RESEARCH ARTICLE

Optimization of music healing effect and analysis of physiological data based on human-computer interaction technology

Shangling Ge*

School of Physical Education and Art, Henan College of Industry and Information Technology, Jiaozuo, Henan, China.

Received: September 29, 2025; **accepted:** January 9, 2026.

Music therapy as a non-pharmacological intervention has shown significant value in the fields of mental health, stress management, and chronic disease rehabilitation in recent years. This study explored the potential application of human-computer interaction (HCI) technology in optimizing the effectiveness of music therapy by addressing the subjectivity and lack of dynamic feedback in traditional music therapy models. An intelligent control system integrating multi-modal physiological sensing and a generative music platform was constructed to achieve accurate perception of users' physiological and psychological states. Wearable devices were used to continuously collect physiological signals including heart rate variability (HRV), skin conductance, and respiratory rate. A support vector machine (SVM) model was applied to achieve high-precision recognition of users' "stress-relaxation" states, and physiological adjustment indexes were further calculated as the basis for intervention decisions. The system integrated the Endel artificial intelligence (AI) music platform, which dynamically adjusted the rhythm, tonality, and acoustic structure of music according to real-time physiological feedback, realizing parameter-level personalized audio generation. To verify the effectiveness of the system, this study included 80 participants for a controlled experimental study. The results showed that, after intervention, the standard deviation of normal-to-normal intervals (SDNN) of the experimental group increased from 48.7 ms at baseline to 63.2 ms, and the root mean square of successive differences (RMSSD) increased by 26.4%, which indicated that the participants' autonomic nerve regulation ability was substantially improved. Furthermore, the experimental group's average usage frequency was 5.8 times per week, which was considerably greater than the control group's 4.2 times per week (*P* < 0.05). The results showed that the dynamic feedback mechanism could enhance users' willingness to participate. This study provided feasible technical support for the intellectualization of music therapy, and revealed the neurophysiological mechanism of music intervention.

Keywords: human-machine interaction; music therapy; effect optimization; physiological data; recommendation system.

***Corresponding author:** Shangling Ge, School of Physical Education and Art, Henan College of Industry and Information Technology, Jiaozuo, Henan 454000, China. Email: 18317590066@163.com.

Introduction

In modern life, the increasingly fast-paced competitive pressure has left people mentally exhausted and anxious. Mental health issues such as anxiety and depression have gradually threatened people's normal lives [1]. Traditional methods for treating or alleviating such mental disorders rely on two approaches including pharmacotherapy and consultations with psychological interventionists. However, pharmacotherapy may lead to side effects that affect patients’ health. Furthermore, due to individual differences, the therapeutic effect of medications may not meet expectations [2-4]. Among these non-invasive, non-pharmaceutical, and side-effect-free treatment approaches, music is a promising means to ease people’s negative emotions and adjust their adverse mental states. Since ancient times, humans have used music to regulate their physical and mental conditions [5], some researchers argue that music can regulate emotions and restore mental health because it acts by modulating neurotransmitter production [6]. Specifically, soothing music reduces the level of adrenaline in the human body and promotes the secretion of endorphins, thereby helping people alleviate anxiety and lower internal stress levels. Music can stimulate the human brain, affect the brain's neural networks, enhance neural plasticity, and through such ways, exert positive effects on human cognition and rehabilitation. Human-computer interaction (HCI) refers to the use of computers and various sensors to realize information exchange between humans and machines. In the process of conducting music therapy for humans, HCI means are used to better provide personalized services. The use of HCI technology can more accurately select more appropriate music according to people's music needs and monitor people's physiological and psychological status in real time. It also contributes to better evaluation of music therapy services and achievement of ideal effects [7-9].

