humanities

MDPI

Article

Toward a Neuro-*Yijing* Model: Musical Imagery, Affect, and Reflective Experience in Chinese Programmatic Music

Dingyuan Zhang

School of Music, East China Normal University, 500 Dongchuan Road, Minhang District, Shanghai 200241, China; 13621921976@163.com

Abstract

Neuroaesthetic research has often examined aesthetic experience through measurable categories such as beauty, pleasure, preference, prediction, and reward. Although these approaches have produced important insights, they do not fully explain how culturally situated musical experience may be organized around sound-induced scene formation, atmosphere, temporal openness, emotional restraint, memory, and reflective awareness. This article brings the Chinese aesthetic concept of *yijing* into dialogue with neuroaesthetics and music cognition and proposes a preliminary Neuro-*Yijing* Model as a conceptual and hypothesis-generating framework. The model comprises three interrelated and recursive dimensions: perceptual–imaginal formation, affective resonance, and reflective integration. Through a conceptually guided analysis of Xiyang Xiaogu (Evening Drum and Flute at Sunset), anchored in a publicly available 1999 pipa performance by Wu Man and its programmatic notes, the article distinguishes between sound-based affordances and verbal–cultural framing in the formation of imagined scene and affective meaning. The term "Neuro" is used here not to claim an identified neural basis for *yijing*, but to position *yijing* as a cultural–aesthetic framework for refining questions about imagery, embodied listening, affect, memory, and self-referential cognition. The model's added value lies in treating sound–scene transformation, culturally mediated affect, and contemplative completion as a unified but testable process rather than as isolated perceptual, emotional, or self-related components.

Keywords: *yijing*; Chinese programmatic music; musical imagery; affect; neuroaesthetics; aesthetic cognition; cross-cultural aesthetics; *Xiyang Xiaogu*

1. Introduction

Over the past three decades, neuroaesthetics has examined the neural and cognitive processes underlying aesthetic experience. Work by Semir Zeki and Anjan Chatterjee helped shift aesthetic inquiry toward questions of perception, emotion, valuation, attention, embodiment, and meaning-making (Chatterjee 2011; Chatterjee and Vartanian 2014; Zeki 1999). Contemporary research no longer locates aesthetic response in a single "beauty centre," but understands it through distributed interactions among sensory, reward, affective, memory-related, and self-referential systems (Vessel et al. 2012). In music, these interactions are further shaped by temporal unfolding, auditory imagery, expectation, embodied listening, affective memory, and culturally learned interpretation (Vuust et al. 2022).

Yet the categories used in neuroaesthetic research influence what counts as an aesthetic outcome. Beauty, liking, preference, pleasure, prediction, reward, symmetry, and

Check for updates

Received: 27 June 2026
Revised: 31 August 2026
Accepted: 2 September 2026
Published: 4 September 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

fluency are useful operational categories, but they do not exhaust aesthetic experience. Music also unfolds through atmosphere, suggestion, memory, visual–spatial imagery, bodily simulation, silence, and reflective attention. The issue is not that existing models are incapable of addressing these dimensions; rather, they often do not organize them through culturally specific aesthetic vocabularies that make these experiences theoretically central.

This article addresses this issue by bringing Chinese musical aesthetics into dialogue with contemporary neuroaesthetics. Its focus is the concept of *yijing* (意境), a central category in Chinese aesthetic thought that is often translated as "artistic conception," "artistic realm," or "poetic mood," but which exceeds any single English equivalent. In the context of music, *yijing* may be understood as a dynamic configuration of sound, imagined scene, emotion, memory, and reflective awareness. It does not refer simply to the mood expressed by a work, nor to a fixed image represented by musical sound. Rather, it describes an experiential process through which auditory materials evoke inner scenes, these scenes become affectively charged, and affective resonance in turn reshapes the listener's perception of the music (Jing 2023). In this sense, *yijing* offers a culturally situated model for thinking about how music generates meaning through the interaction of perception, imagination, emotion, and reflection.

Cross-modal imagery, embodied simulation, musical expectation, affective memory, and reflective listening are already discussed in music cognition and neuroaesthetics (Taruffi and Küssner 2019; Pearson 2019; Gordon et al. 2018). The article therefore treats *yijing* as a historically developed lens for organizing these processes around sound–scene transformation, affective resonance, and contemplative integration, rather than as evidence for a culturally unique neural mechanism.

On this basis, the article proposes a preliminary Neuro-*Yijing* Model as a hypothesis-generating framework. "Neuro" does not mean that the neural basis of *yijing* has already been identified. It indicates that the model connects a Chinese aesthetic concept with questions in neuroaesthetics, music cognition, auditory imagery, multisensory association, embodied listening, emotion, memory, and self-referential processing. The model consists of three interrelated dimensions: perceptual–imaginal formation, in which musical sound prompts visual–spatial and atmospheric imagery; affective resonance, in which imagery becomes emotionally meaningful; and reflective integration, in which memory, self-reference, and contemplative awareness contribute to aesthetic completion. These dimensions are not rigid stages but interacting processes that can be examined and challenged in future empirical work.

To clarify the analytical value of this model, the article examines *Xiyang Xiaogu* (《夕阳箫鼓》, Evening Drum and Flute at Sunset) as a case study. The work is especially relevant to the present inquiry because its title, musical imagery, timbral associations, pacing, and temporal unfolding explicitly invite listeners to experience music as a transformation of sound into landscape. Rather than treating *Xiyang Xiaogu* as representative of all Chinese music, this article approaches it as a focused example through which the proposed framework can be examined conceptually. The analysis considers how melodic contour, rhythmic flexibility, textural spacing, timbral contrast, silence, and programmatic framing may support the emergence of *yijing* as a listening experience.

The article contributes in four ways. First, it reconsiders neuroaesthetic models from the perspective of cross-cultural musical experience without reducing the issue to a "Western" versus "Chinese" opposition. Second, it develops *yijing* as a historically situated concept for asking more precise questions about musical imagery, affective resonance, temporality, and reflective listening. Third, it clarifies the boundaries and testable implications of the proposed framework. Fourth, it uses Xiyang Xiaogu to show how musical sound and

https://doi.org/10.3390/h15090134

programmatic language jointly shape the movement from auditory perception to imagined scene and reflective meaning.

