features known to be associated with different aspects of colour [19, 21, 24, 30, 31]. Higher pitches and major modes have been found to be associated with brighter colours, with lower pitches and minor modes associated with darker colours [8, 19, 30]. Other associations between timbre and saturation [19, 24], loudness and brightness, pitch and size [21] have also been documented.

Spence presents three qualitatively different kinds of cross-modal correspondence, statistical, semantic and structurally mediated, that can have different developmental trajectories and consequences for human perception and behaviour. Structural correspondence occurs because of neural connections present at birth, and are related to dimensions involving magnitude, i.e. loudness-brightness. Statistical correspondence is the associations between two stimulus that correlate in nature, i.e. pitch-size, pitch-elevation. While semantic correspondence can occur when the terms people use to describe stimuli overlap in two dimensions, i.e. pitch-spatial frequency [21].

A study by Thomas et al., shows that adults showed a speed advantage over children for associations that are well-known but no advantage for newly acquired auditory-visual pairing. These findings show that age has no effect on the influence of semantic associations on multi-sensory processing, however it shows that these associations become more robust and influential [23]. Understanding congruent associations between the auditory and visual senses is an important factor in the research presented in this paper. As mentioned in Section 1, research into cross-modal associations benefited from research into the condition called synaesthesia.

2.2 Synaesthesia

Synaesthetes and non-synaesthetes experience cross-modal associations differently. It can be characterised as the difference between general and personal associations, with synaesthetes experiencing personal associations as a consequence of their condition [8]. This personal association that synaesthetes possess has motivated the research presented here and exploring variables such as emotion, personality, age and gender, will make it possible to create these synaesthesia-like personalised auditory-visual associations for the general population.

2.3 Personality Traits

The psychology and music disciplines have produced interesting studies concerning the connection between music and personality [7, 10, 27]. In more recent years, research in computing disciplines has started to focus on using different machine learning techniques to understand the associations between certain audio features and personality, and between personality and emotion [6, 16]. While Carbonneau et al. investigated speech segments and extracting

features using spectrograms to perform classifications of personality traits, the question presented here will focus on non-verbal or musical audio, and the issue of whether or not personality and other variables impact decisions on associations between the auditory and visual senses [6]. These research findings can help take into account how emotion, personality, age and gender can be understood together with auditory-visual associations to achieve the goals presented in this paper.

2.4 Emotion

As mentioned in Section 1, emotion can have a significant impact on the associations between the auditory and visual senses, as set out in the emotion mediation hypothesis [19, 24]. This understanding shows a need for thorough research into how music emotions are understood and the associations with the visual sense. A study looking into musically-induced emotions using discrete, dimensional and music-specific emotion models found that the dimensional emotion model was overall the most consistent in rating the sixteen music excerpts tested [27]. However in the same study, it was found personality-related differences were more pronounced in the discrete emotion model [27]. This shows that further research is needed to enable us to identify the approach that will allow us to best understand emotion and personality in the context of associations between our senses. A multi-modeled approach could be beneficial in gaining a better understanding of both emotion and personality in the proposed research. Other studies have also shown interesting links between emotion and personality in relation to musical stimuli [7, 10, 12, 14, 28]. Participants listening to sad musical excerpts were shown to experience sadness but also more complex emotions such as nostalgia, peacefulness and wonder. The personality traits of empathy and openness to experience were also associated with sad musical excerpts [28]. While these studies show how auditory stimuli impact the perception of emotion and personality, an understanding of auditory and visual senses are needed as well as insights into how age and gender might affect the associations between these two senses.

2.5 Age & Gender

Multi-sensory perception of emotions has been found to be more accurate than perception through one sense, with females displaying more accuracy than males [29]. Similar results were also found in [13], where *"women were found to be more accurate in recognition of emotional prosody"*. With regard to age, it has been shown that age plays a significant role in emotion perception between the auditory and visual senses [17, 32]. A study by Wurm et al. showed that there were latencies in emotional Stroop tasks between incongruent stimuli, but only for older adults [32]. Studies of perception between positive and negative emotions within different age groups have also shown that older participants have increased difficulty recognising some of the basic emotions in the four modalities of faces, voices, bodies/context and matching of faces to voices [20].

