

Affective Bond Between AI-Generated Music and Human Listeners

Nanfei Wang

*Department of Arts Management, Xinghai Conservatory of Music, Guangzhou, China
maggie_wang06@outlook.com*

Abstract. Generative artificial intelligence is being applied to create music in an ever-widening range, but it has not been determined whether the results can effectively convey human emotions. A controlled-variable experimental design is used to explore the causal relationship between information density in prompts and listeners' affective recognition of AI-generated music. Six AI-generated musical works were selected as the experimental stimuli under three increasing levels of prompt conditions: minimal emotional descriptions, prompts that added harmonic progression information, and detailed prompts including parameters such as instrumentation, vocal processing and dynamic changes. The results showed that listeners could not understand the creators' purposes in either the minimal-prompt or harmony-augmented situations. All 30 listeners were able to feel the emotion and structure that were desired, and they could do so after repeated instructions. The study also found that the audio outputs of AI music systems generally have logical discontinuities and perceptual incongruity without fine-grained human encoding intervention. This paper proposes that AI functions as a medium for music creation, and its affective communication performance is based on a closed-loop human intervention system at two stages: prompt engineering (information encoding) and generative evaluation (signal decoding). This study offers empirical support and practical paths for human-AI collaboration models in AI-assisted music creation.

Keywords: AI-generated music, affective bond, prompt engineering, human-AI collaboration, controlled experiment

1. Introduction

Generative artificial intelligence technology is developing at a high speed in recent years and is now changing how music is created and distributed. Text-to-music systems such as Suno AI have changed the traditional barrier of professional musicianship in music creation, and now, only descriptive text prompts are required to generate full-fledged pieces of music in seconds. These technologies have reduced the threshold for entry into music creation. However, when the authority for creation is transferred from professional musicians to ordinary users, and when it moves from the human mind to algorithmic systems, a fundamental problem arises: if the "subject" of music creation is no longer humans but algorithms, can the affective power of music still be effectively transmitted?

Music has always served different emotional purposes for people. Music can move people and give a sense of belonging to a group; it is not merely an organised system of sound. Research on physiological synchrony has shown that, in interaction, people often display coordinated rhythms in their bodies, such as heart rate and breathing; thus, this synchronisation can foster a sense of closeness and affection among individuals [1-3]. Therefore, the emotional effect of music may be due to the connections among people that it helps to create; that is, the music itself may not be the source of emotion. If the above assumption holds true, a serious problem will arise when AI is employed in music creation: What role should AI assume—should it replace human creators as the "subject" of emotion and expression, or become a new form of "intermediary" for connecting with people emotionally?

Some research has already begun to address this problem to some extent. At the theoretical level, Juslin's BRECVEMA framework has systematically presented eight psychological mechanisms by which music elicits emotions and provided an all-encompassing theoretical system for exploring how music affects people's emotions [2, 3]. At the empirical level, researchers have been studying "AI-composer bias" extensively to determine whether listeners systematically give negative evaluations after learning that a piece of music was created by AI [4-6]. However, most of these studies have focused on the acceptance of AI as the "author" and paid less attention to how humans, as "collaborative creators", can compensate for the lack of AI subjectivity through optimisation at the input end. In other words, the existing studies have mainly explored how people evaluate music created by AI, but have rarely investigated how to create genuinely moving music with AI.

This study aims to address the above research deficiency. Through controlled-variable experiments, this study systematically examines how the precision of prompts affects the emotional expression in AI-generated music and attempts to answer the following three questions: (1) Does the density of information in the prompt change listeners' perception of the intended emotional meaning embedded in AI-generated music? (2) Under what prompt conditions can AI-generated music effectively convey emotions? (3) Based on the empirical results, how should the role of AI in music creation be redefined?

The structure of this paper is as follows: Section 2 presents the theoretical basis of musical-affective bonding and previous research on affective perception in AI-generated music; Section 3 reports the research results through a three-level comparative analysis of six case groups; Section 4 concludes the entire paper and proposes future research directions.

2. Literature review

2.1. Theoretical foundations of affective bond

2.1.1. The BRECVEMA framework: eight mechanisms of music-induced emotion

Why is music emotionally moving? It is also one of the problems in music psychology at present. Swedish Psychologist Patrik N. Juslin at Uppsala University and his colleagues have made the first contributions to this area. In 2008, Juslin and Västfjäll put forward the BRECVEM framework in Behavioral and Brain Sciences to systematically integrate the six psychological mechanisms of music-evoked emotions for the first time [2]. In 2013, Juslin modified and extended this model in "Physics of Life Reviews" by adding an "aesthetic judgment" mechanism to create the full BRECVEMA model [3].