2. Conceptual and Analytical Approach

This study adopts a conceptual, interpretive, and hypothesis-generating approach rather than an empirical design. Literature selection was theory-driven and purposive rather than systematic: sources were chosen for their direct relevance to the concepts and models compared here, not to provide an exhaustive review of each field. The model was developed through three linked steps. First, the article synthesizes research in neuroaesthetics, music cognition, music-evoked emotion, auditory and visual mental imagery, embodied music cognition, cross-modal association, and self-referential processing. Foundational neuroaesthetic and music–cognitive models were selected because they define the categories against which the proposed model is positioned; more recent work on music-evoked visual imagery, cross-modal association, and motor engagement was added to strengthen the basis for perceptual–imaginal formation (Taruffi and Küssner 2019; Dahl et al. 2023; Wanke et al. 2025; Gordon et al. 2018).

Second, the article develops *yijing* as a culturally situated analytical concept rather than as a comprehensive label for all aesthetic experience. The three dimensions were derived by aligning recurrent features of *yijing* discourse—scene formation, qing–jing relation, resonance, emptiness, temporality, and inward reflection—with corresponding questions in music cognition and neuroaesthetics. The retained three-part structure was preferred over plausible alternatives, including a two-part structure that would combine affective resonance with reflective integration and a four-part structure that would separate memory/self-reference as an additional dimension. The three-part model was retained because it best preserves the main analytic movement of the argument: from auditory cue to imagined scene, from scene to affective resonance, and from resonance to reflective integration. These dimensions are heuristic and recursive rather than exhaustive.

Third, the case analysis uses the publicly available Smithsonian recording of Wu Man's 1999 Freer Gallery performance as the reference performance, together with the accompanying program notes (Smithsonian's National Museum of Asian Art n.d.). The source lists Xi Yang Siao Gu as occurring from 0:00 to 8:22 and provides eight programmatic subtitles associated with Li Fangyuan's 1895 New Collections. Because the source does not provide internal timing for each subtitle and because the work circulates in multiple versions, the analysis does not make measure-specific claims. Instead, it uses three broad analytical regions within the 0:00–8:22 performance—opening distance and moonrise, middle water/cloud/fishing imagery, and later waves/return—as listening units.

To limit selective interpretation, the analysis separates two kinds of evidence. Sound-based affordances include register, attack, decay, timbre, contour, rhythmic density, silence, pacing, texture, and resonance. Framing affordances include title, section subtitles, program notes, and cultural knowledge. The model is supported only where these levels can be shown to interact; it would be weakened if listeners' imagery were explained entirely by verbal framing or entirely by general acoustic affect without culturally specific organization.

3. Dominant Neuroaesthetic Models: Contributions and Cross-Cultural Limits

Neuroaesthetic inquiry is shaped by the categories through which aesthetic experience is defined. This section reviews influential models in neuroaesthetics and music cognition to clarify their contributions and limits for the present study. It does not oppose "Western" and "Chinese" aesthetics but asks how existing categories may be expanded when applied to culturally specific musical experience.

Early neuroaesthetic research made a major contribution by bringing aesthetic experience into dialogue with the neuroscientific study of perception and valuation. Zeki's work on the "visual brain" emphasized that aesthetic perception is closely related to the functional organization of the visual cortex and to the brain's capacity to process perceptual features such as form, color, and motion (Zeki 1999). Subsequent empirical work by Kawabata and Zeki further suggested that the experience of beauty may involve activity in the medial orbitofrontal cortex, a region associated with subjective valuation and affective reward (Kawabata and Zeki 2004). These studies were important because they helped transform aesthetic experience from a topic addressed primarily through philosophical reflection into one that could also be investigated through observable neural processes.

At the same time, later models broadened the field beyond the search for localized neural correlates of beauty. Chatterjee's "aesthetic triad" model conceptualizes aesthetic experience as the interaction of three partially overlapping systems: a sensory–motor system involved in perceptual and formal processing; an emotion–valuation system associated with affective and reward-related responses; and a meaning–knowledge system that incorporates memory, expertise, context, and cultural understanding (Chatterjee 2011; Chatterjee and Vartanian 2014). Similarly, psychological models of aesthetic appreciation have emphasized that aesthetic judgment develops through multiple levels of processing, including perceptual analysis, implicit memory integration, explicit classification, cognitive mastering, and evaluation (Leder et al. 2004). These models are valuable because they show that aesthetic experience is not a simple stimulus–response event, but a complex process involving perception, affect, cognition, memory, and interpretation.

In music research, this distributed understanding of aesthetic experience has been especially productive. Neuroimaging studies have shown that intense musical pleasure may involve interactions among auditory cortices, limbic and paralimbic structures, and reward-related regions such as the ventral striatum and nucleus accumbens (Blood and Zatorre 2001; Salimpoor et al. 2011; Zatorre and Salimpoor 2013). Broader accounts of music-evoked emotion have further demonstrated that musical experience depends on multiple mechanisms, including expectation, memory, imagery, emotional contagion, rhythmic entrainment, and embodied response (Brattico and Pearce 2013; Juslin and Västfjäll 2008; Koelsch 2014). These findings indicate that musical aesthetics cannot be reduced to sensory pleasure alone. Music becomes aesthetically meaningful through the interaction of sound perception, temporal anticipation, affective response, memory, bodily simulation, and cultural interpretation.

Several limits appear when these models are applied to cross-cultural musical aesthetics. The first is not that neuroaesthetics ignores imagery, memory, affect, or context. Many existing models explicitly include these processes. The limitation is that empirical studies often operationalize aesthetic response through variables such as beauty, liking, pleasure, preference, and reward, while many traditions also value ambiguity, suggestion, restraint, atmosphere, spatial imagination, and contemplative resonance. In Chinese programmatic music, aesthetic meaning may depend less on immediate hedonic response than on the gradual formation of inner scenes and affective associations.