These differences between genders and age groups show that these two variables could provide important information, as well as, emotion and personality when it comes to understanding auditory-visual associations and creating personalised interactions in the proposed framework.

3 RESEARCH METHODOLOGY

To answer the research questions presented in Section 1, a plan will be implemented using the following approach:

(1) Develop a HCI framework for personalised auditory-visual associations.
(2) Extract and analyse auditory features and test if certain features can be connected to certain emotions and personality traits of the users.
(3) Implement an auditory-visual interface that models the relationship between variables such as emotion and personality and presents congruent visual images associated with the auditory stimuli.

To help inform the design of the framework, an extensive investigation into past research in the disciplines of HCI, affective computing, psychology and affective neuroscience will be conducted, with regard to the impact of emotion, personality, age and gender on the associations between the auditory and visual senses.

From this knowledge, a study will be conducted using the Big Five Personality test to analyse participants' personalities while also capturing auditory data played on various instruments. From the resulting experiments, correlations between personality and auditory features will be analysed to find if personality can be understood through analysis of an individual's musical style, timbre, use of major/minor mode, tempo. If the hypothesis can be confirmed, a machine learning approach will be explored to extract personality information in real-time for use in displaying congruent visual imagery associated with the auditory stimuli.

While emotion, in relation to auditory-visual associations, is a more developed area, similar methods will be studied to understand the participants emotional reaction to the stimuli to help present congruent visual imagery to the auditory stimuli elicited. With the information gathered from these studies, an auditory-visual interface will be created that embodies this knowledge in its design, which will be used in answering the main research question.

Figure 1 presents the flow of information in the proposed framework, showing the different interactions between variables. It can be broken up into two parts, on the right is the auditory-visual associations that happen from an action, i.e. playing a note on a piano, and on the left is the variables proposed that can personalise the auditory-visual associations shown on the right. The filled orange arrows show the flow of the process, from the starting disposition, to an action, to the audio-visual associations, back to an action, i.e. playing another note. On the left, the long term (LTV), medium term (MTV), and short term (STV) variables are presented. These variables interact with each other and auditory-visual associations in different ways, the arrow length and colour describe the speed of the interactions between two points in Figure 1.

My previous research focused on the right side of the framework using predefined associations between sound and colour, based on general associations between auditory and visual senses [18]. Future research will involve the left side of Figure 1, and understanding if and how variables, such as, emotion, mood, personality, age and gender impact the associations between the auditory and visual senses shown on the right side of Figure 1.
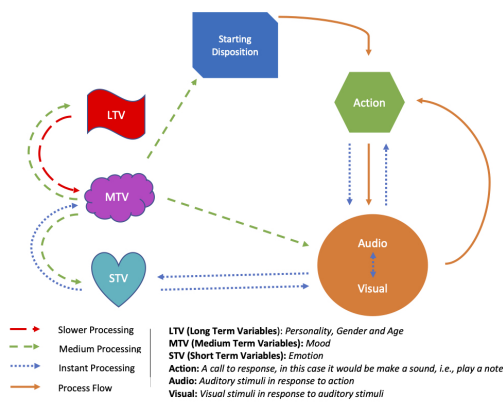
Figure 1: The Flow of the Information Involving Personalised Auditory-Visual Associations

4 RESULTS

My previous research investigated connections between auditory-visual cross-modal associations and Chromaesthesia, which as mentioned in Section 1, is a form of sound-colour synaesthesia [18]. This research explored the theory of synaesthesia in the general population, with the possibility of the condition being present in everyone, and it can be perceived to be on a spectrum with different strengths in the associations between the senses [9, 25].

The experiment conducted in [18] used an ad-hoc program where participants would play musical notes on a guitar and coloured shapes would display on the screen related to the auditory stimuli elicited. Certain well researched associations were implemented in this program, such as pitch to lightness and brightness, pitch to elevation and pitch to size and direction of movement [21, 24]. This study found that, after using the program three times, the relationship between the auditory stimuli and participants' colour choices grew stronger, in line with the sound-colour associations used in the program [18]. The interesting insights from this study, have driven me to explore in more depth the way people experience associations between their senses, and what other variables can impact these connections.

