Lecture Notes in Education, Arts, Management and Social Science,
2026, Volume 4, Issue 5
https://bbwpublisher.com/journal/LNE
Online ISSN: 2705-053X

# Music—Colour—Emotion Cross-Modal Association Research Review——From the Perspective of Cultural Differences and Adolescent Music Aesthetic Education

**Yidan Xie***

Ezhou Vocational University, Ezhou 436000, Hubei, China

***Author to whom correspondence should be addressed.***



**Abstract:** While cross-modal correspondence theory establishes the conceptual foundation for exploring auditory-visual associations, the intricate triadic interplay among music, colour, and emotion has emerged as a pivotal frontier in sensory cognition and pedagogical research. To date, international scholarship has postulated two primary theoretical frameworks: the direct association hypothesis and the emotion-mediated hypothesis. Most of the existing empirical research takes Western cultural groups as its research subjects, but systematic research focusing on Chinese children aged 6-18 years old is relatively rare. Local unique cultural elements, such as Chinese pentatonic music, traditional Chinese musical instruments, and traditional symbolic colours, will create a new audio-visual perception mode for residents, but systematic research on the stratified development of elementary education, junior middle education, and senior middle education stages is relatively rare. Despite the effectiveness of using colour in teaching music in foreign countries, empirical studies about its effect in China are relatively rare. Based on the original research proposal literature materials, this paper categorizes the literature on the relationship between music, colour and emotion from three aspects, which include theoretical mechanism, cross-cultural empirical research and educational application. Through this paper, we can see the agreement and contradictions in current research, the prominent gaps in research in the Chinese adolescent sample group and indigenous culture stimulus and put forward the research directions for follow-up localized music aesthetic education. This review may give theoretical basis for development of cultural adaptation colour-assisted music education instruments and cross-cultural empirical evidence to the cross-modal correspondences theory.

**Keywords:** Cross-modal correspondence; Music-colour association; Emotion mediation; Cross-cultural sensory perception; Adolescent music aesthetic education

**Online publication:** May 26, 2026

## 1. Introduction

The theory of cross-modal correspondence elaborates the universal sensory link between music and colour visual stimuli, and the corresponding relationship among music, colour and emotion has become an interdisciplinary research focus involving experimental psychology and music pedagogy [1]. Two central interpretation models have been constructed about this cross-sensory correspondence phenomenon in current research, namely the direct association hypothesis and the emotion-mediated hypothesis. Of them, the latter hypothesis has been verified empirically more often using child and adolescent samples, where it is proposed that music arouses certain emotions first and emotion mediates to form relatively stable music-colour matching rules [2].

But the current empirical verification experiments all choose Western classical music, a Western universal colour system, and Europeans and Americans as research objects, and the rules extracted cannot be fully extended to non-Western cultural groups [3]. For Chinese children and adolescents aged 6 to 18 years old, the interference of native cultural factors exists in their audio-visual emotion perception model, including pentatonic scale, erhu, guzheng and other national music instruments to form its unique musical emotion clues; at the same time, traditional symbolic colours such as China red and porcelain blue carry collective culture memories and differentiated emotional implications [4, 5, 6]. In contrast to the Western twelve-tone equal temperament scale, the traditional five-note modes (Gong, Shang, Jue, Zhi, and Yu) of China have the unique quality of integrating timbral modality, season, and natural colours as one embodied cultural metaphor. As a case in point, the Jue mode is semantically connected with spring and the colour of green, representing an audio-visual association based on ancient Chinese natural philosophy [4]. On the other hand, traditional Chinese colours such as vermilion red and cerulean blue are not just purely visual stimulation, but deeply ingrained in the conceptual schema of Yin-Yang and the Five Elements, which constructs the structure of Chinese people's embodied multimodal experience of sound and sight. Such culturally ingrained schemas present an inherent logic of emotional correspondence between traditional music modes and symbolic colours in China. Cross-cultural comparative studies, including those with Chinese subjects, have used college students as their sample, which fails to account for the developmental changes that occur in the music-colour association rule at primary, secondary, and senior secondary school levels [6].

In terms of educational practice, visualization methods have been extensively applied to music education in foreign countries to help children comprehend the abstract aspects of music like mode, timbre, and melody. Nevertheless, in general, the development of Chinese music education is overly focused on technical practice and neglects the cultivation of students' sensual and emotional experience, and there is a dearth of empirical studies proving the effectiveness of teaching with colour aids in the local youth group [7,8].