The second limitation concerns modality and temporality. Much early neuroaesthetic work developed through studies of visual art, facial attractiveness, symmetry, and perceptual form. This visual orientation has generated important insights, but it can be less sufficient for music, whose meaning unfolds through time rather than appearing all at once. Musical aesthetic experience depends on duration, recurrence, delay, silence, tension, release, and transformation. It also often involves auditory imagery, embodied expectation, and memory-based association. For this reason, a model adequate to musical aesthetics

must account not only for the perception of formal features but also for how listeners construct meaning across time.

Recent empirical work makes this point increasingly important. Music-evoked visual mental imagery has been shown to be common during listening and to vary with musical and acoustic features, individual differences, and emotional response (Taruffi and Küssner 2019; Dahl et al. 2023). Research on extramusical information also indicates that titles and contextual descriptions can shape emotion and interpretation (Vuoskoski and Eerola 2015). Studies of audiovisual association and auditory morphodynamics further suggest that listeners may map sonic features onto visual–spatial forms in systematic ways (Wanke et al. 2025). These findings support the model's first dimension while also showing why the paper should speak of cross-modal imagery rather than synesthesia in the strict neuropsychological sense.

The third limitation concerns the relationship between neural evidence and aesthetic meaning. Although localized activations in reward-related or valuation-related regions provide valuable information, aesthetic consciousness is unlikely to be fully explained by isolated brain areas. It is better understood as a dynamic process involving distributed interactions among perceptual systems, affective networks, memory mechanisms, attentional control, embodied simulation, and self-referential thought (Chatterjee and Vartanian 2014). This point is especially relevant to musical experiences that are meditative, nostalgic, spiritually suggestive, or emotionally mixed. Such experiences may involve reward, but their aesthetic force often lies in the listener's capacity to form associations, inhabit an imagined scene, and integrate sound with memory and reflective awareness.

These limits are not failures of existing models; their contributions remain foundational. They show that aesthetic experience is distributed, interactive, and context-sensitive. What remains less developed is an account of how different aesthetic traditions organize these shared processes through different conceptual vocabularies. Such concepts do not merely label experience; they shape attention, imagination, feeling, and interpretation.

This is where the concept of *yijing* becomes relevant. *Yijing* does not replace models based on perception, emotion, valuation, memory, or embodied response. Rather, it offers a historically situated aesthetic framework through which these processes can be reorganized around the relation between sound, image, affect, memory, and reflective awareness. In this sense, *yijing* can help refine neuroaesthetic inquiry by directing attention to forms of musical experience that are not fully captured by the categories of beauty, pleasure, or preference alone.

4. *Yijing*, Yin–Yang, and Chinese Aesthetic Cognition

Within Chinese aesthetic thought, *yijing* (意境) describes the relation between artistic form, emotion, imagination, and reflective experience. Often translated as "artistic conception," "artistic realm," or "poetic mood," it is better understood as an experiential field in which sensory perception, imagined scene, affective resonance, and reflection interact. It belongs neither solely to the artwork nor solely to the perceiver, but emerges through artistic materials, imaginative participation, and cultural–aesthetic expectation (Jing 2023; Li 2010; Yu 1987).

For the purposes of this article, *yijing* is treated not as a static category but as an experiential process in which emotion and scene become mutually expressive. The scene is not a neutral background for emotion, nor is emotion simply projected onto the scene. Instead, the two are understood as interdependent within aesthetic experience. This working definition also emphasizes movement, emptiness, relational space, rhythm, and the vitality of lived experience. *Yijing* therefore names a process through which artistic form organizes perception into an internally resonant world (Jing 2023; Jullien 2012; Li 2010).

It is also necessary to distinguish *yijing* from several related concepts. Imagery refers primarily to internal sensory or visual–spatial representation. Qingjing jiaorong (情景交融) refers more specifically to the interpenetration of emotion and scene. Jingjie (境界) may indicate a broader aesthetic or spiritual state. *Yijing* overlaps with these ideas but is not identical to any one of them. In this article, *yijing* refers to a process in which sound, imagined scene, affective resonance, memory, and reflection become organized into an aesthetically meaningful experiential field.

This process is especially significant for music. Unlike painting or poetry, music does not present visual images directly. Its materials are temporal, sonic, and bodily: pitch, rhythm, timbre, register, gesture, resonance, silence, and formal unfolding. Yet music can nevertheless generate powerful imagistic experience. In Chinese programmatic music, titles, melodic gestures, timbral associations, and rhythmic pacing often guide listeners toward imagined landscapes or atmospheres. The listener does not simply decode a fixed meaning already contained in the score. Rather, meaning emerges through a process of association, memory, affective participation, and inner visualization. From this perspective, *yijing* in music may be understood as a process of sound–image transformation: musical sound creates the conditions under which an inner scene becomes experientially present (Yu 1987; Cowen et al. 2020).

This understanding of *yijing* provides a useful bridge between Chinese aesthetics and cognitive neuroscience, but the connection must be made cautiously. It would be inappropriate to claim that *yijing* corresponds to a unique or culturally exclusive neural mechanism. Cross-modal imagery, embodied listening, affective memory, and reflective attention are not limited to Chinese music, and they have been widely discussed in music cognition and neuroaesthetics. The more defensible claim is that *yijing* offers a culturally specific conceptual framework for organizing these processes. It foregrounds the movement from sound to image, from image to affect, and from affect to reflective integration. In doing so, it helps formulate neuroaesthetic questions that are not fully captured by categories such as beauty, pleasure, or preference alone.

Neuroaesthetic research increasingly treats aesthetic experience as distributed and context-sensitive, involving sensory, affective, valuation-related, memory-related, and self-referential systems (Brown et al. 2011; Chatterjee and Vartanian 2014). *Yijing* can be approached as a theoretical model for how such systems may interact during listening: auditory perception may connect with visual–spatial imagination and memory; imagined scenes may acquire affective meaning; and inward resonance may involve self-referential and reflective processes. These remain hypotheses, not established neural correspondences.

The default mode network offers one cautious reference point for the reflective dimension of *yijing*. It has been associated with internally oriented cognition, autobiographical memory, mental simulation, and self-generated thought (Andrews-Hanna et al. 2014). Neuroaesthetic work also suggests that personally meaningful artworks may involve default-mode activity (Vessel et al. 2012). These findings do not prove that *yijing* activates the default mode network, but they support questions about inward imagination, memory, self-related association, and reflection in *yijing*-oriented listening.

