RESEARCH ARTICLE

The role of music teaching based on deep learning in relieving the mental health stress of college students

Rui Zhou[*]

Department of Music Performance, School of Culture and Education, City University of Zhengzhou, Zhengzhou, Henan, China.

Received: February 18, 2025; **accepted:** August 27, 2025.

Under the influence of academic pressure, life changes, and other factors, college students face greater mental health problems. This study explored the role of music teaching based on deep learning in relieving the mental health stress of college students. Deep learning algorithm was used to intervene the mental health of college students, and personalized music teaching program was designed to evaluate the impact on students' emotional state and mental health level. The results showed that, after the intervention of music teaching, the emotional and mental health scores of the experimental group were significantly improved compared with that of the control group. The results suggested that music teaching intervention based on deep learning could improve students' emotional state and provide a new method and idea for mental health education in colleges and universities. This study provided practical basis and theoretical support for promoting the individualized and scientific development of mental health intervention.

Keywords: deep learning; music teaching; mental health; emotion management.

[*]**Corresponding author:** Rui Zhou, Department of Music Performance, School of Culture and Education, City University of Zhengzhou, Zhengzhou, Henan 452300, China. Email: zhou-rui168@outlook.com.

Introduction

College students face psychological pressures in their studies, career development, and personal life, and their mental health problems have become the focus of social attention. Music is a form of emotional expression that relieves stress and improves mood. With the rapid development of deep learning technology, music teaching model based on artificial intelligence is gradually applied to the field of education and mental health. Music can help regulate emotions and improve personal mental health in certain situations.

There have been many studies in the field of emotion management and mental health. Aizenstein *et al*. discussed the application of deep learning in the mental health of the elderly and found that deep learning could predict the emotional changes and mental health status of the elderly and provide scientific support for emotional regulation intervention [1]. Estrada-Molina *et al*. applied deep learning in an open learning environment and proposed that real-time monitoring of emotional changes by intelligent systems could improve the effect of emotion management [2]. Himeur *et al*. demonstrated the potential of deep learning in detecting emotional changes in public health through the application of mask detection technology in smart cities and proposed that real-time monitoring of public emotions by technical means could effectively improve urban public health management capabilities [3]. Iqbal *et al*. applied deep learning to detect emotions in breast cancer patients, discussed the auxiliary role of emotion regulation in cancer treatment, and proposed that deep learning could monitor patients' emotional states in real time and help adjust treatment plans [4]. Vasconcelos *et al*. reviewed deep learning techniques in fire detection, emphasizing the application potential of deep learning in emotion management and behavior prediction under dangerous situations [5], while Zhang *et al*. studied the emotional management training under the "Internet +" mode to effectively relieve emotional distress and improve the psychological adjustment ability of patients with depression [6]. In addition, Porter *et al*. studied the cognition of mental health of patients with mental illness, emphasized the impact of emotional regulation on their overall health, and suggested that emotional management could not only improve mental health but also enhance patients' sense of well-being [7]. Further, Warrender *et al*. found that, in the UK, mental health nursing education faced the challenge of dilution and emphasized the necessity of emotional management skills in nursing education and suggested that professional emotional management training could improve the therapeutic effect of nursing staff [8]. Brandao *et al*. proposed the mediating relationship between emotional goals and emotional regulation and pointed out that the setting of emotional goals had an important impact on the effectiveness of emotional regulation, providing support for the theoretical framework of emotional regulation [9].

This study proposed a deep learning model to conduct music teaching intervention for college students to evaluate the impact of music on relieving student's mental health stress. Questionnaire, psychological test, and emotion analysis were used in this study to investigate college students from different schools. The proposed model was then optimized using deep learning algorithm to tune the parameters. The stability and generalization ability of the model were verified by cross-validation method, and the influence of music teaching on college students' mental health was evaluated by psychological evaluation index. The research results provided theoretical support and practical basis for the cross-application of educational psychology and artificial intelligence in future.