In light of the aforementioned deficiencies in research, this paper attempts to systematically review the current literature on cross-modal correspondences among music, colour and emotion from an original research proposal. In this paper, two fundamental hypotheses in theories will be organized, a comparison of the empirical findings will be made between the Western and Chinese cultural context, the applications of cross-modal association to music education will be summarized, the existing limitations in current research on adolescents in China and cultural stimuli will be identified, and future research directions will be proposed.

## 2. Theoretical mechanisms of music-colour-emotion cross-modal associations

The auditory-visual correspondence of music and colour is based on the common psychological operational mechanism for human multi-sensory perception. There are two theoretical hypotheses suggested by scholars in order to explain the inner mechanism of the auditory-visual correspondence phenomenon.

### 2.1. Direct association hypothesis

The basic idea behind the direct association hypothesis is that there is a natural physical association between the parameters of musical and colour stimuli. The main logic behind this hypothesis is that the frequency of vibrations of musical pitch equals the frequency of light of colour, making up a natural cross-modal association rule that does not depend on any emotional experience of the individual [9]. The basic idea behind this hypothesis is primarily supported in studies of congenital strong synaesthetes who can form visual colour images through hearing certain pitches without emotional mediation.

However, later studies have found that this theory is inadequate to explain the consistent music-colour match findings of normal children and teens who do not suffer from synaesthesia. The matching performance of most teenagers will vary according to the emotions generated by music content, and this kind of matching pattern can hardly be explained through simple physical correspondence [10,11].

On the other hand, the direct association hypothesis proposes a physical deterministic logic of the correspondence between audio and visual through frequency correlation, which is able to provide a robust explanation for congenital strong synaesthetic people who have fixed cross-sensory mapping by birth. But the limitation of this theory is self-evident. For one thing, it fails to explain the flexible colour matching results for non-synaesthetes who do not have congenital strong synaesthesia, because their matching results vary greatly in accordance with their emotional states but not the fixed acoustic conditions. For another, the theory fails to take into consideration the effect of the developed culture and development of humans on the modification of cross-modal correspondence, treating cross-modal correspondence merely as an innately physical response. There is little discussion about the possibility of a dual-path hybrid approach which takes into consideration both the physical correspondence and emotionality.

### 2.2. Emotion-mediated hypothesis

Nowadays, the emotion-mediated model has been recognized as the prevailing theoretical model for the understanding of the poor cross-modal relationships between ordinary adolescents. The fundamental idea of the emotion-mediated model is that music cannot establish stable colour matching relationships; different types of music evoke different emotions and people select the related colour images based on the value and intensity of emotions induced by music [10].

Strong synaesthesia and weak everyday cross-modal associations add more evidence to the rationality of the emotion-mediated theory. Martino & Marks classify cross-modal matching as two types: strong synaesthesia represents an inflexible and involuntary audio-visual trigger mechanism that only exists in congenital synaesthetes, which is consistent with the logic of the direct association theory; weak synaesthesia is a flexible matching mechanism, which is prevalent among young people, and entirely depends on emotions for the connection between music and colour [12].

In combination with results from studies of developmental psychology, 6–12-year-old children belong to the stage of concrete operations and depend on sensory intuition for perceiving emotional meanings of music; 13–18-year-old adolescents are capable of abstract thinking and perceive complex emotions in music

and diverse rules of colour matching. It is further evidence of the key role of emotional experience as an intermediary of music-colour association in youth groups [13,14].

Empirical research findings from both childhood and adolescent groups confirm the greater generalisability of the emotion-mediated hypothesis compared to the direct connection hypothesis, particularly regarding the weaker cross-modal associations that occur frequently among youth populations. The major strength of the former is that it involves the introduction of the intermediary concept of affect into the equation of music-colour connections in developmentally and individually differing populations. However, a clear limit for the theoretical basis has not been addressed in previous research studies. Firstly, the theory tends to overgeneralize the role of emotion as the only possible intermediary and leaves out the influence of the lower-level priming effect of physical senses, making no distinction between the independent effect of sound elements regardless of emotional tone. Besides, most of the previous empirical work on this hypothesis uses the Western scales of emotion and colour and the question arises as to whether the intermediary theory works equally well in the Chinese Yin-Yang Five Elements musical context.

## 3. Empirical research on music-colour-emotion associations: Western and Chinese cultural contexts

### 3.1. Empirical consensus based on Western cultural samples

Numerous studies conducted by researchers in the West have shown that there is a stable emotion-based rule for music-colour matching in the child and adolescent population. The study conducted by Barbiere et al. on adolescents from Western countries involving the genres of classical, pop, and folk music found that music with emotions like joy had a consistent association with bright warm colours like red and yellow, while music with negative emotions like sadness was consistently matched to dark colours such as gray and black[2]. This conclusion was replicated by Cimbalo et al. among 8-12 year old children, and it was stated that there is a certain level of consistency in this matching pattern based on emotional valence in children and adolescents [15].