The concept of *yin–yang* (阴阳) further clarifies the dynamic structure of this aesthetic experience. In traditional Chinese thought, yin and yang do not represent fixed binary oppositions, but mutually dependent and continuously transforming tendencies. When applied to music, this logic can be understood through relationships such as sound and silence, motion and stillness, fullness and emptiness, density and sparsity, tension and release, and presence and absence. These relationships are not merely formal contrasts. They shape the listener's temporal experience by producing alternation, suspension, re-

turn, transformation, and balance. In this sense, *yin–yang* offers an aesthetic vocabulary for describing the dynamic organization of musical experience (Hall and Ames 1998; Li 2010).

A neuroscientific interpretation of these relationships must remain metaphorical and exploratory rather than directly evidential. It would be too strong to argue that *yin–yang* corresponds to a specific neural mechanism. Nevertheless, the concept can help describe how aesthetic experience may be organized dynamically: attention moves between external sound and internal imagery; affect fluctuates between tension and calm; perception alternates with memory and reflection; sound gains meaning through its relation to silence. In this way, *yin–yang* can function as a philosophical model for understanding musical experience as a process of dynamic balance rather than as a linear progression from perception to pleasure.

This framework also explains why Chinese musical aesthetics often values suggestion, restraint, resonance, and incompletion. The force of *yijing* lies not only in what is sounded, but also in what is implied or left for the listener to complete. Silence, temporal spaciousness, subtle timbral change, and restrained affect actively create conditions for imaginative participation (Jullien 2012; Li 2010).

Accordingly, *yijing* should not be treated as a replacement for existing neuroaesthetic models, nor as evidence for an essentially "Chinese" neural paradigm. Its value lies in offering a historically situated and culturally grounded theoretical lens through which the relations among sound, image, affect, memory, embodiment, and self-reflection can be examined more carefully. By bringing *yijing* into dialogue with neuroaesthetic research, this article aims to refine rather than reject existing models. It asks how a concept developed within Chinese aesthetics may generate more precise hypotheses about musical imagery, affective resonance, temporal unfolding, and contemplative integration.

5. Xiyang Xiaogu as a Conceptual Case of Sound–Scene Transformation

The case analysis is anchored in the traditional pipa lineage of Xiyang Xiaogu rather than in a modern orchestral or piano arrangement. The reference material is Wu Man's 1999 Freer Gallery performance, made publicly available by the Smithsonian's National Museum of Asian Art. The source identifies the piece as Xi Yang Siao Gu (Flute and drum music at sunset), lists it at 0:00–8:22, and explains that a seven-section untitled manuscript appeared in 1875 while an eight-section programmatic version was included in Li Fangyuan's 1895 New Collections (Smithsonian's National Museum of Asian Art n.d.). The eight titles move from distant bells and drums and moonlit mountains through breeze, water, flowers, clouds, fishing, waves, and a returning boat.

The following analysis therefore concerns a particular recorded performance and its documented programmatic frame, not an abstract composition detached from performance. This distinction matters because silence, tempo, timbre, resonance, phrasing, and expressive timing depend partly on performance practice. At the same time, the analysis does not treat Wu Man's realization as the only authoritative version. The recording and notes provide a transparent reference point through which readers can evaluate the interpretive claims.

The procedure is conceptually guided rather than computational. The 0:00–8:22 recording is divided into three broad listening regions corresponding to the programmatic trajectory: an opening region of distance and moonrise; a middle region of water, clouds, shadows, and fishing; and a later region of waves and return. For each region, the analysis distinguishes sound-based affordances from programmatic framing. Sound-based affordances include attack, decay, register, timbre, rhythmic density, pauses, resonance, and pacing. Programmatic framing includes the title, section subtitles, and culturally available landscape associations.

In the opening region, the framing of distant bells and drums and moonrise prepares the listener to hear the first musical events spatially. Separated attacks, decaying resonance, register contrast, and pauses can create perceptual distance even before any determinate image is formed. These features do not objectively depict a temple or a moon. Rather, attack, decay, and silence provide auditory conditions from which an inner scene may be constructed when combined with programmatic framing. This region most directly illustrates perceptual–imaginal formation: sound becomes available as spatial and atmospheric imagery.

In the middle region, associated with quiet water, flower shadows, distant clouds, and evening fishing, the analytical emphasis shifts from initial scene construction to continuity with change. Recurrence, subtle alteration in pacing, changes in density, shifting register, and timbral nuance may allow listeners to hear the present event against a recently established memory. Water imagery is useful here not as proof of musical representation but as a culturally available analogy for fluctuation, suspension, and return. As the imagined scene acquires calmness, distance, nostalgia, or intimacy, perceptual imagery becomes affectively colored. This is the process named affective resonance.

In the later region, the programmatic movement toward waves and the returning boat introduces activity and return without requiring a fixed dramatic plot. A performance may realize this trajectory through increased rhythmic motion, denser texture, brighter articulation, registral expansion, or a stronger sense of arrival, followed by easing, thinning, or lingering resonance. The analytical point is relational: increased motion makes later relaxation perceptually meaningful, while the idea of return can activate memory of earlier material. Closure is therefore not merely formal. It becomes an opportunity for recollection and inward completion, which is the principal function of reflective integration in this case.

This distinction between sound-based and framing-based influence is central to the model. Without programmatic cues, listeners might still report contour, pacing, timbre, movement, or calmness; with titles and section names, they may be more likely to describe spatial scenes, water, distance, or return. The model does not claim that one level alone produces *yijing*. It claims that *yijing*-oriented listening emerges from interaction among acoustic affordances, programmatic language, memory, cultural knowledge, and reflective attention.

Table 1 is therefore an analytical map rather than a claim of one-to-one correspondence. The case does not prove a distinct neural state, but it clarifies how a culturally framed musical work can make sound–scene transformation, affective resonance, and reflective completion observable as analytic processes. It also identifies empirical questions: whether programmatic framing increases reports of visual–spatial imagery, whether imagery mediates affective response, and whether reflective integration is shaped by familiarity, training, language, and personal memory.