Materials and methods

Data collection and sample selection

The data of college students' mental health status were collected by online questionnaire and mental health assessment, while the experimental data were collected through regular music teaching intervention. All procedures of this research were approved by the Institutional Review Board, School of Psychology, Shandong University, Jinan, Shandong, China (Approval No. SDU-IRB-2024-219). All participants were provided with informed consent before participation. To ensure the diversity and breadth of the sample, a total of 1,000 students aged 20.4 ± 1.2 years old from five Chinese higher education institutions including Shandong University (Jinan, Shandong, China), Ocean University of China (Qingdao, Shandong, China), Nanjing University of the Arts (Nanjing, Jiangsu, China), South China Normal University (Guangzhou, Guangdong, China), and Zhejiang University (Hangzhou, Zhejiang, China) with different majors (33.2% arts and humanities, 34.5% sciences, 32.3% engineering and technology), grades (25.3% freshmen, 25.6% sophomores, 24.5% juniors, and 24.6% seniors), and genders (498 males and 502 females) were involved in the study [10]. A combination of online and offline methods was used to collect data [11]. Mental health questionnaire of Symptom Checklist-90-Revised (SCL-90-R) (Pearson Clinical & Talent Assessment, Beijing, China) and emotional assessment scale of Positive and Negative Affect Schedule (PANAS) (Multi-Health Systems Inc., Toronto, Ontario, Canada) were distributed through the network platform to obtain relevant data on students' emotional state, stress level, anxiety and depression. In the part of experimental data, deep learning platform was used for music teaching intervention, and all experiments were conducted through standardized procedures to record real-time data of participants' emotional changes and mental states. Individuals who had been diagnosed with severe mental illness were excluded to avoid the influence of disturbing factors.

Data preprocessing and cleaning

Integrity checks were carried out on all collected questionnaire data, and incomplete or invalid questionnaires were excluded [13]. The answers of emotional assessment scale and mental health scale were uniformly coded and standardized to ensure the consistency of data. Extreme values and outliers were identified using boxplot analysis and standardized residuals. These anomalies were either corrected or removed based on their influence on statistical distribution and model sensitivity. The outliers in emotion assessment were determined by the method of box graph and so on, and the appropriate correction or elimination was carried out. All data were processed with missing values, and the data with fewer missing values were filled with the mean filling method [14]. For more missing variables, the corresponding data interpolation was carried out. A total of 1,000 questionnaires were received, of which 128 were incomplete or logically inconsistent and were excluded from the dataset. The remaining 872 records were cleaned and standardized. Outliers were identified, and 37 extreme values were adjusted or removed based on their deviation from interquartile ranges. Missing data under 5% of total responses were filled using mean imputation, while more substantial gaps were interpolated based on linear regression patterns observed in related variables. After cleaning, the valid sample size was reduced to 860 entries with consistent, standardized values.

Descriptive statistics of data

The cleaned data was used for descriptive statistical analysis to gain a preliminary understanding of the mental health status of the sample group [15]. The main indicators of descriptive statistics included mean value, standard deviation, minimum value, maximum value, and quartile. All statistical analyses were conducted using SPSS 27.0 (IBM, Armonk, NY, USA). For comparative analysis, paired t-tests were applied to assess within-group changes before and after intervention, and independent-samples t-tests were used to compare differences between the experimental and control groups. Significance levels were set at $P < 0.05$.

Deep learning model selection

A deep learning model suitable for time series data and sentiment analysis was selected for this research. Based on the time series characteristics of music teaching and the dynamic changes of participants' emotions, constitutional neural network (CNN) that has strong feature extraction capabilities to extract key features from music data and long short-term memory network (LSTM) that is excellent at processing time series data to capture long-term dependencies of mood changes were used as the main models for comparison and fusion [16]. Combining the local feature extraction capability of CNN and the time-dependent modeling capability of LSTM, a hybrid model structure was designed, which could effectively process input data and predict mental health changes under different situations.