I have conduced some preliminary studies on single musical note to emotion associations, and the impact that different variables can have. These results have confirmed the auditory to emotion associations presented in [18, 22] and also produced interesting insights concerning personality, gender and age. These insights will help in further experiments that will be conducted as part of my PhD.

Another project I have been involved in is, investigating how the use of virtual agents instead of humans as leaders affects the perception of transactive memory system (TMS) in a team. Transactive memory is a mental representation of the distribution of knowledge between the members of a team and it has been shown that the two leadership styles of Transformational (TFL) and Transactional (TAL) are positively associated with TMS, and were used in this study [3]. The results show that the intervention from both the Embodied Computational Agent (ECA) and the Human Actor can improve the perceived TMS of a team and that the TFL style was found to perform better for the Human Actor, with both TAL and TFL impacting the perception of TMS with the ECA [4].

This research work has given additional insights into the main research goals of my PhD, where the proposed framework could be updated to reflect an understanding of leadership styles that could help produce more effective communication in the workplace, as well as personality and emotion.

5 FUTURE WORK

As part of my ongoing research, I have been working with data collected as part of the EU-H2020-ICT Project TELMI, that is publicly available on the repoVizz platform. The data collected for this project included synchronised audio, video, motion capture and physiological (EMG) data, however for my research the audio and physiological data from four violin performers is being analysed. Five musical excerpts, played by all performers (two professional and two advanced players) will be used to create a visual representation of the music, and the physiological data will be use to measure muscle contraction and bowing arm movements to add a more in-depth visual representation of the music.

A study will be conducted into participants' understanding of the visual representation of the musical excerpt. Posing two questions:

(1) Can musical style and emotion be understood from auditory stimulation translated into visual only stimulation?
(2) Can musical skill level be understood from only the visual representation of the musical excepts?

Results will provide a greater understanding of how colour and emotion can be implemented into an auditory-visual interface to create effective musical communications.

As mentioned in Section 3, a study will be conducted to find out if personality traits can be measured by analysing the musical performance of a person playing an instrument. Comparing standard Big Five personality test scores with certain audio features such as musical style, timbre, use of major/minor mode, tempo, I hope to find a correlation between certain traits and the audio features captured from analysis of participants' performances. The results from this study will be used to implement personalisation into the program used in the first study mentioned in Section 5.

The research proposed in this paper might be applied to beneficial effect in various areas. For example, the creation of personalised profiles using auditory-visual-emotion-personality associations that can be implemented in adaptive teaching platforms for learning an instrument; entertainment software or video games to enhance the users experience, and also in mental health and wellness applications for more personalised care programs.

ACKNOWLEDGMENTS

REFERENCES

[1] Zeynep Akata, Dan Balliet, Maarten de Rijke, Frank Dignum, Virginia Dignum, Guszti Eiben, Antske Fokkens, Davide Grossi, Koen Hindriks, Holger Hoos, Hayley Hung, Catholijn Jonker, Christof Monz, Mark Neerincx, Frans Oliehoek, Henry Prakken, Stefan Schlobach, Linda van der Gaag, Frank van Harmelen, Herke van Hoof, Birna van Riemsdijk, Aimee van Wynsberghe, Rineke Verbrugge, Bart Verheij, Piek Vossen, and Max Welling. 2020. A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence. *Computer* 53, 8 (Aug. 2020), 18–28. https://doi.org/10.1109/MC.2020.2996587

[2] Liliana Albertazzi, Luisa Canal, and Rocco Micciolo. 2015. Cross-Modal Associations between Materic Painting and Classical Spanish Music. *Frontiers in Psychology* 6 (April 2015), 1–13. https://doi.org/10.3389/fpsyg.2015.00424

[3] Daniel G. Bachrach and Ryan Mullins. 2019. A dual-process contingency model of leadership, transactive memory systems and team performance. *Journal of Business Research* 96, November 2018 (2019), 297–308. https://doi.org/10.1016/j.jbusres.2018.11.029

[4] Beatrice Biancardi, Patrick O'Toole, Ivan Giaccaglia, Brian Ravenet, Maurizio Mancini, and Giovanna Varni. 2021. How ECA vs Human Leaders Affect the Perception of Transactive Memory System (TMS) in a Team. In *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, Nara, Japan, 1–8.

