



OPEN ACCESS

EDITED BY

Xuemin Zhang,
Beijing Normal University, China

REVIEWED BY

Yuan Yuan,
Central South University, China
Xiaolong Li,
Hohai University, China

*CORRESPONDENCE

Yannan Zhang
✉ northxs@163.com

RECEIVED 08 December 2024

ACCEPTED 28 May 2025

PUBLISHED 01 July 2025

CITATION

Zhang Y and Wang X (2025) The impact of sensory modalities and background information on the emotional resonance of Li Bai's classical poetry. *Front. Psychol.* 16:1541680. doi: 10.3389/fpsyg.2025.1541680

COPYRIGHT

© 2025 Zhang and Wang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The impact of sensory modalities and background information on the emotional resonance of Li Bai's classical poetry

Yannan Zhang* and Xiao Wang

College of Arts, Northeastern University, Shenyang, China

Introduction: Emotional resonance (ER) refers to the extent to which readers identify with and emotionally connect to the representations in poetry, often drawing parallels between the emotions conveyed in the poem and their own personal experiences. This phenomenon is influenced by sensory modalities—such as visual, auditory, tactile, gustatory, and olfactory—and the cultural context embedded within the poem. Understanding how these factors interact to enhance emotional resonance in poetry is crucial for advancing poetry education and engagement.

Methods: This study employs a repeated-measures experimental design to investigate how different sensory modalities and contextual background influence emotional responses to classical poetry. The experiment includes five conditions: three focus on sensory modalities (visual, auditory, and visual-auditory), while the remaining two examine the combined effects of background information (background music and background images) in conjunction with sensory modalities.

Results: The results indicate that: The visual modality and the visual-auditory modality yield higher emotional resonance ratings compared to the auditory modality. Background information (background music and images) enhances the emotional resonance of the poetry. The synergy between background information and sensory modalities further increases the emotional resonance ratings.

Discussion: This study highlights the critical role of both sensory modalities and background information in shaping the emotional response to poetry. The findings suggest that integrating different sensory experiences and cultural cues can significantly enhance emotional resonance, offering new theoretical insights for the teaching and appreciation of poetry.

KEYWORDS

sensory modalities, background information, poetry, emotional resonance, Li Bai

Introduction

Blizek (1973) proposed the aesthetic motivation theory, which argues that emotional resonance (ER) in art stems from its ability to evoke emotions in the viewer or reader. In poetry, imagery represents the manifestation of abstract emotions created within subjective consciousness, symbolizing objects that carry

emotional weight (Frame et al., 2024). The stronger the ER elicited by an artwork, the deeper the reader's aesthetic experience of its imagery. Other studies have also suggested that ER enhances the reader's aesthetic evaluation (Blizek, 1973), and recent research indicates that ER significantly contributes to the aesthetic appeal of art (Silvia, 2005). Santos et al. (2021) further pointed out that ER induced by art enhances aesthetic pleasure.

Specifically, in the field of poetry, poetry attracts readers' attention through its unique aesthetic qualities, initially sparking their interest and curiosity about the appreciation of beauty. Through profound emotional expression and the creation of artistic imagery, poetry gradually resonates with the reader's inner world, eliciting emotional engagement and reflective thinking (Leder et al., 2014). ER refers to the reader's ability to find similarities between their own emotions or experiences and those conveyed in poetry, thereby fostering a deep emotional connection and sense of identification (Graber and Sumera, 2020; Liu X. L., 2022). It can be understood as the extent to which an individual identifies with and comprehends the emotional representations in poetry (Coburn, 2001). ER is influenced by an individual's relevance to the emotional aspects of poetry, including the modalities through which they receive and process external information. These modalities primarily include visual, tactile, auditory, gustatory, and olfactory modalities (Frame et al., 2024). Additionally, ER is shaped by the cultural background embedded in poetry (Yuan and Guoyuan, 2022).

The impact of sensory modalities on the emotional resonance of poetry

ER typically occurs within milliseconds and is determined by the "interactive structure," which influences the extent to which a relatively definite feeling or a sense of mutual authenticity is experienced (Stolorow and Atwood, 2014). Previous research has primarily examined the role of ER in the visual modality (Beebe et al., 1997; Coburn, 2001; Pan et al., 2019).

Specifically in the field of poetry, previous research comparing the effects of different sensory modalities on ER has been limited. Ongoing self-reports on the impact of spoken poetry on aesthetic evaluation and ER showed no significant differences in ratings and the distinction between spoken and written poetry has yet to be thoroughly examined (Wagner et al., 2021). Many studies have shown that, compared with pure audio or pure visual presentation, the combination of audio and visual information can enhance emotional perception and appeal (Thompson et al., 2008). A recent study demonstrated that sensory modalities, as important channels of information perception, produce differences in the intensity of ER across modalities (Bishop, 2019). Overall, the finding regarding the limited comparisons among visual, auditory, and audiovisual modalities remain unclear (Frame et al., 2024).