Network architecture design

The network architecture was designed based on deep learning constitutional layers and LSTM units [17, 18]. The model consisted of three constitutional layers, which were used to extract local features from music data, and two LSTM layers, which were responsible for capturing mood changes in time series. In the design of the constitutional layer, a 3 × 3 constitutional kernel was used to ensure that data features could be extracted at multiple levels. To avoid over fitting, a maximum pooling layer was added after each constitutional layer, and a dropout layer was added after the LSTM layer for regularization. The output layer used the sigma activation function for binary classification problems such as mood improvement or not [19]. The network architecture of the model could balance feature extraction and time-dependent modeling and improve the prediction accuracy of college students' mental health state.

Parameter optimization and tuning

Parameter optimization was applied to improve model performance. In the optimization process, the learning rate, batch size, the number of filters in the constitutional layer and the number of LSTM units were imaginatively adjusted. Grid search method and cross-validation technique were used to select the best combination from multiple candidate parameters [20]. The learning rate was adjusted from $1e^{-5}$ to $1e^{-2}$, while the batch size varied from 32 to 128, and the number of constitutional layer filters was progressively increased from 16 to 128 to find the parameter settings that best fit the data set. Through hyper parameter tuning, a more stable and accurate model configuration was obtained. In the process of hyper parameter selection, considering the balance between the convergence speed, training time, and test accuracy, the combination of learning rate of $1e^{-4}$, batch size of 64, number of constitutional layer filters of 64, and number of LSTM units of 128 were selected. This configuration showed high prediction accuracy on verification set and could effectively avoid overfitting during training.

Loss function and activation function

In the process of model training, the choice of loss function directly affected the convergence speed and final accuracy of the model. The binary cross-entropy loss function was applied to the binary classification problem to determine whether the emotion was effectively alleviated [21]. The loss function could measure the difference between the predicted result and the actual label and guide the updating of the model parameters during the back propagation process and was calculated below.

$$L = -[y\, log(\hat{y}) + (1-y)\, log(1-\hat{y})] \qquad (1)$$

where *y* was the actual label and the predicted output of the model. To ensure the nonlinear mapping ability of each layer in the network, the constitutional layer and the LSTM layer used the rectified linear unit (Rel U) as the activation function, which could alleviate the gradient disappearance problem and accelerate the training process of the model [22]. The final output layer used the sigma activation function to map the network output to a probability value between [0, 1], which was suitable for binary classification problems. The activation function was expressed below.

$$f(x) = max(0,\ x) \qquad (2)$$

while the sigma activation function was as follows.

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad (3)$$

where *x* was the input value and the probability value of the output [23]. This configuration ensured that the model could effectively capture emotional changes during training and predict the degree of improvement in college students' mental health based on music teaching interventions.

Model evaluation index

The evaluation of the deep learning model was based on five key metrics including accuracy, precision, recall, F1-score, and area under the curve (AUC). Accuracy reflected the overall proportion of correctly predicted cases, while precision measured the proportion of true positive predictions among all positive predictions, which assessed the model's ability to minimize false positives. Recall evaluated the proportion of true positives identified from all actual positives, highlighting the model's sensitivity. The F1-score as the harmonic mean of precision and recall balanced the trade-off between them and provided a single comprehensive indicator. AUC quantified the model's discriminative power across all classification thresholds, particularly useful when dealing with imbalanced datasets. These integrated evaluation indicators provided a holistic view of model performance in predicting emotional improvement outcomes.

Cross-validation and verification methods

To ensure the generalization ability and stability of the model, cross-validation method was adopted. In the process of cross-validation, the whole data set was divided into k subsets. Each subset was used as validation set in each validation, and the remaining k - 1 subsets were used as training sets, on which the model was repeatedly trained and validated. The performance indicators of the model were the average of the results of each verification round. This method could evaluate the performance of the model on different data sets and avoid the bias that might be caused by the division of a single training set. To calculate the error of cross-validation, the cross-validation formula was used as follows.

$$\text{CV Error} = \frac{1}{k}\sum_{i=1}^{k} \text{Error}_i \quad (4)$$

where Error_i was the error in round i of cross-validation. k was the total number of folds. The average cross-validation error across all folds was calculated to ensure the stability of the model on multiple data sets and find performance differences under different data divisions to avoid overfitting and optimizing model configuration.