Further experimental studies have used different kinds of music stimuli, such as classical, rock, jazz, and others, and proved that the audio-visual matching pattern mediated by emotion does exist among Western cultural individuals. Nevertheless, since all materials used in these experiments were Western music and Western universal colour templates, the universality of the above-mentioned conclusion is difficult to verify [11].

There have been many scientific findings regarding emotional effects on the music-colour associations for children in Western countries. These findings prove that there exists an absolute consistency between emotional valence and colour in various musical styles. Nevertheless, two important restrictions hinder the applicability of the existing literature in cross-cultural settings. The first one is related to the usage of experimental stimuli: all studies use exclusively Western twelve-tone music and Western neutral colours, thus omitting culturally specific auditory and visual symbols. In other words, there exists a hidden cultural bias in the so-called universal matching rules of Western literature. The second problem is associated with age categorization; the vast majority of the works make distinctions solely between children and teenagers and fail to track the developmental changes in the cross-modal association throughout the period of schooling.

### 3.2. Research progress of music-colour association under Chinese cultural background

In recent years, some researchers have made use of Chinese subjects in cross-modal matching studies, including indigenous culture-based stimuli in cross-modal matching, thereby overcoming the drawback of

using one Western sample in the earlier studies. Lin et al. used Chinese university students as experimental subjects, adopted Western and Chinese national instrumental music in cross-modal stimuli, and included traditional Chinese symbolic colours like China Red and Porcelain Blue in their matching test [6].

This study makes an important differentiated finding distinct from Western sample research: Western music makes colour-music rules according to the "like-dislike" emotional dimension, whereas Chinese national instrumental music builds colour matching rules according to the "warm-cold" emotional dimension. This is due to the special nature of the emotions attached to Chinese traditional colours and timbral features in pentatonic music composition [16,17].

The emotional perception of colours in China is distinctive, thus contributing to the differentiation of culture in terms of music-colour associations. The colour red evokes both positive and negative emotions in collective Chinese culture, whereas the colour blue evokes mostly positive emotions; the local audience will also perceive visual colours on the dimensions of "soft-hard" and "turbid-transparent," which are not considered by Western systems of colour perception research [16,17].

In spite of that, the current local research on the topic in China still suffers from clear one sample problems: all participants of research are university students, without any design for stratified sampling involving children and teenagers aged 6-18 in elementary schools, middle schools, and high schools. It does not allow to trace age and cognitive development-related changes in musical-colour correspondence rules, and does not give an opportunity to investigate the interaction of education experience and cultural symbols in forming cross-modal perceptual modes [18].

Early research on Chinese local culture has separated affective aspects of pentatonic music that differ from the warm-cold dimension of Western like-dislike comparison and distinct colour emotions influenced by Chinese traditional cultural symbols. However, there are three main research gaps in this new field of study.

First, subject selection is constrained to college students without stratified sampling, including 6-18-year-old teenagers at different school ages, which does not allow tracking the process of gradual reconstruction of audio-visual match through cognitive development. Developmental psychology theory gives crucial theoretical background for interpreting qualitative distinctions of music-colour-emotion cross-modal associations between age groups [13]. Six to twelve-year-olds remain in the concrete operational stage when cognitive activities depend much on the perceptual experience from the external environment; they can make only simple one-to-one matching between sounds of music and colour visuals, and cross-modal association is primitive and unstable, prone to distraction by visual qualities. On the other hand, teenagers of 13-18 years old have already entered the formal operational stage, developing the ability to dissociate concrete sensory perception from emotions, and associate cultural symbolism (Five Elements, pentatonic modal metaphor) with sound-colour association [13]. Theory of adolescent identity by Erikson serves as additional explanation of this developmental logic, because during 13-18 years teenagers constantly learn native culture and develop self-cultural identity, which means that music-colour association inevitably includes cultural schema and represents qualitative leap from children's one [14]. Local Chinese literature totally ignores this age difference, unable to reveal the psychological mechanism of cross-modal matching.

Second, the current experiments distinguish pentatonic music, cinnabar red, porcelain blue, and the cultural implications of the Five Elements from each other as independent variables, without developing holistic experimental setups for exploring the combined effect of these elements.

Thirdly, although the existing studies focus on explaining the phenomenon of cultural differences, they have not yet formulated a new emotion-mediation theory that is adjusted for Chinese embodied culture (Jue mode-spring-green connection).