Table 1. Sound-based and framing-based affordances in the conceptual analysis of Xiyang Xiaogu.

Framing Affordances and Neuro-*Yijing* Emphasis	Sound-Based Affordances	Analytical Region in the 0:00–8:22 Reference Recording
Distant bells/drums and moonrise; supports perceptual–imaginal formation	Separated attacks, decay, register contrast, pauses, and resonance	Opening: distance and moonrise
Quiet water, clouds, and fishing; supports affective resonance through continuity and emotional coloring	Recurrence with subtle changes in pacing, density, register, articulation, and timbre	Middle: water, clouds, shadows, fishing
Waves and return; supports reflective integration through memory, return, and inward completion	Increased motion, possible textural thickening, easing, thinning, and lingering resonance	Later: waves and returning boat

6. Toward a Preliminary Neuro-*Yijing* Model

Building on the theoretical discussion and the case analysis, this section proposes a preliminary Neuro-*Yijing* Model as a hypothesis-generating framework for cross-cultural neuroaesthetic research. The term "Neuro" is used in a restricted sense. It does not assert that *yijing* is a fixed brain state or that a culturally exclusive neural mechanism has been identified. It means that *yijing* is used to organize research questions about auditory perception, visual–spatial imagery, embodied simulation, affective resonance, memory, and self-referential cognition.

The model is based on three interrelated dimensions: perceptual–imaginal formation, affective resonance, and reflective integration. These dimensions were selected because they correspond to the main transformation described by *yijing*-oriented listening: sound becomes available as imagined scene; scene becomes affectively charged; affective experience becomes integrated with memory, self-reference, and contemplative awareness. The dimensions are recursive rather than linear. A musical sound may evoke an image, that image may reshape attention to the sound, and reflective awareness may retrospectively alter the meaning of earlier musical events. Figure 1 summarizes the model's recursive structure and the principal moderating factors.

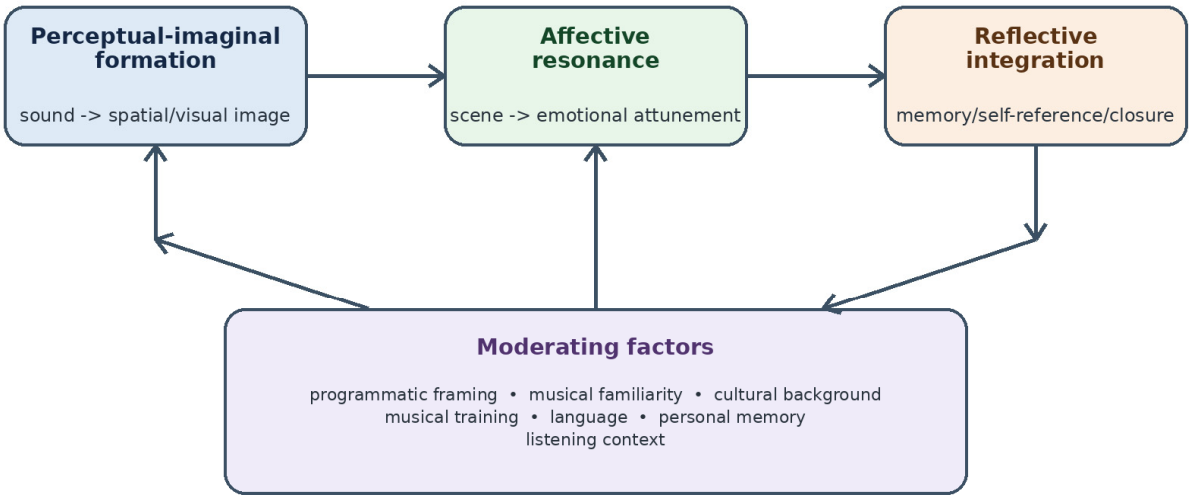

Figure 1. Preliminary Neuro-*Yijing* Model as a recursive framework. Programmatic framing, cultural background, musical familiarity, musical training, language, listening context, and personal memory may moderate the strength and sequence of the three dimensions.

The model has clear boundaries. It is most applicable to musical experiences in which sound, poetic or programmatic framing, imagined scene, restrained affect, temporal spaciousness, and reflective listening interact. It is less appropriate for experiences dominated primarily by technical virtuosity, functional dance regularity, background entertainment, or immediate bodily arousal unless these experiences also generate scene-like imagery and reflective integration. Abstract music without title or narrative can still produce *yijing*-like experience, but the claim would then depend more heavily on sound-based affordances and listener reports rather than on programmatic framing.

The first dimension, perceptual–imaginal formation, concerns the transformation of auditory perception into inner scene construction. Sound may function not only as acoustic information but also as a cue for spatial imagination, atmosphere, memory, and embodied simulation. This claim is supported by work on music-evoked visual mental imagery,

which shows that listeners often report images, places, colors, motion, objects, and affective scenes during music listening (Taruffi and Küssner 2019; Dahl et al. 2023). It is also consistent with broader research on visual mental imagery and cross-modal association (Pearson 2019; Wanke et al. 2025).

In Xiyang Xiaogu, perceptual–imaginal formation is foregrounded through the interaction of title, subtitles, timbre, rhythmic pacing, pauses, and resonance, which encourage listeners to hear sound as an unfolding scene rather than a purely formal sequence. Embodied simulation may also contribute, since passive music listening can recruit motor and premotor systems even without overt movement (Gordon et al. 2018).

The second dimension, affective resonance, concerns the emotional coloring of the imagined scene. In Chinese aesthetic discourse, *yijing* is often associated with the fusion of *qing* (情, emotion) and *jing* (景, scene or realm). This fusion should not be understood as the direct expression of a fixed emotion. Rather, emotion emerges through the listener's engagement with an imagined atmosphere. The scene becomes emotionally meaningful, and this emotional meaning in turn reshapes the perception of the music. Research on music-evoked emotion provides a useful framework for this dimension. Juslin and Västfjäll argue that musical emotion may arise through several mechanisms, including rhythmic entrainment, emotional contagion, visual imagery, episodic memory, and musical expectancy (Juslin and Västfjäll 2008). Koelsch similarly emphasizes that music-evoked emotions involve interactions among auditory processing, limbic structures, memory systems, and higher-order cognitive mechanisms (Koelsch 2014).