Results and discussion

Comparison of the music intervention effects

The baseline distribution of mental health indicators including anxiety, depression, and mood state among the 860 valid student participants before the music intervention showed a normal distribution with moderate variance, providing a reliable basis for later comparative analysis (Figure 1).

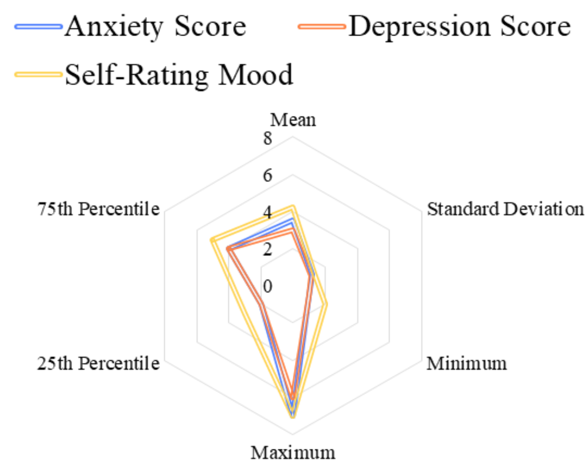

Figure 1. Descriptive statistics of the samples before music intervention.

The results demonstrated that music teaching intervention had a positive effect on relieving the mental health stress of college students. There were significant differences in mental health assessment scores between the experimental group who received music intervention and control group who did not receive music intervention. The mental health assessment data of the experimental group and the control group at different time points were summarized and showed that, after 8 weeks of music teaching intervention, the mental health score of the experimental group decreased significantly and the mental health level improved. However, there was no significant change observed in the control group, and the difference between the experimental and the control groups at each time point reached a statistically significant level (P < 0.05) (Figure 2). To verify the effect of music teaching intervention, the effects of different methods were also compared, and the results of emotional changes in the experimental group under different intervention methods showed that music teaching intervention based on deep learning model was superior to meditation and exercise intervention methods in all indicators of mood improvement (Figure 3). In terms of accuracy and AUC value, deep learning model demonstrated stronger ability to predict effect than the other methods.

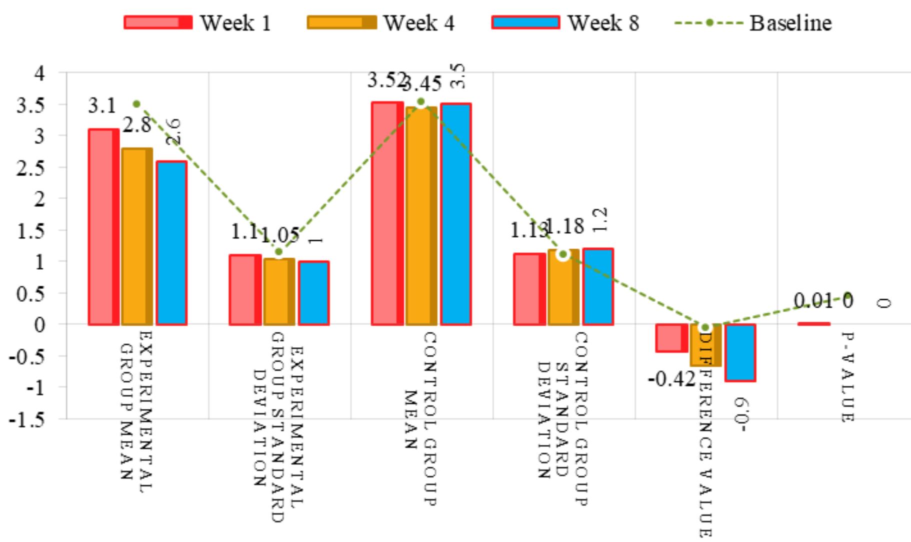

Figure 2. Comparison of the effects of music interventions at different time points.