## 4. Educational application value of music-colour cross-modal associations

The matching scheme mediated by emotions in music-colour has tremendous practical application for addressing the problems faced in adolescent music aesthetics education. Studies abroad have proven that visual prompts through colours have a great positive impact on learners' understanding of abstract music concepts. The match of minor mode sad music and low luminance grey colours will help in increasing the accuracy in the identification of music emotional properties and increase in the memorization rates of music content using colour prompts [3,18].

Children's spontaneous painting creation through musical perception can intuitively demonstrate their proficiency in terms of musical key elements like tone and musical form; while the auxiliary methods of colour notation can be useful for teenagers to feel the subtle change in tones during performance practice [19,20]. Nevertheless, the development of music education worldwide shows a tendency to overvalue skill training and neglect sensory and emotional training of the students, which is especially obvious in Chinese elementary music classes.

Currently, the use of colour by Chinese music teachers to assist their teaching is mostly based on their individual experiences, with no systematic experimental data available to support it. No systematic colour-based learning materials exist that fit the cognitive features of students at various stages of schooling, nor any data that would verify the effects of using colours as an auxiliary method on class participation and learning effectiveness among Chinese adolescents. Understanding the developmental laws of the relationship between music, colour, and emotion of the local youth could lead to creating age-specific cultural teaching schemes [7,8].

Teaching trials abroad have been conducted to confirm the auxiliary role of colour visual information in comprehending emotions and timbre of music, whereas education domestically only stops at the summary experience of teachers without systematic intervention data support. Significantly, the educational study currently conducted possesses three blind spots. First, the vast majority of colour-assisted teaching approaches come straight from Western classroom plans without targeted optimization for the Chinese music curriculum and Five Elements colour cultural cognition of the teenage generation. Second, the performance index is limited, as previous studies only evaluate students' capability of recognizing melodies and emotions, without considering the lasting effects on their aesthetic literacy and multisensory thinking skills. Third, the comparative study among different learning periods lacks comparison, and the effect of colour-assisted teaching on various developmental periods of cognition remains unexplored.

## 5. Overall limitations of existing research

According to the classification of domestic and international literature made in the original research project, four basic weaknesses that are found in the research area of music-colour-emotion cross-modal association are as follows:

First, regarding the sample composition, there exists an obvious flaw in that it does not contain any stratified data involving Chinese youths in all age groups. Almost all the available studies involve either

Western adults or children, while a few existing local studies on China only use college students as samples, and thus fail to identify the developing principle for cross-modal match from childhood to late adolescence.

Second, the localization design of the cultural stimuli is lacking. Currently, most experimental research applies Western classical music and colour schemes that apply to all nations but fail to introduce Chinese pentatonic musical works, instrument playing, and colour schemes of China, unable to examine the influence of indigenous cultural memory on individual cross-modal perception.

Third, the link between the theoretical study and practical study of education is not strong. For example, theoretical studies mainly concentrate on proving the causal relationship of cross-modal matching, but educational studies are limited to classroom experiments. There is a lack of a combination of methodological intervention studies of quantitative surveys and qualitative interviews with teachers in order to prove the practical impact of teaching colour on teaching effectiveness.

Fourth, there is a certain homogeneity of research methods. Most of the previous research has used the method of static music-colour matching score, which lacks longitudinal tracking experiment and classroom intervention design.

## 6. Future research directions

In light of the abovementioned research problems that arise out of the initial research proposal, this article proposes three possible development directions for further localized research on Chinese music aesthetic education:

To begin with, conduct stratified empirical research among children and adolescents in China aged between 6 and 18. This will involve dividing the sample population of research into groups such as primary school, junior high school and senior high school, establishing parallel stimulus groups of Western music and Chinese national instrumental music, and identifying the differences in the music-colour-emotion connection patterns at various age levels.

Secondly, study the coupling method between the cultural symbols of China and cross-modal perception of individuals. Use the traditional colours like China Red and Porcelain Blue and national musical instruments like erhu and guzheng as representative experimental material to analyse the recombination process of audio-visual emotion rules based on cultural memory, and improve the research of cross-modal correspondence theory in cultural perspective.

Thirdly, research localized music education interventions with colour assistance. Design differentiated visual teaching systems with colour differentiation to suit students of various educational levels; employ mixed research methodologies involving teacher surveys and classroom experiments; examine the influence of colour assistance on students' classroom engagement and learning efficacy; and develop an operable teaching tool grounded in evidence-based practice that suits Chinese basic music education.