Music therapy has been used extensively in a variety of sectors in recent years including helping cancer patients with their psychological suffering and helping those with Alzheimer's disease with their symptoms. Dhiman *et al*. proposed an adaptive music system (AMS) based on physiological feedback, which collected heart rate variability (HRV) and skin conductance in real time through wearable devices, used a support vector machine (SVM) classifier to identify the user's emotional state, and dynamically selected matching audio clips from a preset music library [10]. Habashi *et al*. developed a real-time music generation model based on generative adversarial network (GAN), which took the user's real-time EEG signals as input and generated personalized music synchronized with brain wave rhythms. The research adopted binaural beat technology, embedding alpha waves or theta waves into background music to induce brain wave synchronization and found that music therapy could improve the psychological state of patients and promote the recovery of physiological functions to a certain extent [11]. Stress can be reduced by listening to music, and people are influenced differently by different types of music. To solve the problem of automatic music classification, Wang *et al*. used a novel clustering technique, K-MeansH, for pre-gloss processing of music and offered music recommendations based on individual interests using a collaborative filtering process that was individualized [12]. The content-based recommendation engines are used in the majority of current music recommendation systems. However, consumers' musical choices are influenced by their emotions as well as past tastes and musical substance. An emotion-based music recommendation system was presented by Faye *et al*., which used wearable physiological sensors to gather indications about users' feelings [13]. Liu *et al*. designed a personalized music recommendation system that incorporated emotional awareness by linking user data with music and suggested the use of physiological sensors to capture users’ emotional states, which were then utilized to improve the performance of recommendation engines [14]. Similarly, Huang *et al*. enhanced rhythm-based music classification through an ensemble method that combined dynamic classification using long-term modulation spectra with sequence classification using short-term spectra. By aligning music classification with rhythmic patterns, user preference models were better represented, leading to greater satisfaction [15]. In response to the rapid growth of information resources in the era of big data, Hai *et al*. introduced an incremental Slope-One algorithm with the benefit of the ability to adjust to the real-time changes in data, which greatly aided in enhancing recommendation systems' performance [16]. Currently, most research focuses on the development of music recommendation systems and the application of emotion recognition technology, while there is relatively little research on physiological signal monitoring and real-time feedback mechanisms, which is also a key challenge affecting the dynamic adjustment of music intervention programs. Human-computer interaction (HCI) is the process of exchanging information between humans and computers, where certain tasks are completed by using a particular dialogue language and interaction mode to examine how users and systems interact. A variety of machines, as well as computerized systems and software, can be considered systems. The HCI function is mainly realized by input and output external devices and software, which starts with the real-time collection of multimodal physiological signals. Through wearable devices such as smart bracelets and heart rate monitors, the indicators of the user's autonomic nervous system activity are captured [17, 18]. The initial physiological data can be further transformed into emotional states with the aid of machine learning and affective computing technologies. After the system completes the identification of the user's state, it enters the decision-making and execution stage of music intervention.

This study developed an intelligent music therapy system based on HCI technology, which integrated physiological monitoring, music recommendation algorithms, and real-time feedback mechanisms to solve the problems of insufficient personalization, lack of real-time evaluation, and limited professional resources in traditional music therapy. The concept of a music therapy system that could achieve real-time and effective monitoring of user status and obtain user experience and physiological information was introduced, and the algorithm and prototype of the system were presented. By using deep learning, a music recommendation system capable of intelligent recommendation was developed to improve users' music healing experience. Based on the analysis of real physiological data of people, the effectiveness of a new type of music for healing was explored. This study established a correlation model between music intervention and physiological indicators, providing objective evidence for the scientific development of music therapy and filling the gap in research on the neural mechanisms of music therapy. In addition, the proposed adaptive music recommendation algorithm provided new ideas for personalized medical intervention and could be extended to other non-pharmacological intervention fields.

Materials and methods

HCI technology of music therapy

The Unwind program (Mindful Technology Inc, San Francisco, California, USA) was employed in this study, which was a music-based biofeedback system designed to assist relaxation through music. By providing individualized music recommendations and real-time adjustment, this program continuously analyzed the user's physiological data including skin conductance and heart rate and modified the music parameters in response to changes in these signals. The Unwind program mainly consisted of a data processing module, a sound synthesis module, and a sound library. The user's blood volume pulse signal was measured using a finger clip type photoelectric capacitance pulse wave sensor and transmitted to the Unwind program through a biosensor device. In the data processing module, the heart rate interval and heart rate variability index were calculated as biofeedback data. In the sound synthesis module, the biofeedback data was mapped onto the parameters of the sound synthesizer to control the output of the audio. The sound synthesizer adjusted music parameters based on the biofeedback data to achieve personalized music recommendation [19]. The sound synthesis module selected appropriate sounds and adjusted parameters according to the biofeedback data to achieve the optimal relaxation effect. Eventually, the Unwind application used the user's physiological signals to modify the music parameters in real time.