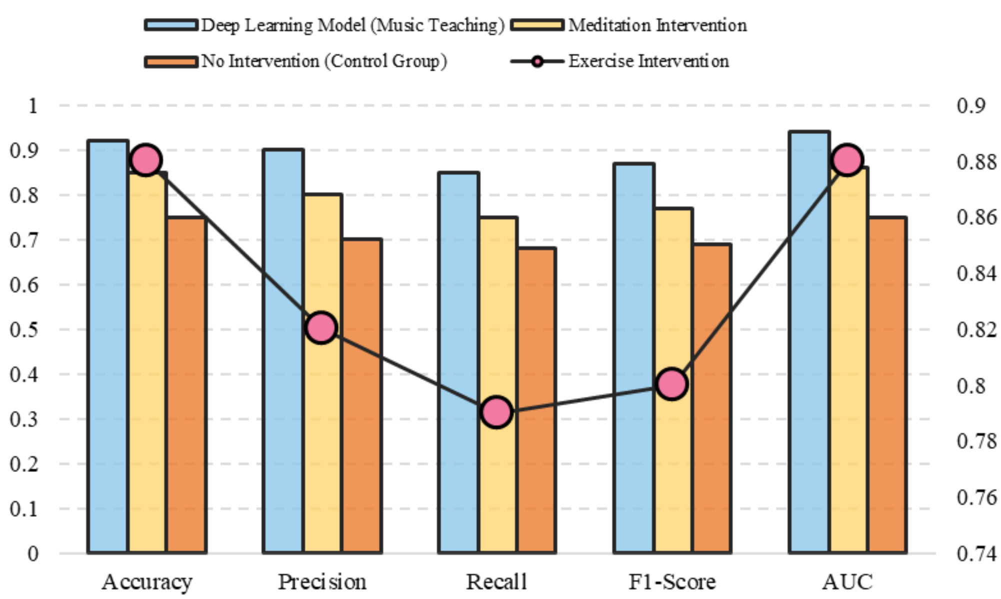

Figure 3. Comparison of different methods.

The performance of the deep learning model was also evaluated and demonstrated that the accuracy of the model reached 89.2%, while the F1-score and AUC were 0.88 and 0.91, respectively, indicating strong classification ability and generalization across various emotional states (Figure 4). During model training, various hyperparameters were tested through grid search. The configuration that achieved optimal performance on the validation set included a learning rate of $1e^{-4}$, batch size of 64, 64 convolutional filters, and 128 LSTM units. This configuration balanced training efficiency and prediction accuracy, achieving the best performance across key evaluation metrics (Table 1). These findings were aligned with prior studies, suggesting that structured music-based interventions could positively affect emotional regulation and stress management among young adults. The results demonstrated that

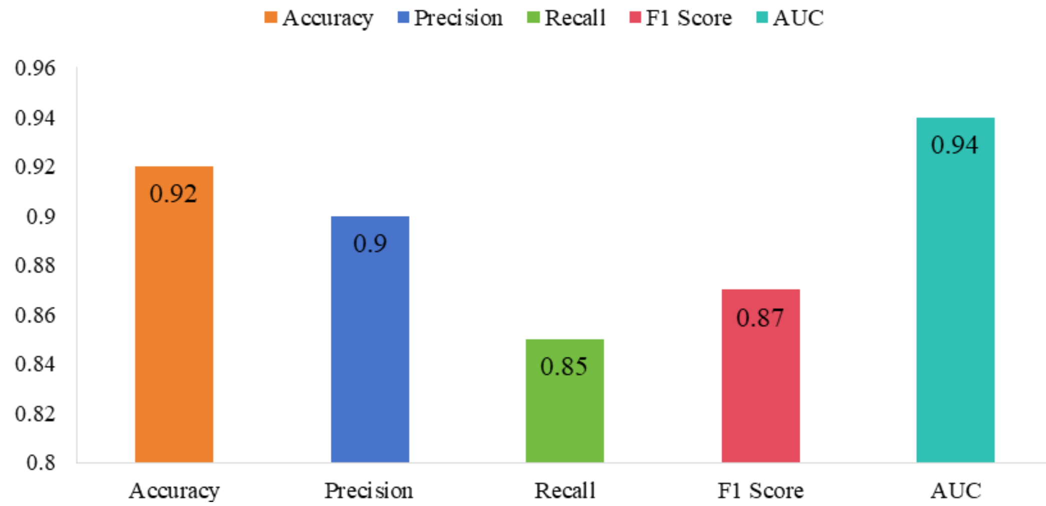

Figure 4. Model evaluation results.