In the Neuro-*Yijing* Model, affective resonance describes how sound-induced imagery becomes emotionally charged. In Xiyang Xiaogu, the affective quality is often restrained, spacious, nostalgic, and contemplative rather than overtly dramatic. Feeling is generated through atmosphere, pacing, timbre, silence, and association. This clarifies why *yijing* should not be equated with pleasure alone. Reward-related systems matter in musical enjoyment (Blood and Zatorre 2001; Salimpoor et al. 2011; Zatorre and Salimpoor 2013), but *yijing* involves a broader affective field: quietness, distance, longing, restraint, and reflective melancholy.

The third dimension, reflective integration, concerns the process through which sound, imagery, emotion, memory, and self-awareness become integrated into a more unified aesthetic experience. Classical discussions of *yijing* often emphasize that the highest aesthetic state is not merely the recognition of a beautiful object, but a form of inward resonance in which the boundary between self and scene becomes less rigid. This should not be interpreted as a literal disappearance of the self. It is better understood as a mode of contemplative attention in which the listener's memory, imagination, affective state, and cultural understanding are drawn into relation with the musical experience.

The default mode network offers a cautious reference point for this reflective dimension. It is associated with internally oriented cognition, autobiographical memory, mental simulation, and self-generated thought (Andrews-Hanna et al. 2014). Vessel, Starr, and Rubin also found that intensely moving aesthetic experiences may involve default-mode activity when artworks are personally meaningful (Vessel et al. 2012). These findings do not establish a neural basis for *yijing*, but they help frame how self-referential memory, internal imagery, and reflection may participate in contemplative musical experience.

Reflective integration is therefore the dimension in which *yijing* becomes most distinctive as an aesthetic concept. Unlike approaches that focus primarily on perceptual fluency, formal beauty, or reward response, the Neuro-*Yijing* Model emphasizes the integration of musical sound with internally generated meaning. The listener's experience is not exhausted by recognizing structure or feeling pleasure. Rather, the listener enters a reflective process in which auditory events become connected with imagined space, emotional mem-

ory, and broader cultural patterns of meaning. This dimension also explains why temporal openness is so important in *yijing*-oriented listening. Silence, lingering resonance, delayed resolution, and gradual transformation allow the listener time to form associations and integrate them internally. The aesthetic value of the music lies not only in what is sounded, but also in what is suggested and imaginatively completed.

The Neuro-*Yijing* Model therefore proposes that the aesthetic experience of Chinese programmatic music can be analyzed through the interaction of three levels. At the musical level, features such as melody, rhythm, timbre, register, texture, silence, resonance, and formal unfolding provide the perceptual basis for aesthetic experience. At the cognitive–affective level, auditory imagery, visual–spatial association, memory, expectation, embodied simulation, and emotional resonance contribute to the formation of inner scenes. At the cultural–aesthetic level, concepts such as *yijing*, *qingjing* jiaorong (情景交融), suggestion, restraint, emptiness, and dynamic balance provide interpretive frameworks through which listeners understand and value these experiences.

Within this framework, culturally situated aesthetic concepts shape how widely shared perceptual, affective, mnemonic, embodied, and self-reflective processes are organized and interpreted. *Yijing* contributes at this conceptual level by linking sound–scene transformation, qing–jing relation, restraint, emptiness, and reflective integration within a historically developed aesthetic vocabulary.

To clarify the model's added value, Table 2 positions the proposed Neuro-*Yijing* Model in relation to frameworks that already address perception, imagery, emotion, memory, reward, and self-reference. Its distinctive contribution is organizational: it treats sound–scene transformation, culturally mediated affect, and contemplative completion as a recursive aesthetic process guided by a historically developed cultural concept.

On this basis, the model yields testable propositions and possible disconfirming evidence. First, programmatic or poetic framing should increase reports of scene, atmosphere, spatiality, and temporal openness compared with neutral presentation; if framing has little effect, the model's emphasis on cultural–aesthetic vocabulary would be weakened. Second, listener-reported imagery should mediate the relation between musical affordances such as timbre, resonance, pacing, silence, and reported affective resonance; if imagery and affect remain unrelated, the qing–jing component of the model would require revision. Third, reflective integration should be associated with memory, personal meaning, and inward attention; if *yijing*-oriented reports are predicted only by liking or reward ratings, the model's distinct contribution would be reduced. These are research questions, not established findings.

The model also has methodological implications. If aesthetic concepts influence how listeners attend to, describe, and value musical experience, then neuroaesthetic research should not rely only on universal rating categories such as beauty, pleasure, liking, or preference. It should also examine culturally specific aesthetic vocabularies that direct attention toward other dimensions of experience. In the case of Chinese programmatic music, *yijing* highlights the role of sound–image transformation, atmosphere, temporal spaciousness, emotional restraint, and contemplative resonance. These dimensions can become testable variables in future empirical research.

Future work could test these propositions through behavioral, mixed-method, and neuroimaging designs. Behavioral studies could compare listener descriptions under neutral versus programmatic framing; mixed-method work could combine imagery reports, affect ratings, cultural-background questionnaires, and listening-history measures; and neuroimaging research could examine whether imagery-oriented and reflective listening is associated with interaction among auditory, visual–associative, motor, memory-related,

affective, and self-referential systems. Such studies would need to control carefully for familiarity, training, language, listening context, and individual differences.

Table 2. Comparative positioning and added value of the Neuro-*Yijing* Model.