After entering the formal therapy phase, the Unwind program was switched to continuous monitoring and dynamic response mode. The system adopted a 30 second sliding window to perform feature extraction and state recognition on the collected physiological data. When the algorithm detected the physiological stress index value higher than the baseline level, which indicated that the user was in a state of stress or tension, the program activated the music parameter adjustment mechanism. At the end of the entire therapy session, Unwind automatically generated a visual report, enabling users to intuitively understand the impact of music intervention on their own state and thus forming a complete therapy loop. In the state recognition module, this system adopted SVM as the core classifier to construct a two-dimensional emotion space of "stress-relaxation". Owing to its excellent classification performance in small-sample and high-dimensional data scenarios, SVM was particularly suitable for individualized physiological data analysis. The system collected physiological data under different psychological tasks through preliminary experiments and used subjective scales for label annotation to train an SVM model to distinguish between high stress, neutral, and relaxed states. To improve the robustness of the model, radial basis function kernels were used in combination with grid search to optimize hyperparameters. After training, the model could classify the current physiological feature vector in real time, output the probability distribution of the user's emotional state, and further calculate the continuous physiological regulation index (PRI) as the decision-making basis for music intervention. The music parameter regulation module dynamically adjusted the rhythm, tonality, acoustic structure, and other parameters of the music based on the PRI value, achieving personalized music intervention. Iterative training of the neural network model using the random gradient descent approach could be accomplished by extracting a batch of samples and adjusting the parameters. Each sample's granularity was represented by the loss function as follows.

$$J(\theta_0,\theta_1) = \frac{1}{2}\left(h_\theta\left(x^{(i)}\right) - \left(y^{(i)}\right)\right)^2 \tag{1}$$

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0,\theta_1) \tag{2}$$

where $J(\theta_0,\theta_1)$ was the squared error loss. $h_\theta\left(x^{(i)}\right)$ was the predicted value of the model. $y^{(i)}$ was the true value. θ_j was the parameters to be updated. α was the learning rate. $\frac{\partial}{\partial \theta_j}$ was for the gradient of the loss function. Only binary classification issues could be solved using the logistic function. However, multicategory classification problems could be solved by using its polynomial regression. If the Softmax function was $\sigma(z)$ and the input z was a vector of C dimension, the output of the Softmax function was also a vector y of C dimension as follows.

$$y_c = \sigma(z)_c = \frac{e^{z}c}{\sum_{d=1}^{C} e^{z}d} \tag{3}$$

$$\sum_{c=1}^{C} y_c = 1 \tag{4}$$

Softmax was the output layer of neural network, and C neurons in the function represented C classifications. Because the sum of the classification probabilities was 1, the categories classified by Softmax function were mutually exclusive. SVM is a potent regression and classification method that can handle high-dimensional data and spot intricate patterns, which makes it appropriate for user preference learning and the categorization of musical features in music recommendation. Finding the ideal hyperplane that optimized the margin between samples of various classes on both sides of the hyperplane was the aim of SVM. The choice of Kernel functions and optimization algorithms could result in the best classification model. After converting the multiclassification problem into a convex optimization problem, the optimal solution was calculated as follows.

$$y = min\frac{1}{2}\|w\|^2 + C\sum_{I=1}^{R}\xi_i \tag{5}$$

where C was the penalty coefficient, which usually took the default value of 1. w was the normal vector of a hyperplane, which determined the direction of the hyperplane. $\|w\|$ was the L2 norm of a vector w. ξ_i was the relaxation variable of the *i*-th sample. R was a real number. Lagrange objective function was defined as below.

$$L(w,b,a,\xi)=\frac{1}{2}\|w\|^2-\sum_{i=1}^{n}a_i\left(y_i\left(w\cdot x_i+b\right)-1+\xi_i\right) \quad (6)$$

where b was the intercept of a hyperplane that determined its position in space. a_i was the Lagrange multiplier of the i-th sample. y_i was the label of the i-th sample. x_i was the feature vector of the i-th sample. Its dual problem was then solved as follows.

$$\min\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}a_ia_jy_iy_jK\left(x_i,x_j\right)-\sum_{i=1}^{N}a_i \quad (7)$$

where a_i and a_j were the Lagrange multiplier of the i-th and j-th samples. y_i and y_j were the labels of the i-th and j-th samples. $K\left(x_i,x_j\right)$ was Kernel function used to calculate the inner product of samples x_i and x_j in high-dimensional feature space. N was the total number of training samples. $a*$, $b*$, and $w*$ were solved below.