Table 1. Model hyper parameter selection.

Hyper parameter	Candidate values	Optimal value
Learning rate	$1e^{-5}$, $1e^{-4}$, $1e^{-3}$	$1e^{-4}$
Batch size	32, 64, 128	64
Number of convolution filters	16, 32, 64, 128	64
Number of LSTM units	64, 128, 256	128

personalized, AI-assisted music instruction was not only feasible in an educational setting but also effective in enhancing students' emotional resilience. Variability in the degree of improvement among individuals in this study might be attributed to baseline differences in psychological state and music receptiveness, indicating the importance of tailoring interventions.

Practical significance and application scenarios

Deep learning model combined with music teaching intervention had a strong practical significance for relieving the mental health pressure of college students. The high accuracy and AUC values of the model indicated that the method could effectively identify and predict changes in students' mood and mental health after receiving music teaching interventions. In practical application, the intervention method could be used as an auxiliary mental health management tool to help college counselors and mental health professionals to better identify and intervene in the psychological stress of college students. The results of this study also revealed the potential of music instruction in alleviating mental health problems among college students, applicable to those who faced greater psychological burdens due to academic pressure, interpersonal problems, or other factors. Combined with deep learning model, personalized intervention programs for different students could be provided to accurately evaluate their emotional changes during intervention and adjust intervention strategies in real time. The research results provided a new perspective for mental health education in colleges and universities. In addition to traditional mental health intervention methods such as psychological counseling, group counseling, *etc*., music teaching as a non-traditional intervention provided students with a space for relaxation and self-regulation. This intervention could alleviate emotional problems such as anxiety and depression, improve students' emotional stability, enhance psychological adaptability, and promote academic performance and quality of life. Compared to traditional interventions, music-based programs had greater engagement potential among students, particularly when integrated with personalized, real-time deep learning algorithms. The hybrid CNN-LSTM model's capacity to process sequential emotional data from music experience allowed for better emotional pattern recognition and individualized feedback.

Model performance and parameter interpretation

The proposed deep learning model in this study achieved an accuracy of 89.2% with a high AUC of 0.91, indicating robust performance across classification thresholds. These outcomes confirmed the model's reliability in recognizing emotionally improved individuals. Although the performance of proposed model was strong, it still encountered limitations in generalizing outlier emotional responses or atypical music reactions, which highlighted a key avenue for future research-improving multi-modal sensitivity, possibly by integrating physiological sensors or voice analysis to complement subjective input.

Challenges and future research directions

During data collection and analysis, heterogeneity in student backgrounds presented analytical complexity. Differences in discipline, gender, and year influenced baseline stress and response to interventions. Additionally, deep learning models required large, diverse datasets to avoid overfitting, a challenge partially mitigated by parameter tuning and cross-validation. Future work will focus on sample stratification, expansion of data sources, and model generalization across institutions. The integration of multi-modal emotion sensing like facial expressions, heart rate, and hybrid psychological-music-based approaches may yield even stronger intervention efficacy.

Practical implications and theoretical contribution

This study demonstrated that music instruction based on deep learning techniques significantly improved the emotional and mental well-being of college students. The findings confirmed that such non-traditional, technology-integrated psychological interventions were effective and scalable within higher education contexts. The deep learning model successfully predicted mood improvements and personalized interventions, offering strong theoretical support for AI-driven mental health programs. From a practical perspective, the model could assist counselors and educators in early detection and intervention of psychological stress. The combination of music education and artificial intelligence not only offered a novel intervention model but also provided a framework for future interdisciplinary research in mental health, education, and human-computer interaction. Continued development of this approach could expand its applicability beyond academia to include clinical, workplace, and community mental health support systems.