The impact of background information on the emotional resonance of poetry

The concept of culture is an abstract notion that evokes differences by separating subjectivity, identity, and language from

lived behaviors (O'Connor, 2024). Artists engage with culture, and culture, in turn, influences individuals, generating ER (Paiva, 2020).

Specifically in the field of poetry, research has primarily focused on poetry education. Previous research found that the signifiers of background music and its direct identity evoke an immediate intuitive understanding (Gu, 2024), and music in poetry often evoked stronger ER (Küssner and Eerola, 2019). Similarly, the symbols, composition, and colors of background images can convey specific cultural emotional values and modes of expression (Langacker, 2002). In the process of poetry education, background images enhance the ER of poetry (Zhao, 2024). However, research using experimental methods to compare the impact of different cultural backgrounds on ER remains limited (Wang and Wang, 2024).

Aims

Based on the above discussion, ER is the phenomenon in which individuals perceive the emotions of others, causing a change in their own emotional state. Sensory modalities serve as important channels through which humans perceive external information (Frame et al., 2024). Previous research has shown that emotional intensity varies depending on the modality of information reception (Johnson-Laird and Oatley, 2022). This study focuses on a limited set of sensory modalities (visual, auditory, and visual-auditory) and compares their impact on ER in poetry.

Moreover, existing studies have found that ER is greatly influenced by cultural background (Alshaer, 2025). Since individual development is significantly affected by social norms and cultural traditions (Guo, 2024), people tend to show stronger ER when exposed to situations belonging to their own cultural background (Miller, 2015).

Therefore, the hypotheses of this study are as follows:

H1: Sensory modalities (visual, auditory, and visual-auditory) differ in their impact on the ER of poetry.

H2: Culture background information (background music and background image) enhances the ER of poetry.

H3: Culture background information synergizes with sensory modalities, collectively enhancing the ER of poetry.

These hypotheses form the core framework of this study, which aims to explore how different dimensions of sensory modalities and background information individually and synergistically affect the ER of poetry, offering new perspectives for research in related fields.

Materials and methods

Participants

In our study, thirty-seven participants completed the experiment, including five males (22.00 ± 3.94 years) and thirty-two females (18.40 ± 1.81 years). Using G*Power calculations, a sample size of thirty-seven yielded a statistical power of 0.96 under an effect size of $f = 0.25$ and a level of $\alpha = 0.05$, demonstrating that the study was both practical and adequately powered.

All participants were right-handed, had normal or corrected vision without any visual impairments or color blindness and had none reported any history of neurological conditions or nerve-related disorders. The participants were all students from the School of Fine Arts, with Mandarin Chinese as their native language. They had received training in the poetry of Li Bai during their primary and secondary education in China, which enabled them to have a profound understanding of the rhythm, rhyme, and connotations of Li Bai's poetry (Li et al., 2009). Each participant was compensated with 10 yuan for their involvement after the experiment.

Ethics approval and consent to participate

This study was approved by the Institutional Review Board (IRB) of Northeastern University, China. Prior to participation, all participants provided written informed consent, confirming their voluntary involvement and full comprehension of their rights. All research procedures strictly adhered to the ethical standards outlined in the Belmont Report and the Declaration of Helsinki, in compliance with relevant ethical guidelines and regulations.

Stimuli and procedure

As shown in Figure 1, the study used a total of five poems by Li Bai. The poems were selected based on three criteria: first, relative consistency in emotional tone, choosing works with calm and serene tone to match the neutral reading style; second, thematic diversity, covering common emotional and philosophical themes in Li Bai's poetry, such as admiration of nature, homesickness, and friendship; third, poem length ranging between 20 and 30 Chinese characters, which ensured moderate length and comparable cognitive load across poems. Every poem was delivered by the same female voice, generated by AI with a neutral tone and no emotional inflections, including: Tianmen Mountain (天门山), Gazing upon the Waterfall at Mount Lu (望庐山瀑布), Reflections during a Tranquil Night (静夜思), A Farewell to Wang Lun (赠汪伦), and Bidding Meng Haoran Goodbye at Yellow Crane Tower (黄鹤楼送孟浩然之广陵). The audio files had an average duration of 7.15 s (SD = 0.83 s), an average sound pressure level (SPL) of 27.10 dB(A), and a speech rate of approximately 3.68 syllables per second.

As shown in Figure 2, five background images related to Li Bai's poems were utilized. The selected poems express Li Bai's deep love for mountains and rivers, his admiration for nature, his longing

天门中断楚江开 碧水东流至此回 两岸青山相对出 孤帆一片日边来
The Mount Tianmen Pass splits the Chu River wide,
Where jade waters eastward flow and then turn aside.
Green hills on either bank rise opposite and near,
A lone sail appears, approaching the sun's frontier.