$$a*=\left(a_1^*,a_2^*,\ldots,a_N^*\right)^T \quad (8)$$

$$b*=y_j-\sum_{i=1}^{N}a_i^*y_iK\left(x_i,x_j\right) \quad (9)$$

$$w*=\sum_{i=1}^{N}a_i^*y_ix_i \quad (10)$$

where $a*$ was the optimal solution of Lagrange multipliers. a_i^* was the optimal value of the Lagrange multiplier corresponding to the i-th sample. T was the transposition symbol, representing the conversion of a vector from a row vector to a column vector. $b*$ was the optimal solution for hyperplane interception. y_j was the label corresponding to the j-th support vector. a_i^* was the optimal value of the Lagrange multiplier corresponding to the i-th sample. y_i was the label of the i-th sample. $K\left(x_i,x_j\right)$ was the Kernel function. $w*$ was the optimal solution of hyperplane normal vector. The optimal classification function was eventually obtained as follows.

$$f(x)=sign(w*x+b*) \quad (11)$$

where $f(x)$ was the classification function of SVM. $sign$ was sign function. $w*$ was the optimal solution of hyperplane normal vector. x was feature vectors of input samples to be classified. $b*$ was the optimal solution for hyperplane interception.

Music recommendation based on deep learning

Commonly used recommendation systems relied on three techniques including collaborative filtering-based recommendation, content-based recommendation, and hybrid recommendation. Collaborative filtering-based recommendation started with measuring similarity and used the similarity between users or items to generate relevant recommendations. Within this area, memory-based and model-based methods were two often employed techniques. Content-based recommendation generated user and item representations directly from the intrinsic attributes of the items. In contrast, collaborative filtering built these representations by analyzing aggregate interaction patterns between users and items, which assumed that consumers would be drawn to products that were comparable to ones they had previously dealt with. Hybrid recommendation systems could mitigate the issues that collaborative filtering approaches faced such as cold start or sparse interaction matrix data by utilizing user and item information from content-based recommendation systems [20–22]. Hybrid recommendation systems could improve suggestion performance by including user and item content information, i.e. item auxiliary information and user auxiliary information, into the collaborative filtering framework. In music recommendation, this meant that the model could learn users' preferences and the intrinsic characteristics of music by analyzing a large amount of user music listening history data, audio features of music, as well as users' emotional and physiological responses during music therapy. By analyzing the audio signals of music, the model could extract features such as melody, rhythm, and harmony. By analyzing users' behavioral data, the model could learn users' preferences for different types of music and their music selection patterns in different scenarios. Thanks to this automatic feature learning process, deep learning models could now identify complex patterns that were difficult to extract manually using traditional techniques, resulting in more accurate and personalized recommendations. Simulating the neural network of the human brain to detect external inputs and extract the prominent features of objects from samples was comparable to deep learning. Recommendation systems could represent vast amounts of user and item data by learning deep-level network topologies. The cold start issue in conventional recommendation systems had been somewhat mitigated by using automatic feature learning to identify intricate correlations between data and then combining it with conventional recommendation techniques to finish recommendations. Deep learning-based recommendation algorithms typically started with user-related data, trained a deep learning model to extract the latent representation of users, and then used this latent representation to provide customized recommendation lists. In HCI content recommendation method, a user model was established for each user, which included information such as the user's historical behaviors, interests, and preferences, which might cover the user's browsing history, purchase history, ratings, social relationships, and other related data. The ripple network could dynamically expand and update the knowledge graph through ripple propagation on the knowledge graph. The process of ripple propagation was similar to the diffusion of water waves. It started from a central node, expanded gradually outward, and affected the surrounding nodes. In the human-computer dialogue system, the ripple network could start with the user's current needs, expand the relevant knowledge and information step by step, and thus generate more personalized and intelligent responses. Through the ripple network propagation model with the user's model as the starting point, the user's interests were automatically propagated to the potentially interesting content. During the propagation process, the entities and relationships in the knowledge graph were considered to help select the propagation path and strengthen the relevance. Based on the results of ripple network propagation, a set of candidates recommended contents was generated, which might include movies, music, articles, products, *etc*. to match the user's interests and preferences. The candidate content was sorted and filtered to provide the most relevant and personalized recommendation list. Various sorting methods could be adopted such as collaborative filtering, deep learning, or methods embedding knowledge graphs. The human-machine dialogue model was constructed by clarifying that the input was the content and knowledge graph of the k-th participant and the output was the reply content of the (k+1)-th robot before entering the loop process, which included 7 steps. Step 1 was the filtering out n responses with the highest semantic confidence based on the k-th interaction input of the participants and converting them into vector form to obtain their feature representations. Meanwhile, entity connection method was used to perform entity extraction and disambiguation processing of the dialogue content. Step 2 was the calculation of the emotional friendliness R(k) of the k-th participant's interactive input. Step 3 was to obtain the ripple set of associated entities based on the extracted entities and knowledge graph. Step 4 was the calculation of the content response probability of candidate replies. Step 5 was the calculation of the satisfaction value of candidate responses. Step 6 used the content corresponding to the maximum satisfaction as the reply content. Step 7 updated the value of k (k = k + 1). All those steps were repeated until the participants stopped interactive input and the session terminated. Basic period was typically thought to be the position that corresponded to the first maximum autocorrelation function. Musical notes and the fundamental frequency had a matching relationship. The sampling frequency divided by the fundamental period was known as the fundamental frequency. Data frame shifting was utilized to determine the peak ratio to select the proper peak points for this study. Data frame shifting was the process of raising the chosen signal interval's top and lower bounds by 64, respectively. The shifted signal interval was expressed as below.