[5] David Brang and V. S. Ramachandran. 2011. Survival of the Synesthesia Gene: Why Do People Hear Colors and Taste Words? *PLoS Biology* 9, 11 (Nov. 2011), e1001205. https://doi.org/10.1371/journal.pbio.1001205

[6] Marc-Andre Carbonneau, Eric Granger, Yazid Attabi, and Ghyslain Gagnon. 2020. Feature Learning from Spectrograms for Assessment of Personality Traits. *IEEE Transactions on Affective Computing* 11, 1 (Jan. 2020), 25–31. https://doi.org/10.1109/TAFFC.2017.2763132

[7] Tomas Chamorro-Premuzic and Adrian Furnham. 2007. Personality and Music: Can Traits Explain How People Use Music in Everyday Life? *British Journal of Psychology* 98, 2 (2007), 175–185. https://doi.org/10.1348/000712606X111177

[8] Caroline Curwen. 2018. Music-Colour Synaesthesia: Concept, Context and Qualia. *Consciousness and Cognition* 61 (May 2018), 94–106. https://doi.org/10.1016/j.concog.2018.04.005

[9] Caroline Curwen. 2020. Music-Colour Synaesthesia: A Sensorimotor Account. *Musicae Scientiae* (Oct. 2020), 1–20. https://doi.org/10.1177/1029864920956295

[10] Snježana Dobrota and Ina Reić Ercegovac. 2015. The Relationship between Music Preferences of Different Mode and Tempo and Personality Traits – Implications for Music Pedagogy. *Music Education Research* 17, 2 (April 2015), 234–247. https://doi.org/10.1080/14613808.2014.933790

[11] Jonathan Gratch. 2021. The Field of Affective Computing: An Interdisciplinary Perspective. *Transactions of the Japanese Society for Artificial Intelligence* 36, 1 (2021), 13.

[12] Patrik N. Juslin. 2013. What Does Music Express? Basic Emotions and Beyond. *Frontiers in Psychology* 4 (2013), 1–14. https://doi.org/10.3389/fpsyg.2013.00596

[13] Lena Lambrecht, Benjamin Kreifelts, and Dirk Wildgruber. 2014. Gender Differences in Emotion Recognition: Impact of Sensory Modality and Emotional Category. *Cognition and Emotion* 28, 3 (April 2014), 452–469. https://doi.org/10.1080/02699931.2013.837378

[14] Solange Mardaga and Michel Hansenne. 2009. Do Personality Traits Modulate the Effect of Emotional Visual Stimuli on Auditory Information Processing? *Journal of Individual Differences* 30, 1 (Jan. 2009), 28–34. https://doi.org/10.1027/1614-0001.30.1.28

[15] Lawrence E Marks. 2004. Cross-Modal Interactions in Speeded Classification. In *The Handbook of Multisensory Processes*, Gemma Calvert, Charles Spence, and Barry E. Stein (Eds.). MIT Press, Cambridge, Mass, 85–105.

[16] Laura Alejandra Martínez-Tejada, Yasuhisa Maruyama, Natsue Yoshimura, and Yasuharu Koike. 2020. Analysis of Personality and EEG Features in Emotion Recognition Using Machine Learning Techniques to Classify Arousal and Valence Labels. *Machine Learning and Knowledge Extraction* 2, 2 (April 2020), 99–124. https://doi.org/10.3390/make2020007

[17] Matthew W. E. Murry and Derek M. Isaacowitz. 2017. Age Differences in Emotion Perception: The Effects of the Social Environment. *International Journal of Behavioral Development* 41, 5 (Sept. 2017), 597–604. https://doi.org/10.1177/0165025416667493

[18] Patrick O'Toole, Donald Glowinski, and Maurizio Mancini. 2019. Understanding Chromaesthesia by Strengthening Auditory -Visual-Emotional Associations. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, Cambridge, United Kingdom, 1–7. https://doi.org/10.1109/ACII.2019.8925465

[19] S. E. Palmer, K. B. Schloss, Z. Xu, and L. R. Prado-Leon. 2013. Music-Color Associations Are Mediated by Emotion. *Proceedings of the National Academy of Sciences* 110, 22 (May 2013), 8836–8841. https://doi.org/10.1073/pnas.1212562110