日照香炉生紫烟 遥看瀑布挂前川 飞流直下三千尺 疑是银河落九天
Sunlight strikes the Incense Burner Peak, and purple smoke arises,
From afar the waterfall seems like a hanging veil across the stream.
Its torrent plunges straight down three thousand feet,
As if the Milky Way were falling from the ninth heaven.

床前明月光 疑是地上霜 举头望明月 低头思故乡
Before my bed, the moonlight glistens bright,
As if the ground were covered with frost's white.
I raise my head to gaze upon the moon,
Then lower it, recalling my hometown soon.

李白乘舟将欲行 忽闻岸上踏歌声 桃花潭水深千尺 不及汪伦送我情
Li Bai boarded the boat, preparing to depart,
When suddenly from shore came songs, filling his heart.
Though Peach Blossom Pool runs a thousand feet deep,
It cannot match the friendship Wang Lun did keep.

故人西辞黄鹤楼 烟花三月下扬州 孤帆远影碧空尽 唯见长江天际流
An old friend takes leave westward from Yellow Crane Tower,
In the misty, blossom-filled third month, bound for Yangzhou's bower.
A lone sail fades into the boundless azure sky,
Only the Yangtze flows on, merging with the horizon's eye.

FIGURE 1

Sample of poetry.

for his hometown, his profound feelings of friendship, and his reluctance to part with friends. Participants evaluated the extent to which they perceived the emotional dispositions of the poetry and achieved ER with the poet, using a seven-point scale ranging from complete disagreement to complete agreement. Finally, the background music consisted of traditional Chinese guqin music, including: The Resonance of the Qin and Soul (琴心共鸣), Hearing Rain Amidst the Hills (半山闻雨), Song of Qingwen (晴雯吟), The Orchid-Scented Ravine (兰香涧), and Seated Contemplation of the Misty Peaks (坐看云峰). All music was purely instrumental with a neutral emotional tone, avoiding interference with participants' emotions. The music had an average tempo of 101 beats per minute (BPM), and playback volume was controlled at 28.52 dB to maintain an acoustic environment consistent with the poetry recordings.

As shown in Figure 3, the experiment comprised two parts: the pilot study and the formal experiment. In the pilot study, to minimize the impact of existing knowledge and help soothe the participants' emotions, they were asked to familiarize themselves with the experiment in the same setting as the formal experiment. Subsequently, participants were asked to rate their familiarity with the textual poetry on a seven-point scale. In the formal experiment, participants completed the experiment based on five conditions.

Condition 1 (visual): with the five poetic texts in Figure 1, each trial randomly presented one on the computer screen.

Condition 2 (auditory): with the audio versions of the five poetic texts, each trial randomly presented one on the computer screen.

Condition 3 (visual-auditory): With five poetic texts in Figure 1 and the corresponding audio versions, each trial randomly presented one on the computer screen.

Condition 4 (visual-auditory and background music): With five poetic texts in Figure 1 and the corresponding audio versions, each



FIGURE 2 Sample of background images.

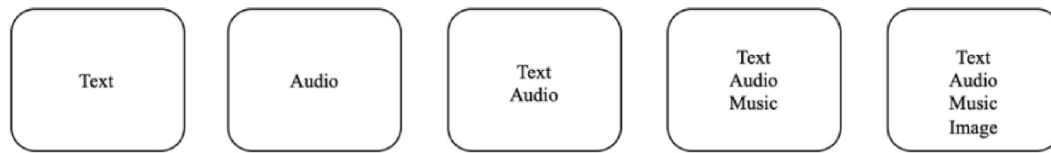


FIGURE 3 Five conditions of experimental procedure.

trial randomly presented one on the computer screen. Additionally, a randomly selected piece of background music from guqin music.

Condition 5 (visual-auditory, background music, and background image): With five poetic texts in Figure 1 and the corresponding audio versions, and the corresponding background image in Figure 2, each trial randomly presented one on the computer screen. Additionally, a randomly selected piece of background music from guqin music.

At the end of each trial, using a 7-point Likert scale, participants evaluated the poem’s ER. After submitting their response, they proceeded to the next trial. The experiment was carried out using Microsoft PowerPoint and automatically recorded via a linked Excel spreadsheet. All experimental were completed within a continuous time on the same day. Depending on individual pacing, the entire experiment took approximately 10 to 20 min. Throughout the experiment, silence was maintained. If participants had any questions, the experimenter focused on providing reassurance and encouraged participants to rate their feelings based on their genuine experiences. After the experiment concluded, five participants were randomly selected to inquire about their subjective feelings. Through verbal questioning, the experimenter recorded their main points in writing.

Results with poetic texts

Familiarity with poetic texts

On the familiarity evaluation of the textual poems, the participants’ average rating was 33.95 (SD = 2.32). This indicated that the participants had a high level of familiarity with the five poetic texts.