$$seg(i) = w[S(i) + 64, S(i) + len + 64] \quad (12)$$

According to the basic theory of music, the relationship between note i and standard frequency $F(i)$ could be expressed as below.

$$F(i) = f_{a^1} * 2^{n/12} \quad (13)$$

where n was the number of semitones between note i and note a^1. $f_{a^1} = 440$ was the first international height. The relationship between pitch period $T(i)$ and pitch frequency $F'(i)$ could be expressed as follows.

$$F'(i) = f_s / T(i) \quad (14)$$

where f_s was the sampling frequency. The partial deviation $O(i)$ was expressed as below.

$$O(i) = \log_k \left(F'(i) / F(i) \right) \quad (15)$$

where $k = \sqrt[1200]{2}$. If $U = \{x \mid -50 < x < 50\}$, when $O(i) \in U$, x was considered that the note i was recognized correctly.

Design of music healing based on Endel platform

Endel (Endel B.V., Amsterdam, North Holland, Netherlands) as an artificial intelligence (AI)-based ambient sound generation platform provided an innovative technical path for this study. The core mechanism of Endel lied in combining ambient sounds, synthetic sound effects, and neuroscience through algorithms to generate a continuously changing audio stream, which was designed to adapt to the user's biological rhythms and psychological needs. In this study, Endel was integrated into the overall HCI framework as the execution terminal for personalized music intervention. The system transmitted the collected physiological data to the Endel platform in real time through an application programming interface (API), and the platform dynamically adjusted its audio parameters. When an increase in the user's sympathetic nerve activity was detected, the system would automatically reduce the rhythm density of the audio and increase the weight of low-frequency components. During the deep relaxation stage, the system would moderately enhance the richness of mid-to-high frequency overtones to prevent the user's consciousness from wandering. This closed-loop regulation mechanism enabled Endel to function as an intelligent acoustic regulator. In terms of sound design, the Endel system avoided complex musical narratives and adopted a simplified acoustic structure to reduce the consumption of neural energy. In the design of the sleep mode, the "sleep cue tone" introduced by the system was not a traditional melodic segment, but a functional acoustic signal. Its motif was based on the pentatonic scale of the C major gong mode, and a slowly decaying pitch contour was constructed through the descending step progression of pure octave intervals. This descending movement was generally associated with a sense of relaxation and finality in psychoacoustics, and its physical manifestation was a gradual decrease in frequency, which corresponded to the release of energy.

Experimental procedures

A pre-test-post-test control group approach was used in this six-week experimental study to comprehensively assess the therapeutic effect of the Endel-based HCI system. A total of 80 individuals including 48 males and 32 females, aged between 25 and 45 years old, with moderate to high levels of work-related stress