[20] Ted Ruffman, Julie D. Henry, Vicki Livingstone, and Louise H. Phillips. 2008. A Meta-Analytic Review of Emotion Recognition and Aging: Implications for Neuropsychological Models of Aging. *Neuroscience & Biobehavioral Reviews* 32, 4 (Jan. 2008), 863–881. https://doi.org/10.1016/j.neubiorev.2008.01.001

[21] Charles Spence. 2011. Crossmodal Correspondences: A Tutorial Review. *Attention, Perception, & Psychophysics* 73, 4 (May 2011), 971–995. https://doi.org/10.3758/s13414-010-0073-7

[22] Charles Spence. 2020. Assessing the Role of Emotional Mediation in Explaining Crossmodal Correspondences Involving Musical Stimuli. *Multisensory Research* 33, 1 (July 2020), 1–29. https://doi.org/10.1163/22134808-20191469

[23] Rhiannon L. Thomas, Marko Nardini, and Denis Mareschal. 2017. The Impact of Semantically Congruent and Incongruent Visual Information on Auditory Object Recognition across Development. *Journal of Experimental Child Psychology* 162 (Oct. 2017), 72–88. https://doi.org/10.1016/j.jecp.2017.04.020

[24] Tawney Tsang and Karen B. Schloss. 2010. Associations between Color and Music Are Mediated by Emotion and Influenced by Tempo: (525772013-006). https://doi.org/10.1037/e525772013-006

[25] Crétien van Campen. 2010. *The Hidden Sense: Synesthesia in Art and Science*. MIT Press, Cambridge, Mass.

[26] Veijo Virsu, Henna Oksanen-Hennah, Anita Vedenpää, Pentti Jaatinen, and Pekka Lahti-Nuuttila. 2008. Simultaneity Learning in Vision, Audition, Tactile Sense and Their Cross-Modal Combinations. *Experimental Brain Research* 186, 4 (April 2008), 525–537. https://doi.org/10.1007/s00221-007-1254-z

[27] Jonna K. Vuoskoski and Tuomas Eerola. 2011. Measuring Music-Induced Emotion: A Comparison of Emotion Models, Personality Biases, and Intensity of Experiences. *Musicae Scientiae* 15, 2 (July 2011), 159–173. https://doi.org/10.1177/1029864911403367

[28] Jonna K. Vuoskoski, William F. Thompson, Doris McIlwain, and Tuomas Eerola. 2012. Who Enjoys Listening to Sad Music and Why? *Music Perception* 29, 3 (Feb. 2012), 311–317. https://doi.org/10.1525/mp.2012.29.3.311

[29] Teija Waaramaa. 2017. Gender Differences in Identifying Emotions from Auditory and Visual Stimuli. *Logopedics Phoniatrics Vocology* 42, 4 (Oct. 2017), 160–166. https://doi.org/10.1080/14015439.2016.1243725

[30] J Ward, B Huckstep, and E Tsakanikos. 2006. Sound-Colour Synaesthesia: To What Extent Does It Use Cross-Modal Mechanisms Common to Us All? *Cortex* 42, 2 (2006), 264–280. https://doi.org/10.1016/S0010-9452(08)70352-6

[31] Kelly L. Whiteford, Karen B. Schloss, Nathaniel E. Helwig, and Stephen E. Palmer. 2018. Color, Music, and Emotion: Bach to the Blues. *i-Perception* 9, 6 (Nov. 2018), 1–27. https://doi.org/10.1177/2041669518808535

[32] Lee H. Wurm, Gisela Labouvie-Vief, Joanna Aycock, Kristine A. Rebucal, and Heather E. Koch. 2004. Performance in Auditory and Visual Emotional Stroop Tasks: A Comparison of Older and Younger Adults. *Psychology and Aging* 19, 3 (2004), 523–535. https://doi.org/10.1037/0882-7974.19.3.523

[33] Zhihong Zeng, M. Pantic, G.I. Roisman, and T.S. Huang. 2009. A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 1 (Jan. 2009), 39–58. https://doi.org/10.1109/TPAMI.2008.52