The eight modules of the BRECVEMA system are:

(1) Brain Stem Reflex. This is the most primitive mechanism, referring to the automatic affective responses of listeners to sudden, intense, or extreme sound stimuli, mediated directly by the brain stem.

(2) Rhythmic Entrainment. This refers to the gradual synchronization process between the listener's internal physiological rhythms—such as heart rate, breathing rate, and motor rhythms—and the external musical rhythms. This mechanism explains why music with a distinct rhythm can trigger bodily movement and affective arousal.

(3) Evaluative Conditioning. This refers to music acquiring the affective valence of a specific stimulus through repeated pairing with that stimulus.

(4) Emotional Contagion. This describes the process by which listeners "catch" emotional expression from a musical performance.

(5) Visual Imagery. This describes the visual images evoked in the listener's mind by music, along with the emotional associations that accompany them.

(6) Episodic Memory. This occurs when musical features trigger the listener's unique personal memories, and the emotions connected to those memories that are thereby evoked—this is the mechanism with the greatest individual variability among the eight.

(7) Musical Expectancy. This refers to the affective responses generated by the establishment, violation, or confirmation of expectations during the progression of music. When music proceeds in the direction the listener anticipates, it produces satisfaction; when expectations are skillfully delayed and then ultimately fulfilled, it produces intense emotional release. This mechanism has the most direct relevance to the case analysis in the following sections—it explains how structured sectional design and dynamic contours amplify listeners' affective engagement.

(8) Aesthetic Judgment. This denotes conscious aesthetic appraisal based on cognitive evaluation, relying on the involvement of higher brain regions such as the prefrontal cortex.

In line with evolutionary psychology, the eight mechanisms of the BRECHEMA framework are believed to be information-processing systems at different levels of the brain. The system no longer distinguishes between the ideas of innate and learned musical emotions; instead, it believes that the appearance of musical emotion is a multi-pathway complex phenomenon.

2.1.2. The dual-pathway neural mechanism of music aesthetics

Neuroscience has provided the theoretical basis for the BRECHEMA framework. The generation of music-induced aesthetic pleasure is supported by two parallel neural pathways [7].

The first is a "bottom-up" way; therefore, acoustic stimuli directly excite the emotional areas of the brainstem and limbic system. This way is quick and highly automated; it does not require higher-order thinking. The two kinds of brainstem reflexes and rhythmic entrainment are shown above.

The second is the "top-down" path: higher brain areas, such as the prefrontal cortex, integrate, anticipate and evaluate musical information. This way is slower but more accurate, and it needs cognitive functions such as working memory, attention allocation and predictive processing. The way in which music evokes anticipation and beauty is generally through this path.

The two paths need to work together to provide a full musical-aesthetic experience. A fast bottom-up response has a sense of immediacy and intensity in emotion; a well-organised top-down process is deep and prolonged. The two-pathway model has important implications for understanding affective communication in AI-generated music; if the AI-generated music lacks sufficient regularity and predictability at the structural level (i.e., it cannot support top-down processing), a full and lasting aesthetic experience will not be achieved even after an initial affective response [8-10].

2.1.3. Physiological synchrony: the social functional basis of music's affective connection

Physiological synchrony is the coordination of the internal rhythms of heart rate and breathing among interacting people, and it has been linked to the formation of social relationships and emotions.

A large-scale study in 2024 published in *Scientific Reports* found that the heart rate, breathing rate and skin conductance responses of people listening to classical music changed simultaneously. The study found that the extent of this "induced synchrony" was positively correlated with listeners' self-reported aesthetic immersion experience, and that listeners' personality traits (such as openness and extraversion) could predict the degree of synchronisation. Another study published in the same journal also found that music-induced heart rate synchronisation is due to the reliability of physiological responses to the musical work, not to the listener's mood or motivation at that time.