## 7. Conclusion

The emergence of the music-colour-emotion cross-modal association process is controlled jointly by the natural function of sensory mechanisms, emotion-related intermediate variables and the environmental factor of cultural acquisition. Direct association and emotion mediation are two key theories that have been able to explain the audiovisual matching phenomena, respectively, of which the emotion mediation theory has

greater interpretative validity towards ordinary child and adolescent samples. Current Western experimental studies have achieved relatively stable music-colour matching results according to the emotional valence, while the limited local study in China has proved that native music creation and colour cultural symbols would create distinct cross-modal associations.

Currently, there are several important research gaps in the entire field, including a lack of stratified samples of Chinese youth, poor localization of cultural stimulus designs, as well as a gap between theoretical research and educational practices. Conducting systematic empirical research on Chinese youth in the age group of 6 to 18 years old, by utilizing instrumental music in the nation and traditional symbolic colour variable design, will help to not only add more empirical research for cross-modal correspondence theory but also serve as a valuable theoretical base for China's basic music classroom practices.

## Disclosure statement

The author declares no conflict of interest.

## References


[1] Spence C, 2011, Crossmodal Correspondences: A Tutorial Review. Attention, Perception, & Psychophysics, 73(4): 971–995.

[2] Barbiere JM, Vidal A, Zellner DA, 2007, The Colour of Music: Correspondence Through Emotion. Empirical Studies of the Arts, 25(2): 193–208.

[3] Lindborg PM, Friberg AK, 2015, Colour Association with Music Is Mediated by Emotion: Evidence From an Experiment Using a CIE Lab Interface and Interviews. PLOS ONE, 10(12): e0144013.

[4] Yau V, 1994, Colour in China. British Journal of Aesthetics, 34(2): 151–162.

[5] Bogushevskaya V, 2024, A Study in Scarlet: Cultural Memory of the Tropes Related to the Colour Red, Female Countenance, and Onstage Makeup in the Sinophone World. Semiotica, 261: 209–225.

[6] Lin K, Zhang D, Chen R, 2024, The Emotional Bridge: Exploring the Association Between Colour and Western and Chinese Music. Proceedings of the 17th International Symposium on Visual Information Communication and Interaction: 112–118.

[7] Carioti F, Berlingeri M, Bodrato M, 2025, Music Education at School: Too Little and Too Late? Evidence From a Longitudinal Study on Music Training in Preadolescents. Bollettino dell'Associazione Italiana di Musicologia, 10(1): 2704.

[8] Newton M, Crawford R, Southcott J, 2025, The Problem of Secondary School Music Enrolments: Tensions Between Relevance, Engagement and Improving Musical Skills and Knowledge. Music Education Research, 27(1): 1–18.

[9] Pridmore R, 1992, Music and Colour: Relations in the Psychophysical Perspective. Colour Research & Application, 17(1): 57–61.

[10] Palmer SE, Schloss KB, Xu Z, et al., 2013, Music–Colour Associations Are Mediated by Emotion. PNAS, 110(22): 8836–8841.

[11] Whiteford KL, Schloss KB, Helwig NE, et al., 2018, Colour, Music, and Emotion: Bach to the Blues. i-Perception, 9(6): 1–27.

[12] Martino G, Marks LE, 2001, Synesthesia: Strong and Weak. Current Directions in Psychological Science, 10(2):

61–65.

[13] Piaget J, 1952, The Origins of Intelligence. International Universities Press.

[14] Steinberg L, 2005, Cognitive and Affective Development in Adolescence. Trends in Cognitive Sciences, 9(2): 69–74.

[15] Cimbalo RS, Beck L, Weiner MA, 1978, Colour Responses to Musical Emotion in Children Ages Eight Through Twelve. Journal of Psychology, 133(2): 303–304.

[16] Wang T, Shu S, Mo L, 2014, Blue or Red? The Effects of Colour on the Emotions of Chinese People. Asian Journal of Social Psychology, 17(2): 152–158.

[17] Gao XP, Xin JH, Sato T, et al., 2007, Analysis of Cross Cultural Colour Emotion. Colour Research and Application, 32(3): 223–229.

[18] Cutietta RA, Haggerty K, 1987, A Comparative Study of Colour Association With Music at Various Age Levels. Journal of Research in Music Education, 35(2): 78–91.

[19] Martinović Bogojević J, 2021, Hearing Colour, Seeing Sound—Visual Expression of Music Perception. Proceedings of the International Scientific Professional Meeting Art and Education, Academy of Arts in Novi Sad: 382–393.

[20] Poast M, 2000, Colour Music: Visual Colour Notation for Musical Expression. Leonardo, 33(3): 215–221.