Correlation analysis

As shown in Table 1, the ER ratings exhibited strong positive correlations across the five conditions. Condition 1 and Condition

TABLE 1 Correlational analysis.

Condition	1	2	3	4	5
1					
2	0.600				
3	0.593	0.882			
4	0.679	0.602	0.701		
5	0.673	0.682	0.779	0.759	

2 showed a significant correlation ($r = 0.600, p < 0.01$), as did Condition 1 and Condition 3 ($r = 0.593, p < 0.01$), Condition 1 and Condition 4 ($r = 0.679, p < 0.01$), and Condition 1 and Condition 5 ($r = 0.673, p < 0.01$). Condition 2 and Condition 3 were also significantly correlated ($r = 0.882, p < 0.01$), as were Condition 2 and Condition 4 ($r = 0.602, p < 0.01$) and Condition 2 and Condition 5 ($r = 0.682, p < 0.01$). Similarly, Condition 3 and Condition 4 showed a significant correlation ($r = 0.701, p < 0.01$), as did Condition 3 and Condition 5 ($r = 0.779, p < 0.01$). Finally, Condition 4 and Condition 5 were significantly correlated ($r = 0.759, p < 0.01$). Overall, these results indicate that the emotional resonance of the poetry remained largely stable across different sensory modalities and background information conditions.

Analysis of variance

As shown in Table 2, a one-way repeated measures ANOVA indicated a significant effect across the five conditions [$F(4, 144) = 16.821, p < 0.001, \eta^2 = 0.318$] and *post hoc* comparisons were conducted. The ratings of ER were higher when participants were presented with Condition 1 than with Condition 2 (MD = 2.162, $p = 0.004$). The ratings were higher when participants were presented with Condition 3 than with Condition 2 (MD = 1.784, $p < 0.001$). No significant difference was found between Condition 1 and Condition 3 (MD = 0.378, $p = 0.563$). Overall, in line with hypothesis H1, audiovisual and visual modalities

TABLE 2 Paired comparisons.

Condition		MD	SE	P	CI	
					Lower	Upper
1	2	2.162	0.712	0.004	0.718	3.606
	3	0.378	0.648	0.563	-0.935	1.692
	4	-2.054	0.478	< 0.001	-3.023	-1.085
	5	-1.622	0.547	0.005	-2.731	-0.512
2	1	-2.162	0.712	0.004	-3.606	-0.718
	3	-1.784	0.419	< 0.001	-2.634	-0.933
	4	-4.216	0.714	< 0.001	-5.663	-2.769
	5	-3.784	0.661	< 0.001	-5.125	-2.443
3	1	-0.378	0.648	0.563	-1.692	0.935
	2	1.784	0.419	< 0.001	0.933	2.634
	4	-2.432	0.569	< 0.001	-3.586	-1.279
	5	-2.000	0.511	< 0.001	-3.037	-0.963
4	1	2.054	0.478	< 0.001	1.085	3.023
	2	4.216	0.714	< 0.001	2.769	5.663
	3	2.432	0.569	< 0.001	1.279	3.586
	5	0.432	0.480	0.374	-0.541	1.406
5	1	1.622	0.547	0.005	0.512	2.731
	2	3.784	0.661	< 0.001	2.443	5.125
	3	2.000	0.511	< 0.001	0.963	3.037
	4	-0.432	0.480	0.374	-1.406	0.541

resulted in higher ratings for ER compared to the auditory modality.

The ratings were higher when participants were presented with Condition 4 than with Condition 3 ($MD = 2.432$, $p < 0.001$), and also higher than Condition 1 ($MD = 2.054$, $p < 0.001$). Ratings were higher when participants were presented with Condition 5 than with Condition 3 ($MD = 2.000$, $p < 0.001$) and Condition 1 ($MD = 1.622$, $p = 0.005$). No significant difference was found between Condition 4 and Condition 5 ($MD = 0.432$, $p = 0.374$). Overall, consistent with hypotheses H2 and H3, background information beyond sensory modalities enhanced ratings of ER for the poetry (Supplementary Table 1).

Qualitative analysis

In line with hypotheses H1, H2, and H3, participants generally expressed that the textual poetry allowed them to more directly perceive the aesthetic features of the poetry, making it easier to evoke ER. In contrast, auditory forms of poetry, due to the transient nature of auditory information and the potential for attention to be diverted, made it difficult for participants to capture specific poetic content, with ER occurring more slowly. Furthermore, background music helped engage participants' emotions, while background images enhanced visual perception and increased the sense of immersion. This made the poetic imagery more vivid and intuitive, leading to stronger ER with the poetry.

Discussion

The impact of sensory modalities on the emotional resonance of poetry

In line with hypothesis H1, our study found that the audiovisual modality and the visual modality had a greater impact on ER in poetry compared to the auditory modality.