Ng, Sargent and Snell proposed the Cyborg Synchrony framework in *Frontiers in Computer Science* to directly extend research on physiological synchrony to AI-music research [8]. The two-stage AI training framework works as follows: First, a foundation model builds a map between musical features (tempo, loudness, timbre, etc.) and people's physiological responses based on a database of physiological data from thousands of individuals; second, the system dynamically creates personalised music according to the user's real-time biometric data. The goal of this system is not to replace people's emotions with AI-generated music, but to promote social harmony among the people through the cultivation of cross-individual physiological synchronisation.

The above studies have shown that, to some extent, the emotional effect of music is not due to the music itself but rather to the connections among people that it helps to create. Research on physiological synchrony provides strong physiological support for the above claim: when people experience changes in emotion simultaneously through music, their bodies also respond in the same way. This "bodily resonance" is precisely the most fundamental basis for socialisation.

2.2. Existing research on AI-generated music and affective perception

2.2.1. The debate on "AI-composer bias"

Can AI-generated Music Be Enjoyed by Listeners? Does awareness of an AI origin for music lead to systematic bias in auditory perception? It is one of the more contentious problems in the area of AI music research.

Hong, Fischer and Ha published a study in *Computers in Human Behaviour* that mentioned this problem at an early stage [4]. Experimental work in the study examined how machine-related attributes affect evaluations of AI music, and it was found that participants who regarded AI music generators as legitimate "musicians" were more likely to like the AI-generated music. Therefore, the bias of "AI-composer bias" is not universal and depends on how listeners perceive the role of AI.

In 2026, Xu, Xu and Xie published a larger-scale experimental study (253 participants, 24 music excerpts) in the same journal [5]. Using linear mixed models for analysis, the study found no systematic "AI-composer bias effect". A moderator was also found: for complex or calming music, human-composed works were more likely to arouse awe in listeners; for intense or fast-paced music, AI-generated works performed better. In other words, the impact of the composer's works on listeners at different levels of musicality (affective intensity and complexity) was not the same.

There are also opposing results. Wu and Holmes indicated in *Cognitive Research: Principles and Implications* that attributing music to AI may interfere with listeners' narrative-meaning construction [6]. Their research found that the belief in an AI origin for musical material reduces the inclination

to construct a story or search for an implicit "narrator" during music reception. As the construction of narrative meaning is required for the emotional effect of music, this inhibition will directly reduce the affective power of AI music.

The above results are in contradiction with the assumption that the source of AI-generated music determines how it is emotionally received; instead, multiple factors such as the features of the music and the cognitive systems of listeners also influence this reception.

2.2.2. Ontological-philosophical dilemmas of AI-generated music

In addition to the empirical controversy, AI-generated music also raises a deep philosophical problem of ontology for affective communication in music. The reason music can move people is not that the sound signals themselves carry some objective affective attributes; rather, it is that listeners naturally assume—or rather, instinctively believe—that behind these organised sounds is a subject with the ability to feel emotions and a personal life history. The idea of "intentionality" in music aesthetics offers theoretical support for this: when we listen to music, we are not merely perceiving sound; instead, we are actively creating an imaginative and intersubjective space with an "expresser". When listeners are aware that the source of the sound is not a person with lived experiences of joy and sorrow, but an algorithmic system devoid of life history and inner experience, the cognitive foundation for such affective narration is structurally weakened.

The above judgment can also be viewed from the perspective of phenomenology. Husserl's theory of intersubjectivity holds that human affective understanding fundamentally depends on the recognition of and empathy with the subjectivity of the "other"; thus, the perception of others' sadness stems from the assumption that an inner flow of consciousness exists in other people. Heidegger further indicated that the existential structure of Dasein is Mitsein (being-with-others), and thus our perception of the world is always relative to others. As an expressive art, the power of music to move people's emotions lies in this presupposition of "being-with": when listeners feel that music is the expression of another "Dasein", they are moved emotionally. However, as a non-living thing, AI does not have the quality of "Dasein" at the level of ontology; it lacks an existential situation of "being-in-the-world" and therefore cannot become the "other" in an emotional relationship. At its core, AI music generation is a statistical fit of musical styles and structures by algorithms, not an expression of emotions from a particular subject. In other words, the affective communication effect of AI-generated music is inherently limited at the level of philosophy by a paradox: the more independently AI attempts to express emotion, the more conscious listeners become of the absence of "someone" behind it, and only when listeners actively forget that the music was created by AI can emotional resonance occur. This contradiction is the philosophical boundary that the emotional connection of AI-generated music has not yet crossed.