Specific Added Value of the Neuro-*Yijing* Model	What It Can Already Explain	Existing Framework
Specifies how these systems may be culturally organized around sound–scene transformation and contemplative completion.	Interaction of sensory–motor, emotion–valuation, and meaning–knowledge systems.	Aesthetic triad model (Chatterjee and Vartanian 2014)
Explains how imagery and affect become recursively linked through qing–jing relation and *yijing*-oriented listening.	Imagery, memory, expectancy, contagion, entrainment, and emotional response.	Music-emotion mechanisms (Juslin and Västfjäll 2008; Koelsch 2014)
Adds atmospheric openness, restraint, silence, and meaning that may not be reducible to hedonic reward.	Expectation, prediction, pleasure, and reward in musical experience.	Reward and predictive accounts of music (Salimpoor et al. 2011; Vuust et al. 2022; Zatorre and Salimpoor 2013)
Links inwardness to programmatic scene formation, cultural vocabulary, memory, and contemplative closure.	Self-referential and internally oriented responses to moving artworks.	Default-mode and personal-meaning accounts (Vessel et al. 2012)
Accounts for why scene-like imagery is aesthetically valued and integrated with affect and reflection.	Occurrence, content, and affective functions of visual imagery during music listening.	Music-evoked visual imagery research (Taruffi and Küssner 2019; Dahl et al. 2023)

7. Discussion and Implications

The Neuro-*Yijing* Model proposed in this article offers a preliminary framework for reconsidering how culturally specific aesthetic concepts can refine neuroaesthetic inquiry. Its significance lies in clarifying how *yijing* organizes perception, imagery, affect, memory, embodiment, and reflection into a culturally meaningful listening process.

This contribution is consistent with broader developments in music neuroaesthetics. Existing research has shown that musical aesthetic experience involves distributed interactions among perceptual, affective, cognitive, motor, and reward-related systems (Brattico and Pearce 2013; Chatterjee and Vartanian 2014; Vuust et al. 2022). Music does not become aesthetically meaningful through acoustic perception alone. It also engages memory, expectation, bodily response, emotional appraisal, attention, and cultural interpretation (Juslin and Västfjäll 2008; Koelsch 2014). The Neuro-*Yijing* Model builds on these findings by foregrounding a particular configuration of these processes: sound becomes associated with inner scene formation, imagined scenes acquire emotional resonance, and affective experience is integrated through memory and reflective awareness.

The first implication is theoretical. The model shifts the organizing question from how strongly music is liked or valued to how sound becomes scene, how scene acquires affective resonance, and how that resonance is reflectively integrated.

The second implication is methodological. If aesthetic concepts shape how listeners attend to and describe music, cross-cultural neuroaesthetics must examine the terms through which experience is elicited and measured. Asking only whether a passage is "beautiful" or "pleasant" may miss atmosphere, spatial imagery, emotional restraint, or contemplative resonance. Future research could combine rating scales with qualitative reports, imagery descriptions, cultural-background questionnaires, and listening-history measures, especially because cross-cultural work suggests both shared affective dimensions and culturally variable descriptors for musical experience (Cowen et al. 2020).

The third implication concerns cross-cultural comparison. Cultural background should not be treated as a fixed essence that divides listeners into simple "Eastern" and "Western" groups. A more rigorous approach would examine specific variables such as musical familiarity, language, training, aesthetic vocabulary, exposure to programmatic traditions, and interpretive habits. Research in cultural psychology has shown that perception, attention, self-construal, and meaning-making may be shaped by cultural experience and learned patterns of interpretation (Nisbett et al. 2001). Applied to music, this suggests that listeners may differ not only in preference, but also in the kinds of imagery, memories, bodily associations, and reflective meanings they bring to the listening process.

This perspective also requires care in the use of neuroscientific terminology. Default-mode activity, cross-modal integration, embodied simulation, and reward circuitry are used here as sources of testable questions, not as direct equivalents of traditional aesthetic categories. The methodological task is to examine how perception, imagery, memory, emotion, embodiment, and self-related cognition interact during aesthetic experience without reducing *yijing* or qingjing jiaorong (情景交融) to a single neural correlate (Andrews-Hanna et al. 2014; Vessel et al. 2012).

The model further suggests empirical directions. Behavioral studies could compare how listeners from different backgrounds describe Chinese programmatic works. Experiments could test whether titles, poetic descriptions, or contextual framing alter imagery reports, emotional ratings, memory associations, and perceived meaningfulness. Neuroimaging studies could examine whether imagery-oriented and contemplative listening involves interaction among auditory, visual–associative, memory-related, affective, and self-referential networks, while controlling for familiarity, training, language, context, and individual differences.

Finally, the model has implications for pedagogy and performance practice. If *yijing* involves the integration of sound, imagery, emotion, and reflection, teaching and performance can attend not only to technical accuracy but also to phrase direction, breath, resonance, silence, timbral nuance, and temporal spacing. This practical point should be understood as an interpretive implication rather than a fixed rule: the aim is not to impose a single "Chinese" response, but to show how performers may create conditions for listeners to participate imaginatively in the music.

At the same time, the limitations of the present study must be clearly acknowledged. The Neuro-*Yijing* Model remains conceptual and hypothesis-generating. It is based on theoretical synthesis and a single interpretive case rather than direct experimental evidence. The case study of Xiyang Xiaogu clarifies the model's analytical potential, but it cannot by itself establish general conclusions about Chinese music as a whole or determine whether the proposed dimensions correspond to specific neural patterns.

In sum, the model's contribution is to make culturally mediated sound–scene transformation available as a neuroaesthetic research problem. Its claims are modest but useful: aesthetic experience is biologically grounded, culturally organized, and methodologically sensitive to the vocabulary through which it is elicited and described. By examining how sound, imagery, emotion, memory, embodiment, and reflection are organized through *yijing*, this article offers a pathway toward a more pluralistic theory of musical aesthetic cognition.

8. Conclusions

This article has examined how the Chinese aesthetic concept of *yijing* may contribute to a more culturally responsive understanding of musical aesthetic cognition. It identifies *yijing* as a historically developed way of organizing sound, imagined scene, affective resonance, memory, and reflective awareness, thereby foregrounding dimensions of mu-

sical experience that may be obscured when inquiry is limited to beauty, pleasure, liking, or reward.

The revised Neuro-*Yijing* Model clarifies three recursive dimensions: perceptual–imaginal formation, affective resonance, and reflective integration. The case of Xiyang Xiaogu illustrates how sound-based affordances and programmatic framing may interact in a culturally situated listening process. At the same time, the model has boundaries. It is most appropriate for musical experiences involving scene-like imagery, atmosphere, temporal openness, restraint, and inward completion; it is less suitable for experiences organized primarily around virtuosity, dance function, or immediate arousal unless these also generate reflective sound–scene integration.