Visual modality

The first possible explanation is the objectivity of visual information, which can present the scenes and details in the poetry, allowing readers to more deeply feel the emotions conveyed in the text (Mehl et al., 2023). For example, in the sentence “天门中断楚江开，碧水东流至此回” (The Tianmen Pass splits the Chu River wide, Where jade waters eastward flow and then turn aside.), the word “断” (splits) directly separates the river, which helps readers immediately grasp the emotional disruption and fluctuation.

The second possible explanation is the richness of visual cues, which can stimulate readers' imagination and emotions (Jelinčić and Šveb, 2021). For instance, in the sentence “两岸青山相对出，孤帆一片日边来” (Green hills on either bank rise opposite and near, A lone sail appears, approaching the sun's frontier.), the imagery of green hills, lone sail, setting sun, and both banks provides rich visual elements, helping readers gain a deeper understanding of ER.

The third explanation may be that, during the process of reading a poetic text, readers construct a basic emotional framework based on their existing knowledge. Textual reading allows readers ample time to reflect on the connotations of the poem. The ancient Chinese philosopher Liu Xie, in his work *The Literary Mind and the Carving of Dragons*, states, “Emotion is the foundation of literature, while words are the threads of reasoning; when the foundation is correct, the threads follow, and when the reasoning is settled, the words flow” (Liu X. L., 2022). The text itself contains a wealth of emotional elements, and in the process of interpreting the poem, readers also experience a process of emotional sublimation. For instance, in the sentence “孤帆远影碧空尽, 唯见长江天际流” (Green hills on either bank rise opposite and near, A lone sail appears, approaching the sun’s frontier.), the lone sail and faraway blue sky in traditional Chinese culture represent feelings of coldness and solitude, while the use of “唯见” (appears) in the contrast highlights the poet’s melancholic mood, revealing a sense of solitude and loss. Therefore, textual reading and textual analysis occur almost simultaneously, making it easier to evoke ER in readers.

The fourth possibility lies in the richness of the meanings of Chinese characters within the visual system. The same character can express different emotions and imagery depending on the poet’s creation. In textual reading, the words are more likely to evoke associations and thoughts in readers, thus generating a stronger emotional response (Guo, 2018). For example, in the sentence “两岸青山相对出, 孤帆一片日边来” (Green hills on either bank rise opposite and near, A lone sail appears, approaching the sun’s frontier.), words like both banks, green hills, lone sail, and setting sun may be neutral in isolation, but when combined with the imagery, they convey a profound emotional depth.

Auditory modality

The first possible explanation is the subjectivity of auditory information. The auditory channel primarily uses sound and rhythm to express emotions (Belfi, 2019), which may have a lesser impact on ER. For example, in the sentence “日照香炉生紫烟, 遥看瀑布挂前川” (Sunlight strikes the Incense Burner Peak, and purple smoke arises, From afar the waterfall seems like a hanging veil across the stream.), the word 香炉 (Burner Peak) might be overlooked during auditory processing of information, and different individuals may have different emotional reactions to the same sound or rhythm. Some might associate the neutral word 香炉 with a certain scene, while others may have no associations at all.

The second possible explanation is the semantic nature of auditory information. Processing semantic information requires more cognitive resources. The brain must transform and understand the content of the poetry, which takes time. In this process, if poetry is presented quickly in an auditory form, it can make it more difficult for individuals to grasp the underlying emotions and form ER (Pişur and Miu, 2025). For example, in the sentence “飞流直下三千尺, 疑是银河落九天” (Its torrent plunges straight down three thousand feet, As if the Milky Way were falling from the ninth heaven.), the scene of 落九天 (straight down three thousand feet) appears unclear in the auditory format, and individuals may find it difficult to form a clear mental image, resulting in a more neutral ER.

Visual-auditory modality

In the visual-auditory modality, the integration of visual and audio may improve comprehension of the poem. The visual modality focuses on spatial structure and imagery, while the auditory modality emphasizes temporal changes and sound features. Together, they provide a more comprehensive understanding of the poem (Coutinho and Scherer, 2017). For example, in the sentence “床前明月光, 疑是地上霜” (Before my bed, the moonlight glistens bright, As if the ground were covered with frost’s white.), the visual imagery of the bright moon becomes more vivid, while the passage of time in the auditory channel adds an additional dimension. Together, they contribute to the deep emotional layers of the poem.

The impact of background information on the emotional resonance of poetry

In line with hypothesis H2, our study found that both background music and background images enhanced the ER of the poetry.

Imagery interpretation and cultural empathy

During the process of interpreting poetry, readers encounter keywords that carry the emotional essence of the entire poem. The surrounding imagery serves to enhance this emotional focal point (Li, 2012). For example, in the sentence “举头望明月, 低头思故乡” (I raise my head to gaze upon the moon, Then lower it, recalling my hometown soon.), the emotional core lies in the word “思” (gaze). The imagery of the moon and bowing my head is recreated in the background images, which deepens the interpretation of the imagery and highlight the emotional point. This progression moves from basic reading and understanding of the text to a more holistic perception of the work. It deepens the higher-level understanding of the poem, thereby fostering a deeper comprehension of the poem.