2.2.3. Empirical explorations of human-machine connection: music and empathetic robots

At present, many efforts are being made to explore the emotions conveyed by AI music through human-computer interaction. Johan F. Johan's studies show that Hoorn's team at The Hong Kong Polytechnic University, in cooperation with The Chinese University of Hong Kong, found that combining "music" and "empathetic language" in the design of AI robots can enhance the emotional connection between people and robots [11].

The study had the participants engage in several rounds of interaction with a robot on a screen, and the robot was randomly assigned to one of four conditions: empathetic language with music, empathetic language without music, music only, or no intervention. Based on the above results,

adding expressions of empathy and music can increase people's sense of empathy for robots. The researchers provided a reason for this phenomenon and indicated that music can create a sense of closeness to "real" conversation with personality, and human psychological counselors also use music to soothe distressed people.

However, the study also found that as the participants became more accustomed to the repeated interactions, the benefit of musical cues gradually declined over time. Therefore, the emotional design of an AI system should not be fixed; instead, it needs to be able to change dynamically according to changes in user feedback and the environment, just as humans do.

In theory, this study proposes that the role of AI in affective systems should be changed from that of an "object" of emotion to that of a "medium" for emotional experience.

2.2.4. Research gap and positioning of this study

Based on the literature reviewed above, a serious deficiency in the current research has been identified.

Based on theory, the BRECVEMA framework and a dual-pathway neural mechanism have proposed that the whole experience of musical emotion is a result of "top-down" cognition that leverages organised acoustic information. Research on physiological synchronisation has also shown that the emotional effects of music are related to a sense of closeness among people. Therefore, if AI music is to achieve genuine emotional expression, it needs to have a certain regularity and predictability in its structure, and make listeners feel as if there is a "person" behind the sound.

At the empirical level, most of the existing research falls into two categories: studies on the acceptance of "AI-composer bias" and philosophical and ethical reflections on AI-generated music. However, there is a missing link between the two threads: how to compensate for the lack of AI subjectivity through optimisation at the input end. In other words, the current research has focused more on "what AI does" and "how listeners respond", but has paid little attention to the practical problem of "how humans should operate AI".

This study is based on the above deficiency. The question of whether AI can create moving music has been partially answered through technological development, and it is no longer a problem to be solved. The main problem is how to create AI that can produce music with real emotions and meanings. Prompts are the primary interfaces for human-AI interaction, so they attract attention. Controlled-variable experiments systematically investigate how the information density and dimensional richness of prompts affect listeners' recognition of the emotional intention in AI-generated music. This research perspective changes the focus of the discussion on the affective connection of AI music from "algorithmic ability" to "how humans and algorithms cooperate", and offers empirical support for the human-machine collaboration model of AI-generated music.

3. Case Analysis

3.1. Case-selection principles

To explore the limits of AI music tools in expressing emotions and creating art in depth, six pieces of music from different styles were selected as the research objects. All tracks start from human-created versions as the baseline, and based on the same underlying prompt framework, two versions—simple prompts and complex prompts—are generated. The Design will ensure that the AI-generated results are comparable and allow for an experiment with human modification.

3.2. Analytical dimensions

This paper will systematically study each set of comparative works in the following four aspects.

- (1) Harmonic and modal framework: Compare the core harmonic progressions, tonal unity, and degree of harmonic variation between the two;
- (2) Musical form and sectional structure: Examine the structural completeness, functional division of sections, and affective progression logic of both;
- (3) Orchestration, timbre, and performance details: Analyze differences in instrumentation, timbre design, performance techniques, and dynamic changes;
- (4) Vocal processing and audio details: Evaluate the refinement of vocal delivery, mixing parameters, spatial effects, and ending design.

3.3. Overview of core differences across the six case groups

In addition to the typical features of their respective genres, the six cases show a very consistent pattern in all styles.

All the simple-prompt versions have a single section or two, use fixed template-based harmonic loops, specify only instrument types and basic performance forms in orchestration, have vocals without effects processing, and maintain a consistent rhythmic groove from start to finish.

The detailed-prompt versions, under extended instructions, all achieve a structural plan of three to five sections, introduce specific techniques and parameters such as arpeggiated chords, filter processing, volume fades, and sidechain compression, and the vocals begin to show differences between breathy and chest voice along with reverb settings; however, a sense of "assembly" still exists among the elements.