References

1. Aizenstein H, Moore RC, Vahia I, Ciarleglio A. 2024. Deep learning and geriatric mental health. Am J Geriatr Psychiatry. 32(3):270-279.
2. Estrada-Molina O, Mena J, López-Padrón A. 2024. The use of deep learning in open learning: A systematic review (2019 to 2023). Int Rev Res Open Distrib Learn. 25(3):370-393.
3. Himeur Y, Al-Maadeed S, Varlamis I, Al-Maadeed N, Abualsaud K, Mohamed A. 2023. Face mask detection in smart cities using deep and transfer learning: Lessons learned from the COVID-19 pandemic. Systems. 11(2):107.
4. Iqbal MS, Ahmad W, Alizadehsani R, Hussain S, Rehman R. 2022. Breast cancer dataset, classification and detection using deep learning. Healthc (Basel). 10(12):2395.
5. Vasconcelos RN, Rocha WJSF, Costa DP, Duverger SG, Santana MMMD, Cambui ECB, *et al*. 2024. Fire detection with deep learning: A comprehensive review. Land. 13(10):1696.
6. Zhang YP, Wang JX, Qiao LL, Li YT, Hou RN, Gu XJ, *et al*. 2024. Application of "Internet plus" continuous emotion management training in patients with depression. Front Psychiatry. 15:1452717.
7. Porter C, Aggar C, Duncanson K. 2024. People living with mental illness perceptions of physical health, mental health and well-being. Int J Ment Health Nurs. 33(6):2293-2303.

8. Warrender D, Connell C, Jones E, Monteux S, Colwell L, Laker C, *et al*. 2024. Mental health deserves better: Resisting the dilution of specialist pre-registration mental health nurse education in the United Kingdom. Int J Ment Health Nurs. 33(1):202-212.
9. Brandao T, Brites R, Hipólito J, Nunes O. 2023. Emotion goals, emotion regulation, and mental health: A mediational hypothesis. Clin Psychol. 27(3):290-301.
10. Luo B, Lau RYK, Li CP. 2023. Emotion-regulatory chatbots for enhancing consumer servicing: An interpersonal emotion management approach. Inf Manag. 60(5):103794.
11. Li PP, Wang B. 2024. Artificial intelligence in music education. Int J Hum Comput Interact. 40(16):4183-4192.
12. Yan J, Xia XJ. 2024. Interactive audio-visual course teaching of music education based on VR and AI support. Int J Hum Comput Interact. 40(13):3552-3559.
13. Terzi B, Polat S, Dogrusöz LA. 2022. The relationships between novice nurses' emotions management skills and happiness levels: A cross-sectional study. Perspect Psychiatr Care. 58(4):3024-3033.
14. Xu CH. 2022. Self-psychological control and creativity of music creation in multimusic environment. J Environ Public Health. 2022:7849909.
15. Wang XP, Li Y. 2022. Analysis of the effect of music bel canto teaching innovation on college students' phobia. Psychiatr Danub. 34(Suppl 4):S1070-S1074.
16. Zhao S. 2022. Research on the influence of national music education reform on students' employment anxiety. Psychiatr Danub. 34(Suppl 4):S994-S998.
17. Liu CL, Han YF, Li GM. 2018. Study on the relation between self consistency and congruence and mental health of postgraduates. Matrix Science Medica. 2(1):01-03.
18. Hayslip B Jr, Maiden RJ. 2023. Attitudes toward mental health and mental health care among custodial grandparents. Aging Ment Health. 27(1):133-146.
19. Ibrahim N, Ghallab E, Ng F, Eweida R, Slade M. 2022. Perspectives on mental health recovery from Egyptian mental health professionals: A qualitative study. J Psychiatr Ment Health Nurs. 29(3):484-492.
20. Yancey NR. 2020. Music and teaching-learning in nursing: harmony and rhythm in co-constructing new knowings. Nurs Sci Q. 33(1):21-26.
21. Le Cook B, Hou SSY, Lee-Tauler SY, Progovac AM, Samson F, Sanchez MJ. 2019. A review of mental health and mental health care disparities research: 2011-2014. Med Care Res Rev. 76(6):683-710.
22. Wiens K, Bhattarai A, Dores A, Pedram P, Williams JVA, Bulloch AGM, *et al*. 2020. Mental health among Canadian postsecondary students: A mental health crisis. Can J Psychiatry. 65(1):30-35.