Imagery representation, and cultural metaphors

These images are carriers of the poet’s emotions and thoughts, containing rich cultural backgrounds and symbolic meanings (Kim and Wedell, 2016). The poet’s time, environment, and influences shape the creation of these images. The imagery in poetry goes beyond words, and background images serve to re-present these poetic visuals, blending the poet’s inner world with the surrounding environment. This enhances the emotional expression of the poetry, making it easier for readers to ER (Marin and Leder, 2018). For instance, in the sentence “举头望明月, 低头思故乡” (I raise my head to gaze upon the moon, Then lower it, recalling my hometown soon.), the imagery of the bright moon and the action of looking up are clearly depicted. The cultural metaphor of the moon symbolizing longing and reunion is made tangible, which helps readers better understand the cultural meaning of the poem and increases their emotional connection with the text.

Emotional deepening, and cultural fusion

Playing ancient Chinese guqin music from the Tang Dynasty while reading the poetry transports individuals back to the

era of Li Bai, allowing them to immerse themselves more deeply in the poet's time (Bennett and Ginsborg, 2018). This is not just an auditory experience but a profound spiritual journey. The guqin music itself represents purity and transcendence, symbolizing the noble and refined qualities of the time. It enhances the emotional depth of the poem by enriching its artistic expression, representing both poetic sentiment and beauty (Menninghaus et al., 2018). During the reading process, the rhythm, melody, and tranquility of the music allow the reader to feel the depth and richness of traditional Chinese culture, making it easier to understand the poetic imagery and atmosphere, thus enhancing their comprehension of the poem's artistic essence (Juslin et al., 2023).

The impact of sensory modalities and background information on the emotional resonance of poetry

Situated conceptualization

In line with hypothesis H3, according to the situated conceptualization theory, individuals construct psychological representations of time, space, entities, and their relationships while understanding and experiencing poetry (Antović, 2021). In the process of appreciating poetry, sensory modalities provide rich perceptual dimensions that form the preliminary imagery of the poem. Background information, on the other hand, provides insights into the poet's era and cultural background, offering a broader framework. This allows the poem to be interpreted within a broader cultural background, thereby deepening its emotional meaning. For example, in the sentence “故人西辞黄鹤楼，烟花三月下扬州” (An old friend takes leave westward from Yellow Crane Tower, In the misty, blossom-filled third month, bound for Yangzhou's bower.), when readers see or hear the term “黄鹤楼” (Yellow Crane Tower), there is an immediate emotional recognition. As a famous cultural landmark in China, the Yellow Crane Tower transcends mere scenery and carries deep cultural significance, representing a longing for home. When the imagery of the Yellow Crane Tower reappeared in the background image, readers were instantly transported back to the poet's time. Coupled with the serene sound of guqin music, readers felt fully immersed in Li Bai's era, completely entering the scene. This fusion of sensory and cultural background enhanced ER, allowing the reader to experience the poem's emotional depth to its fullest.

Multisensory synesthesia

In classical Chinese aesthetics, synesthesia refers to the mutual blending, transformation, and fusion between different sensory experiences, which enhances readers' ER with imagery (Sheng and Wu, 2024). The creation of artistic conception in poetry often relies on the integration of sound (specifically music in this study), color (specifically background images here), and meaning (the inherent connotations and emotions of the poem itself), achieving the artistic conception of painting in poetry, sound in painting, and emotion in sound (Liu, 2021). For example, in the sentence “天门中斷楚江開，碧水東流至此回” (The Mount Tianmen

Pass splits the Chu River wide, Where jade waters eastward flow and then turn aside.), the emotions and imagery carried by “Mount Tianmen” are guided by background information. When the distant and clear sound of guqin music and the imagery of Mount Tianmen in the background image reappear, readers not only see the landscape depicted in the poem but also hear the emotional rhythm behind it. This aesthetic experience not only deepens the understanding of the poem's emotional meaning but also reflects a return to traditional Chinese aesthetics. Although the term synesthesia often holds explanatory power in cross-cultural aesthetic studies (non-native Chinese speakers) as well (Simner, 2005), we can further explore this in future research.

Conclusion

In the contrast between visual and auditory modalities, the results indicate that the visual modality, through its direct imagery presentation, significantly dominates the activation of ER, and the ER of poetry depends not only on the performance of sensory modalities but also on the collaborative effect of different background information, highlighting the integrative and layered nature of perceptual experience. This finding deepens our understanding of the core role of cross-modal interaction in ER.