The human-created versions use the same framework and achieve integration at three levels: aesthetic consistency (all details serve the theme), parameter personalisation (volume slopes, reverb times, etc., changed from "correct values" to "appropriate values"), and inter-voice dialogue (voices shift from synchronous execution to interactive response).

The particular differences among the case groups are shown in Table 1.

Table 1. Summary of core differences in the three versions of six case studies

Work	Style	Simple Version	Detailed Version	Human Version
"Midnight Rain Walk"	Lo-fi Hip Hop	Single-section loop, Em7-Am7-Bm7-Em7 root-position chords, no dynamics	Three-part form, arpeggiated chords, rain sound 8 bars 0 dB→-6 dB, hi-hat 2 kHz filtering	Dynamic arc integration, outro piano decay + rain sound fade-out
"Neon Skyline"	Synthwave	Two-part form, Fm-Db-Ab-Eb loop, bass root octaves	Intro pad filter sweep, doubled vocals in the chorus, sawtooth lead, sidechain compression	Parameter-specific adjustments, arpeggio density tapering off for the ending

Table 1. (continued)

"Empty Frame"	Indie folk	Two-part form, I-vi-IV-V progression, fingerpicking without fine detail	Five-part form, 12th-fret harmonics, string noise retained, tempo fluctuating ± 3 BPM, soft brush drums	Micro-dynamic consistency, outro fade-out at -1 dB per bar
"Broken Metronome"	Experimental Electronic	Single section, Cm9 pedal tone + dissonant tone cluster, regular rhythm	Three-part form, irregular kick drum placements, syncopation, 12-bit bit-crushing	Glitch sound effects organized narratively, 2 seconds of silence as a pause
"Paper Crown"	Pop Ballad	Two-part form, D-A-Bm-F#m root-position chords, vocals without effects	Four-part form, intro in first inversion, viola texture, breathy/chest voice layering	Vocal dynamic arc (breathy $\rightarrow$ distorted $\rightarrow$ fade-out), precise part coordination
"Roof Jazz"	Swing Jazz	Single section, 12-bar blues, no accent syncopation	Four sections, stride pattern, 2/4 beat accents, syncopation	Swing eighth-note timing adjustments, cymbal choking + bass descending closure

3.4. The rule of three-level ascension

Across the six case groups, AI-generated music shows a three-tier increase under different prompt conditions:

Level One: Simple prompts—"stylistic shell". It only responds to superficial emotional labels and stylistic descriptions, generates rigid template-based harmonic loops, a single section, and lacks performance details or dynamic design. Listeners do not know the emotion.

Level Two: Detailed prompts—"recognizable intent". When the prompt adds multi-dimensional information, such as harmonic progression, section division, performance technique, dynamic parameter and vocal processing, the AI output will have sectional layering, diverse techniques, volume gradation and vocal differentiation. Listeners can clearly understand the creative purpose (86% of the subjects in this study did). However, the musical details still show obvious stylistic patchwork artifacts, and the parameters tend to be "generic templates" without personalised calibration.

Level Three: Human post-refinement—"emotional resonance". The three levels above the AI framework for human creators' integration are: (1) aesthetic coherence—all details serve the theme; (2) parameter personalisation—changing from "correct values" to "appropriate values"; (3) conversational voice interaction—switching from "synchronized execution" to "interactive response". Listeners move from recognising the intention to feeling.

3.5. Theoretical echoes

Based on the above results, they are consistent with the "musical-expectancy" mechanism in the BRECVEMA framework: the mechanical cyclic repetition of the basic prompt version does not form valid expectation-fulfillment musical patterns; the detailed version builds a multi-layered expectancy network through sectional layering; and the human version adds emotional warmth to the satisfaction of expectancy by making slight alterations to performance dynamics and note duration and thus evokes empathy [2, 3]. A dual-pathway model is employed here: at the bottom, a path can be applied to the listener's immediate response to basic acoustic features, but for a complete aesthetic experience, fine-grained processing of structure via a top-down path is required, which is only feasible if the music contains rich structural information [7].

3.6. Practical implications

The three-tier comparative results of the six case groups are in strong agreement: the main threshold for affective communication in AI-generated music is the information density and dimensional richness of the prompt. Simple prompt outputs are relatively random in terms of emotion; multi-dimensional prompts increase the comprehensibility of AI output to a "clear" degree; and human artistic refinement achieves a leap from "recognizable" to "resonant". The core of prompt engineering is not "writing more", but "encoding more executable musical decision dimensions" to transform abstract emotions into executable harmonic, rhythmic, timbral, dynamic and spatial parameters for AI. This change is, in itself, a high-level form of musical creation.