The model's value will depend on future testing. Research should examine whether programmatic framing increases scene imagery, whether imagery mediates affective resonance, and whether reflective integration is associated with memory, personal meaning, and inward attention beyond preference or pleasure ratings. In this sense, the Neuro-*Yijing* Model should be understood as a starting point for further behavioral, music–psychological, cross-cultural, and neuroscientific inquiry rather than as a finished explanation.

Funding: This research received no external funding.

Data Availability Statement: No new data were created or analyzed in this study. The publicly available recording used for the conceptual case analysis is cited in the References.

Conflicts of Interest: The author declares no conflicts of interest.

References

Andrews-Hanna, Jessica R., Jonathan Smallwood, and R. Nathan Spreng. 2014. The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences* 1316: 29–52. [CrossRef] [PubMed]

Blood, Anne J., and Robert J. Zatorre. 2001. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proceedings of the National Academy of Sciences of the United States of America* 98: 11818–23. [CrossRef] [PubMed]

Brattico, Elvira, and Marcus Pearce. 2013. The neuroaesthetics of music. *Psychology of Aesthetics, Creativity, and the Arts* 7: 48–61. [CrossRef]

Brown, Steven, Xiaoqing Gao, Loren Tisdelle, Simon B. Eickhoff, and Mario Liotti. 2011. Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *NeuroImage* 58: 250–58. [CrossRef] [PubMed]

Chatterjee, Anjan. 2011. Neuroaesthetics: A coming of age story. *Journal of Cognitive Neuroscience* 23: 53–62. [CrossRef] [PubMed]

Chatterjee, Anjan, and Oshin Vartanian. 2014. Neuroaesthetics. *Trends in Cognitive Sciences* 18: 370–75. [CrossRef] [PubMed]

Cowen, Alan S., Xia Fang, Disa Sauter, and Dacher Keltner. 2020. What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences of the United States of America* 117: 1924–34. [CrossRef] [PubMed]

Dahl, Sofia, Antonio Stella, and Thomas Bjørner. 2023. Tell me what you see: An exploratory investigation of visual mental imagery evoked by music. *Musicae Scientiae* 27: 717–40. [CrossRef]

Gordon, Chelsea L., Patrice R. Cobb, and Ramesh Balasubramaniam. 2018. Recruitment of the motor system during music listening: An ALE meta-analysis of fMRI data. *PLoS ONE* 13: e0207213. [CrossRef] [PubMed]

Hall, David L., and Roger T. Ames. 1998. *Thinking from the Han: Self, Truth, and Transcendence in Chinese and Western Culture*. Albany: State University of New York Press.

Jing, Hongmei. 2023. The ideological origins and aesthetic construction of yijing (artistic conception). *International Communication of Chinese Culture* 10: 151–69. [CrossRef]

Jullien, François. 2012. *The Great Image Has No Form, or on the Nonobject Through Painting*. Translated by Jane Marie Todd. Chicago: University of Chicago Press.

Juslin, Patrik N., and Daniel Västfjäll. 2008. Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences* 31: 559–75. [CrossRef] [PubMed]

Kawabata, Hideaki, and Semir Zeki. 2004. Neural correlates of beauty. *Journal of Neurophysiology* 91: 1699–705. [CrossRef] [PubMed]

Koelsch, Stefan. 2014. Brain correlates of music-evoked emotions. *Nature Reviews Neuroscience* 15: 170–80. [CrossRef] [PubMed]

Leder, Helmut, Benno Belke, Andries Oeberst, and Dorothee Augustin. 2004. A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology* 95: 489–508. [CrossRef] [PubMed]

Li, Zehou. 2010. *The Chinese Aesthetic Tradition*. Honolulu: University of Hawaii Press.

Nisbett, Richard E., Kaiping Peng, Incheol Choi, and Ara Norenzayan. 2001. Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review* 108: 291–310. [CrossRef] [PubMed]

Pearson, Joel. 2019. The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience* 20: 624–34. [CrossRef] [PubMed]

Salimpoor, Valorie N., Mitchel Benovoy, Kevin Larcher, Alain Dagher, and Robert J. Zatorre. 2011. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience* 14: 257–62. [CrossRef] [PubMed]

Smithsonian's National Museum of Asian Art. n.d. Master of the Chinese Pipa: Wu Man. Available online: https://asia-archive.si.edu/podcast/master-of-the-chinese-pipa-wu-man-2/ (accessed on 27 June 2026).

Taruffi, Liila, and Mats B. Küssner. 2019. A review of music-evoked visual mental imagery: Conceptual issues, relation to emotion, and functional outcome. *Psychomusicology: Music, Mind, and Brain* 29: 62–74. [CrossRef]

Vessel, Edward A., G. Gabrielle Starr, and Nava Rubin. 2012. The brain on art: Intense aesthetic experience activates the default mode network. *Frontiers in Human Neuroscience* 6: 66. [CrossRef] [PubMed]

Vuoskoski, Jonna K., and Tuomas Eerola. 2015. Extramusical information contributes to emotions induced by music. *Psychology of Music* 43: 262–74. [CrossRef]

Vuust, Peter, Ole A. Heggli, Karl J. Friston, and Morten L. Kringelbach. 2022. Music in the brain. *Nature Reviews Neuroscience* 23: 287–305. [CrossRef] [PubMed]

Wanke, Riccardo, Alessandro Ansani, Nicola Di Stefano, and Charles Spence. 2025. Exploring auditory morphodynamics: Audiovisual associations in sound-based music. *i-Perception* 16: 1–20. [CrossRef] [PubMed]

Yu, Pauline. 1987. *The Reading of Imagery in the Chinese Poetic Tradition*. Princeton: Princeton University Press.

Zatorre, Robert J., and Valorie N. Salimpoor. 2013. From perception to pleasure: Music and its neural substrates. *Proceedings of the National Academy of Sciences of the United States of America* 110: 10430–37. [CrossRef] [PubMed]

Zeki, Sémir. 1999. *Inner Vision: An Exploration of Art and the Brain*. Oxford: Oxford University Press.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.