Limitation

Despite this study's exploration of different sensory modalities and background information on ER in poetry, several limitations remain. The experimental design includes five conditions, and the sample size is relatively limited, which may restrict the generalizability of the results. Future research can expand the sample size and include representative groups across different age ranges and populations to enhance the external validity of the findings. Additionally, the types of background music and background images used in this study are relatively limited and do not cover a richer variety of background information. Subsequent studies should consider introducing diversified audiovisual background materials to further investigate whether subtle differences among different background information exist. Finally, ER exhibits strong individual variability. Future research needs to explore individual characteristics (such as cultural background and personal experiences) as moderating factors to gain a more comprehensive understanding of this complex psychological process.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Northeastern University in China. The studies were conducted in

accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

YZ: Writing – review and editing. XW: Writing – original draft.

Funding

The authors declare that financial support was received for the research and/or publication of this article. Ministry of Education Humanities and Social Sciences Fund 22YJA760109, and The Fundamental Research Funds for the Central Universities N2422002.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Alshaer, A. (2025). “The role of sensory experiences in evoking emotional responses in virtual reality,” in *Umm Al-Qura University Journal of Engineering & Architecture*, Vol. 16, ed. H. M. Youssef (Berlin: Springer Nature).
- Antović, M. (2021). Multilevel grounded semantics across cognitive modalities: Music, vision, poetry. *Lang. Literature* 30, 147–173. doi: 10.1177/0963947021999182
- Beebe, B., Lachmann, F., and Jaffe, J. (1997). Mother–infant interaction structures and. presymbolic self-and object representations. *Psychoanal. Dialogues* 7, 133–182. doi: 10.1080/10481889709539172
- Belfi, A. M. (2019). Emotional valence and vividness of imagery predict aesthetic appeal in music. *Psychomusicol. Music Mind Brain* 29, 128–135. doi: 10.1037/pmu0000232
- Bennett, D., and Ginsborg, J. (2018). Audience reactions to the program notes of unfamiliar music. *Psychol. Music* 46, 588–605. doi: 10.1177/0305735617721339
- Bishop, J. (2019). Lyric poetry: Resonance and attunement. *Axon* 9, 1–11.
- Blizsek, W. (1973). Aesthetics and Psychobiology by DE Berlyne. *J. Aesthetics Art Criticism* 31:553. doi: 10.2307/429334
- Coburn, W. J. (2001). Subjectivity, emotional resonance, and the sense of the real. *Psychoanal. Psychol.* 18:303. doi: 10.1037//0736-9735.18.2.303
- Coutinho, E., and Scherer, K. R. (2017). The effect of context and audio-visual modality on emotions elicited by a musical performance. *Psychol. Music* 45, 550–569. doi: 10.1177/0305735616670496
- Frame, J., Mehl, K., Head, K., and Belfi, A. M. (2024). The influence of sensory modality on aesthetic judgments of poetry. *Psychol. Aesthetics Creativity Arts* doi: 10.1037/aca0000673 Advance online publication.
- Graber, K. J., and Sumera, M. (2020). Interpretation, resonance, embodiment: Affect theory and ethnomusicology. *Ethnomusicol. Forum* 29, 3–20. doi: 10.1080/17411912.2020.1808501
- Gu, X. (2024). Enhancing social media engagement using AI-modified background music: Examining the roles of event relevance, lyric resonance, AI-singer origins, audience interpretation, emotional resonance, and social media engagement. *Front. Psychol.* 15:1267516. doi: 10.3389/fpsyg.2024.1267516
- Guo, S. (2024). Emotional expression in artworks and psychological reactions of audiences. *J. Art Cult. Philosophical Stud.* 1, 1–6. doi: 10.70767/jacps.v1i3.388
- Guo, X. H. (2018). Reflections on teaching “Chinese classical poetry” in primary school Chinese language classes. *Teenagers* 2, 42–43.
- Jelinčić, D. A., and Šveb, M. (2021). The influence of visual stimuli on emotions in the creation of cultural and tourist experiences. *Acta Turistica* 33, 39–74. doi: 10.22598/at/2021.33.1.39
- Johnson-Laird, P. N., and Oatley, K. (2022). How poetry evokes emotions. *Acta Psychol.* 224:103506. doi: 10.1016/j.actpsy.2022.103506
- Juslin, P. N., Ingmar, E., and Danielsson, J. (2023). Aesthetic judgments of music: Reliability, consistency, criteria, self-insight, and expertise. *Psych. Aesthet. CreativityArts* 17, 193–210. doi: 10.1037/aca0000403
- Kim, J., and Wedell, D. H. (2016). Comparison of physiological responses to affect eliciting pictures and music. *Int. J. Psychophysiol.* 101, 9–17. doi: 10.1016/j.ijpsycho.2015.12.011
- Küssner, M. B., and Eerola, T. (2019). The content and functions of vivid and soothing visual imagery during music listening: Findings from a survey study. *Psychomusicol. Music MindBrain* 29, 90–99. doi: 10.1037/pmu0000238
- Langacker, R. W. (2002). *Concept, Image, and Symbol*. Berlin: Walter de Gruyter Inc.
- Leder, H., Gerger, G., Briber, D., and Schwarz, N. (2014). What makes an art expert? Emotion and evaluation in art appreciation. *Cogn. Emot.* 28, 37–41. doi: 10.1080/02699931.2013.870132
- Li, F. Y., Gu, Q. Y., and Li, W. J. (2009). Primary school students’ implicit learning of rules about poetry’s level and oblique tones. *Psychol. Sci.* 5, 1255–1258.
- Li, Y. (2012). *Thoroughly Understanding the Key Words in Classical Poetry. Reference for Middle School Chinese Language Teaching: Senior High Edition*. 54–55.
- Liu, T. (2021). The poetic form construction of contemporary Chinese new poetry’s aesthetic beauty. *Beijing Cult. Creative* 2, 31–37.
- Liu, X. L. (2022). How to guide students to appreciate emotions in poetry: A case study of Du Fu’s poetry. *Gansu Educ.* 19, 99–102.
- Marin, M. M., and Leder, H. (2018). Exploring aesthetic experiences of females: Affect-related traits predict complexity and arousal responses to music and affective pictures. *Pers. Individ. Dif.* 125, 80–90. doi: 10.1016/j.paid.2017.12.027
- Mehl, K., Gugliano, M., and Belfi, A. M. (2023). The role of imagery and emotion in the aesthetic appeal of music, poetry, and paintings. *Psychol. Aesthet. CreativityArts* 13, 259–260. doi: 10.1037/aca0000245
- Menningshaus, W., Wagner, V., Knoop, C. A., and Scharinger, M. (2018). Poetic speech melody: A crucial link between music and language. *PLoS One* 13:e0205980. doi: 10.1371/journal.pone.0205980