4. Discussion

The theoretical contribution of this study is to shift the discussion of AI-music-affective connection from "algorithmic ability" to "the way humans and algorithms cooperate", and to fill the research gap in existing literature on how humans compensate for the lack of subjectivity in AI through input-side optimisation. This offers a clear prompt design strategy for AI music creators in practice: switch from "writing more" to "encoding more executable dimensions of musical decision-making".

The deficiencies of this study are as follows: the six sets of cases focused on specific styles and did not cover a wider range of music genres; the size of the listener sample was small (N=30), and no synchronous measurement of physiological signals was performed. In the future, more samples will be collected, physiological data such as heart rate and skin conductance added, and a dynamic adaptation mechanism for AI-generated music and its application in mental health intervention will also be explored.

5. Conclusion

Based on the three-tier comparative analysis of the six case groups, this paper investigated in an organised manner how different degrees of prompting affect the effectiveness of emotion transmission in AI-generated music and derived the following main conclusions:

First, the information density of the prompts is a necessary condition for the AI-generated music to express emotions well. Simple prompts containing only affective descriptions and style labels produce outputs that are close to random for listeners; however, when these prompts are extended to include multi-dimensional information such as harmonic progressions, sectional divisions,

performance techniques, dynamic parameters and vocal processing, the AI-generated music develops creative intent and structural logic that listeners can clearly perceive.

Second, the AI-music-generation system is in fact a "transmissive instrument" rather than an independent "creative subject". AI lacks aesthetic judgment and structural musical organisation; therefore, the quality of its output requires human closed-loop intervention in "prompt engineering" (information encoding) and "generation evaluation" (signal decoding). Without the above closed-loop human intervention, AI-generated audio generally has structural logical flaws and perceptual inconsistencies. The deepest wish for AI-generated music is not to replace human artists in conveying emotions, but to create a new form of communication that can evoke emotions and foster empathy among the people through music.

Thirdly, the future direction of development for AI-generated music should move from "work creation" to "relationship building". AI should not be regarded as a replacement for human emotions in expression, but rather as a medium to help people connect emotionally; through high-level cooperation between humans and machines, emotions can be conveyed more accurately and warmly in music. In the future, AI music will likely take the form of personalised empathetic mirroring, affective heirlooms and inheritance, co-creative therapy, and collective affective resonators. At the core is to use AI as a bridge to convey human emotions more effectively through music.

References

- [1] Tschacher, W., Greenwood, S., Weining, C., Wald-Fuhrmann, M., Ramakrishnan, C., Seibert, C., & Tröndle, M. (2024). Physiological audience synchrony in classical concerts linked with listeners' experiences and attitudes. *Scientific Reports*, 14(1), 16412.
- [2] Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31(5), 559-575.
- [3] Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10(3), 235-266.
- [4] Hong, J.-W., Fischer, K., & Ha, Y. (2022). Human, I wrote a song for you: An experiment testing the influence of machines' attributes on the AI-composed music evaluation. *Computers in Human Behavior*, 131, 107239.
- [5] Xu, K., Xu, B., & Xie, Z. (2026). Beyond composer labels: Aesthetic awe in music composed by humans and AI. *Computers in Human Behavior*, 182, 109020.
- [6] Wu, S. H., & Holmes, K. J. (2026). Is there a "mind" behind the music? Attributing music to AI can suppress narrative meaning-making. *Cognitive Research: Principles and Implications*.
- [7] Song, B., Zhang, Z. Y., & Hou, J. C. (2022). An investigation to the neural mechanism of aesthetic emotional experience in music from the perspective of neuroaesthetics. *Journal of Xinghai Conservatory of Music*, (4), 125-138.
- [8] Ng, S., Sargent, K., & Snell, J. J. (2025). Cyborg synchrony: Integrating human physiology into affective generative music AI. *Frontiers in Computer Science*, 7, 1593905.
- [9] Husserl, E. (1931). *Cartesian meditations*. Martinus Nijhoff.
- [10] Heidegger, M. (1927). *Being and time*. Max Niemeyer Verlag.
- [11] Hoorn, J. F., et al. (2026). Music and empathetic speech in AI robots. *ACM Transactions on Human-Robot Interaction*.