(PSS-10 score > 20), were recruited through a number of internet enterprises and financial institutions in Shenzhen, Guangdong, China. The participants were randomly divided into an experimental group (25 males and 15 females) and a control group (23 males and 17 females). The experimental group used intelligent therapy system integrated with the Endel platform, conducting 20 minutes interventions daily, and the system dynamically adjusted the audio content according to their physiological status. The control group used a static playlist without real-time feedback function. All participants completed standardized psychological scale assessments before and after the intervention and wore physiological monitoring devices continuously during the intervention. The collected data included multi-dimensional indicators including heart rate using Polar H10 heart rate sensor (Polar Electro Oy, Espoo, Finland), standard deviation of normal-to-normal intervals (SDNN), root mean square of successive differences (RMSSD), Galvanic skin response (GSR) (GSR Technologies Inc, Pasadena, California, USA) baseline levels, subjective stress perception using perceived stress scale-10 (PSS-10) (Cohen & Associates, New York, NY, USA), and positive and negative affect schedule (PANAS) (University of Illinois, Chicago, Illinois, USA). During the experimental period, the system automatically recorded information such as each participant's physiological response curve and usage compliance data. In the pre-experimental stage (week 1), participants underwent a physical examination to confirm that they had no underlying diseases such as cardiovascular disease or mental illness that might affect the experimental results and had no history of music allergies, completed baseline testing of PSS-10 and PANAS scales, while wearing all physiological monitoring devices, and collected physiological data at rest for 30 minutes as baseline data before intervention. Participants would receive training on experimental operations including the correct wearing and use of physiological monitoring equipment, as well as the operation methods of intelligent healing systems/static playlists. During the intervention phase (weeks 2-7 for 6 weeks), the experimental group received a 20-minute music therapy intervention using an intelligent healing system integrated with the Endel platform at a fixed time, recommended 8 - 10 pm, every day. The system collected physiological data in real-time and dynamically adjusted music parameters. During the intervention process, the device continuously recorded physiological response curves. Control group used a static playlist for 20 minutes of music listening at a fixed time every day (consistent with the experimental group) without real-time physiological feedback or music parameter adjustment. The device also continuously recorded physiological data. During the experiment, participants were required to fill out daily intervention logs and record information such as intervention time and subjective feelings. The system would automatically calculate the frequency of use. In the post test phase (week 8), participants completed PSS-10 and PANAS scale tests again and compared them with baseline data before intervention, wore physiological monitoring equipment, collected physiological data at rest for 30 minutes as post intervention data, and conducted comprehensive analysis with baseline data and dynamic data during the intervention process. All procedures of this study were approved by the Medical Ethics Committee of Shenzhen University (Shenzhen, Guangdong, China) (Approval No. 3688).

Statistical analysis

All data in this study were analyzed using SPSS 26.0 statistical software (IBM Corp., Armonk, NY, USA) with a significance level set at $\alpha = 0.05$. Independent sample t-test was used to compare the baseline data balance between the experimental group and the control group before intervention, ensuring comparability between the two groups. After the intervention, covariance analysis was used to compare the differences between the two groups in physiological indicators, psychological scale scores, and usage frequency. Baseline data before the intervention were used as covariates to control the influence of baseline levels.

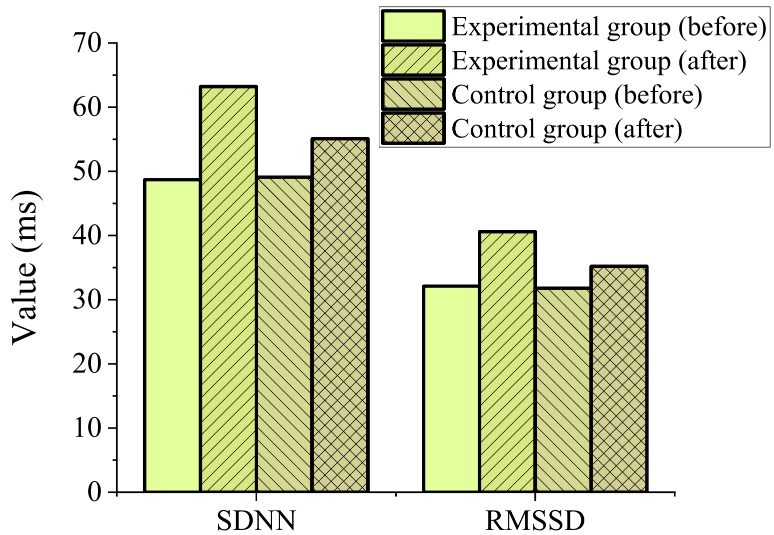

Figure 1. Comparison of HRV indexes between two groups of participants.

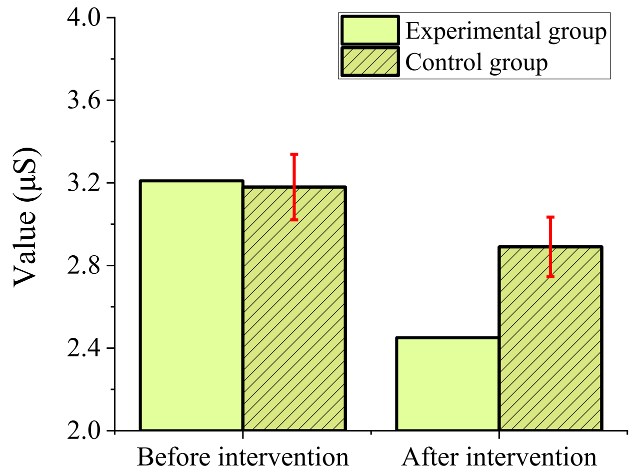

Figure 2. Comparison of baseline levels of skin conductance between two groups of participants.