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2025.1541680/full#supplementary-material>

- Miller, V. (2015). Resonance as a social phenomenon. *Sociol. Res. Online* 20, 58–70. doi: 10.5153/sro.3557
- O'Connor, J. (2024). *Culture is Not An Industry: Reclaiming Art and Culture for the Common Good*. Manchester: Manchester University Press.
- Paiva, D. (2020). Poetry as a resonant method for multi-sensory research. *Emot. Space Soc.* 34:100655. doi: 10.1016/j.emospa.2020.100655
- Pan, F., Zhang, L., Ou, Y., and Zhang, X. (2019). The audio-visual integration effect on music emotion: Behavioral and physiological evidence. *PLoS One* 14:e0217040. doi: 10.1371/journal.pone.0217040
- Pițur, S., and Miu, A. C. (2025). Poetry-elicited emotions: Reading experience and psychological mechanisms. *Psychol. Aesthet. CreativityArts* 19, 218–231. doi: 10.1037/aca0000525
- Santos, D., Requero, B., and Martín-Fernández, M. (2021). Individual differences in thinking style and dealing with contradiction: The mediating role of mixed emotions. *PLoS One* 16:e0257864. doi: 10.1371/journal.pone.0257864
- Sheng, X., and Wu, X. (2024). Poetry and painting: The artistic expression of synaesthesia in Tang poetry and painting. *Art Educ. Res.* 6, 14–16. doi: 10.3969/j.issn.1674-9286.2024.06.010
- Silvia, P. (2005). Emotional responses to art: From collation and arousal to cognition and emotion. *Rev. General Psychol.* 9, 342–357. doi: 10.1037/1089-2680.9.4.342
- Simner, J. (2005). Synesthesia: Perspectives from cognitive neuroscience. *J. Consciousness Stud.* 12, 141–143.
- Stolorow, R. D., and Atwood, G. E. (2014). *Contexts of Being: The Intersubjective Foundations of Psychological Life*. Milton Park: Routledge.
- Thompson, W. F., Russo, F. A., and Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cogn. Emot.* 22, 1457–1470. doi: 10.1080/02699930701813974
- Wagner, V., Scharinger, M., Knoop, C. A., and Menninghaus, W. (2021). Effects of continuous self-reporting on aesthetic evaluation and emotional responses. *Poetics* 85:101497. doi: 10.1016/j.poetic.2020.101497
- Wang, Z., and Wang, Q. (2024). Research on the influencing factors and common laws of emotional resonance in visual art design. *Pak. J. Life Soc. Sci.* 22, 923–932. doi: 10.57239/PJLSS-2024-22.2.0065
- Yuan, Y., and Guoyuan, T. (2022). An empirical study on imagery and emotional response in Chinese poetry translation—The visual grammar perspective. *Front. Psychol.* 13:872497. doi: 10.3389/fpsyg.2022.872497
- Zhao, Y. M. (2024). The perfect integration of poetry and music in the context of literature and art: A review of opera: A dialogue between music and poetry. *Yuwen Construction* 13:10008. doi: 10.16412/j.cnki.1001-8476.2024.13.025