Results and discussion

Impact of music therapy on HRV

The changes in HRV indicators of participants in the two groups before and after the intervention demonstrated that, in the experimental group, SDNN increased from a baseline value of 48.7 ms to 63.2 ms, an increase of 29.8%, while RMSSD increased by 26.4% compared to that of the control group with 12.3% and 9.7% increase, respectively ($P < 0.01$), indicating a substantial improvement in the participants' autonomic nervous regulation ability (Figure 1). The result verified superiority of dynamic music intervention based on physiological feedback in promoting physiological relaxation. The changing trend of GSR levels showed that the baseline GSR value of the experimental group decreased significantly from 3.21 μS before the intervention to 2.45 μS with a reduction rate of 23.7%, which reflected a continuous decrease in the level of sympathetic nerve arousal. However, the control group only showed a 9.3% decrease, and some participants experienced a rebound in the later stage, which implied that the long-term effect of static intervention was limited (Figure 2).

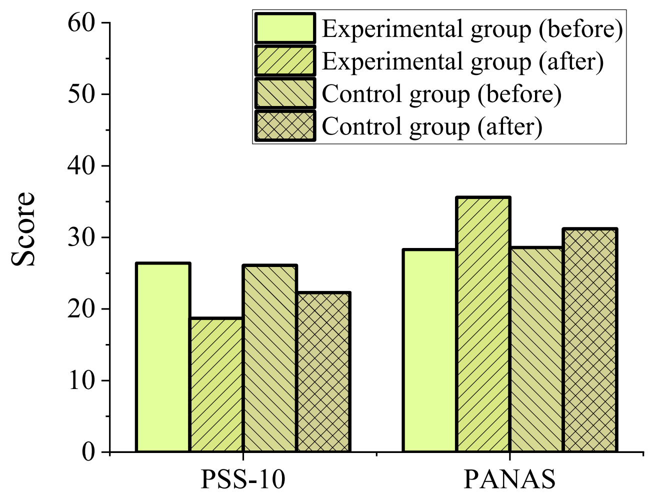

Figure 3. Comparison of scores of psychological scales between the two groups.

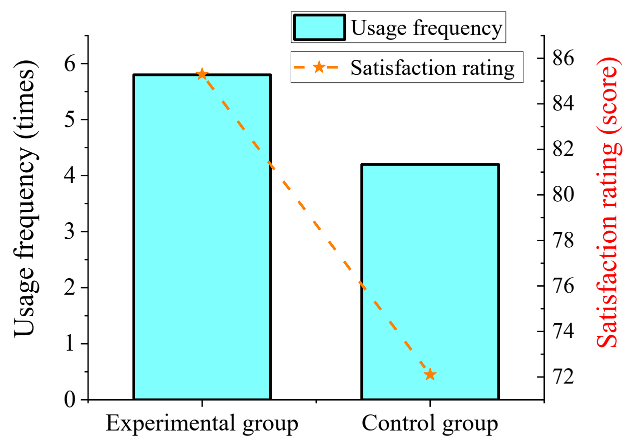

Figure 4. Comparison of intervention compliance between two groups of participants.

Effect of music healing system on mental state improvement

In terms of subjective psychological state, the psychological scale comparison results showed that the PSS-10 score of the experimental group decreased from 26.4 to 18.7, a reduction of 29.2%, which was close to the clinically significant improvement standard (≥ 30%). The PANAS score also increased, indicating an improvement in the emotional state. In contrast, the improvement magnitude of the control group was relatively slight (Figure 3). The system's user compliance demonstrated that the experimental group's average weekly usage frequency was 5.8 times, which was substantially greater than the control group's 4.2 times ($P < 0.05$) (Figure 4). The results confirmed that the dynamic feedback mechanism could enhance users' willingness to participate.

Conclusion

A music therapy system based on HCI was constructed in this study, which integrated multiple music therapy methods and was integrated into the standard music therapy management system through a group management component. By virtue of emotion recognition, physiological signal monitoring, and

SVM-based music recommendation, the system realized the personalization and dynamic adjustment of music therapy. Even though this study combined physiological signal monitoring and emotion recognition algorithms, multimodal data fusion could yet be improved. More efficient multimodal data fusion techniques must be investigated in future studies to improve the system's accuracy and resilience. Furthermore, to improve recommendation performance in the future, deep learning algorithms like convolutional neural network and recurrent neural network can be further tuned.

References

1. Shah B, Mukherjee B, Sundar S. 2024. Use of technologies in music therapy clinical practice. SBV J Basic Clin Appl Health Sci. 7(2):80-95.
2. Osorio Torres J, Shojaei F, Shih PC. 2024. Classifying technologies during the assessment, treatment planning, documentation and evaluation phases of music therapy: A survey of board-certified practitioners. Proc ACM Hum Comput Interact. 8(CSCW2):1-23.
3. Rahman JS, Gedeon T, Caldwell S, Jones R, Jin Z. 2021. Towards effective music therapy for mental health care using machine learning tools: Human affective reasoning and music genres. J Artif Intell Soft Comput Res. 11(1):5-20.
4. Li Y, Liu Y, Yuan R, Zhang H. 2025. Meditation and music therapy in the era of digitalization: Challenges and future considerations. J Posthumanism. 5(5):4127–4138.
5. Kumar V, Pavitra KS, Bhattacharya R. 2024. Creative pursuits for mental health and well-being. Indian J Psychiatry. 66(Suppl 2):S283-S303.
6. Nunes IB, de Santana MA, Charron N, Silva HSE, de Lima Simões CM, Lins C, *et al*. 2024. Automatic identification of preferred music genres: An exploratory machine learning approach to support personalized music therapy. Multimed Tools Appl. 83(35):82515-82531.
7. Agres KR, Schaefer RS, Volk A, van Hooren S, Holzapfel A, Bella SD, *et al*. 2021. Music, computing, and health: A roadmap for the current and future roles of music technology for health care and well-being. Music Sci. 4:2059204321997709.
8. Bao SC, Li F, Xiao Y, Niu L, Zheng H. 2023. Peripheral focused ultrasound stimulation and its applications: From therapeutics to human–computer interaction. Front Neurosci. 17:1115946.
9. Ragone G, Good J, Howland K. 2021. How technology applied to music-therapy and sound-based activities addresses motor and social skills in autistic children. Multimodal Technol Interact. 5(3):11.
10. Dhiman HS, Deb D, Muyeen SM, Kamwa I. 2021. Wind turbine gearbox anomaly detection based on adaptive threshold and twin support vector machines. IEEE Trans Energy Convers. 36(4):3462-3469.
11. Habashi AG, Azab AM, Eldawlatly S, Aly GM. 2023. Generative adversarial networks in EEG analysis: An overview. J Neuroeng Rehabil. 20(1):40.
12. Wang E, Lee H, Do K, Lee M, Chung S. 2022. Recommendation of music based on DASS-21 (Depression, Anxiety, Stress Scales) using fuzzy clustering. Electronics. 12(1):168.
13. Faye P, Sonar S, Shahare S, Ambhore R, Chafle V. 2025. Emotion based music recommendation system. Int J Adv Electr Comput Eng. 14(1):186-191.
14. Liu Z, Xu W, Zhang W, Jiang Q. 2023. An emotion-based personalized music recommendation framework for emotion improvement. Inf Process Manag. 60(3):103256.
15. Huang B, Wang Z, Shang J, Chen G. 2022. A spectral sequence-based nonlocal long short-term memory network for hyperspectral image classification. IEEE J Sel Top Appl Earth Obs Remote Sens. 15:3041-3051.
16. Hai T, Zhou J, Srividhya SR, Jain SK, Young P, Agrawal S. 2022. BVFLEMR: An integrated federated learning and blockchain technology for cloud-based medical records recommendation system. J Cloud Comput. 11(1):22.
17. Su T, Ding Z, Cui L, Bu L. 2024. System development and evaluation of human–computer interaction approach for assessing functional impairment for people with mild cognitive impairment: A pilot study. Int J Hum Comput Interact. 40(8):1906-1920.
18. Lyu Z. 2024. State-of-the-art human-computer-interaction in metaverse. Int J Hum Comput Interact. 40(21):6690-6708.
19. Xu L, Hou J, Gao J. 2021. A novel smart depression recognition method using human-computer interaction system. Wirel Commun Mob Comput. 2021(1):5565967.
20. Liu S, Li G. 2023. Analysis of the effect of music therapy interventions on college students with excessive anxiety. Occup Ther Int. 2023(1):3351918.
21. Bondioli M, Chessa S, Kocian A, Pelagatti S. 2024. A survey on technological tools and systems for diagnosis and therapy of autism spectrum disorder. Hum Comput Interact. 39(3-4):145-173.
22. Parthasarathy G, Sathiya Devi S. 2024. Hybrid recommendation system based on collaborative and content-based filtering. Cybern Syst. 54(4):432-453.