How Generative Music Affects the ISO Principle-Based Emotion-Focused Therapy: An EEG Study

Jiayu Bao[1,4,*], **Yaxin Lyu**[1,5,*], **Jingyi Yang**[3✉], **Yucheng Jin**[2✉] **and Jiangtao Gong**[1✉]
[1]Tsinghua University, Beijing, China. Email: gongjiangtao2@gmail.com
[2]Duke Kunshan University, China. Email: yucheng.jin@dukekunshan.edu.cn
[3] Beijing University of Chemical Technology, China. Email: 2024700068@buct.edu.cn
[4] Soochow University, China. [5] Xiamen University Malaysia, Malaysia. * Both authors contributed equally.

Abstract

Recently, AI-generated content (AIGC) technologies have made remarkable advancements, even achieving superhuman performance across various domains. However, few previous studies have investigated its impact on emotion-focused therapy with artistic content, e.g., music. In this paper, we conducted an EEG experiment to explore the effects of generative music on emotion-focused music therapy based on the ISO principle. This experiment compared AI-generated and human-created music regarding the changes in participants' valence and arousal following negative emotion induction with the ISO principle adherence and non-adherence. The results show that generative music, with its harmonic consistency and simple rhythm, is more effective in supporting positive emotions and improving temporal lobe activity. Besides, the therapeutic effectiveness of generative music adhering to the ISO principle has also been validated. This study highlights the distinct emotional and neural mechanisms of AI-generated music, offering valuable insights into future AI-powered emotion-focused therapy strategies.

Keywords: Music Therapy; AI-generated Music; Emotion-focused therapy; ISO Principle; EEG

Introduction

Emotion-focused therapy is a critical component of mental health and well-being, with effective therapeutic interventions playing a pivotal role in addressing conditions such as anxiety, depression, and stress-related emotional disorders (Gross, 2002). Among the various therapeutic approaches, art therapy has emerged as a powerful medium for emotional expression and regulation, leveraging creative processes to facilitate psychological healing (Malchiodi, 2012). Within this domain, music therapy stands out as an effective modality, utilizing intrinsic emotional and structural properties of music to modulate affective states (Thaut & Hoemberg, 2014). Music's unique ability to evoke and regulate emotions has been extensively documented, with evidence suggesting that it can significantly influence both physiological and psychological responses (Ahmad & Rana, 2015; Saarikallio, 2008).

In music therapy, generative AI technologies hold the potential to revolutionize therapeutic practices by enabling the generation of highly personalized and adaptive media tailored to individual emotional and psychological needs (Hung et al., 2021; Sun et al., 2024). While it holds great potential, the role of AI-generated music (AIGM) in emotion-focused therapy remains largely unexplored. Specifically, it is unclear whether AIGM can replicate or even surpass the therapeutic effects traditionally achieved through human-created music (HMCM), especially within a well-established music therapy framework - the ISO principle (Davis, 2008).

Past research has extensively explored the neural mechanisms of music therapy, particularly for HMCM. EEG studies reveal distinct brain activity patterns, such as prefrontal and temporal lobe activation, during emotion-focused therapy tasks (Altenmüller et al., 2002; Lin et al., 2010). Features like power spectral density (PSD) and differential entropy (DE) further serve as reliable markers of emotional states (Zhang et al., 2019). The ISO principle, which aligns music sequences with emotional states, has proven effective in enhancing therapeutic outcomes (Heiderscheit & Madson, 2015; Starcke et al., 2021). However, these insights are largely based on HMCM, leaving a gap in understanding how AIGM influences these neural mechanisms. Specifically, it remains unclear whether AIGM elicits similar neural and self-report responses and integrates effectively into frameworks like the ISO principle. This gap highlights the need for comparative research to assess AIGM's role in emotion-focused therapy.

In this study, we employed a dual-method approach combining EEG neuroimaging and standardized psychometric assessments to systematically compare the effects of AI-generated and HMCM on emotion-focused therapy. Our results demonstrate distinct patterns in how these two types of music influence emotional states, supported by both neural and self-report data. The key contributions of this work are: (1) establishing the differential neural and self-report effects of AIGM versus HMCM in emotion-focused therapy, (2) demonstrating the adaptability of AIGM to sequential therapeutic frameworks like the ISO principle, and (3) through acoustic analysis, uncovering the underlying mechanisms driving the observed neural and self-report differences between AIGM and HMCM. These insights advance the understanding of AI-generated media in therapeutic contexts and provide a foundation for future research on emotion-focused therapy strategies.

Methods

Materials

Human-created Music HMCM was selected from the EMOPIA dataset (Hung et al., 2021), which includes 387 piano solo pieces and has undergone cross-validation based on Russell's Circumplex Model of Affect. We randomly chose three pieces from each combination of valence and arousal conditions to represent both negative and positive music as experimental materials (Figure 1 A, the procedure above). The high cosine similarity (99.97%) between the selected tracks and the original EMOPIA dataset in terms of

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

average music features ensures the selection reliability. Each selected track is one minute long, with some manually edited for consistency. The neutral tracks, however, were chosen based on recommendations from psychotherapists.

AI-generated Music In selecting generative music models for this study, a systematic search was conducted using key terms such as "generative music models", "AI-based music generation", and "machine learning music generation" in Google Scholar (Figure 1 B, the procedure below). Based on the citation and documented performance in various applications, the following generative music models were selected: Udio[1], Audio Craft[2], Suno[3], Stable Audio[4], SongR[5], Mousai[6] and CaiMAP[7]. To ensure the quality and relevance of the selected music, we considered factors such as musical quality as assessed by psychologists and the duration of the compositions. We excluded models that either generated music tracks shorter than one minute or produced audio with significant repetition after extended play, as well as those that failed to meet the prompt requirements (e.g., piano only). As a result, we selected three platforms (Udio, Suno and Stable Audio), with each platform generating piano music based on specific prompts. To ensure the effectiveness of the AIGM, we recruited 155 participants who wouldn't participate in the latter experiment. They were asked to rank the music based on the emotional responses they personally experienced. The ranking was done according to four distinct valence-arousal conditions: 2 (valence: high, low) × 2 (arousal: high, low). Forty data entries were excluded due to incorrect responses on trap questions, leading to 115 available data. Based on the rankings, we selected the three tracks from each emotional condition that most effectively conveyed the intended emotion, resulting in a total of nine music tracks.

Emotion-inducing Movies The movie clips, validated in culturally relevant studies (Deng et al., 2017; Ge et al., 2019), were employed to establish a sadness baseline. This approach ensures three objectives: (1) mirroring real-world music therapy scenarios where negative emotions are modulated; (2) segregating induction (film) from regulation (music) prevents overlap between emotion elicitation and intervention mechanisms; (3) standardized valence and arousal levels before music exposure isolate therapeutic effects, ensuring observed outcomes reflect music-driven modulation rather than baseline variability.

Scale We employed the Self-Assessment Manikin (SAM), a well-known scale that the emotions are visually expressed (Bradley & Lang, 1994), to measure the timely emotional state before and after the implementation of the stimulus.

[1]https://www.udio.com/
[2]https://github.com/facebookresearch/audiocraft
[3]https://suno.com/
[4]https://www.stableaudio.com/generate
[5]https://www.songr.ai/
[6]https://github.com/archinetai/audio-diffusion-pytorch
[7]https://github.com/CarlWangChina/MuChin

Participants

The experiment recruited 22 participants (9 men and 13 women), between 20 and 30 years of age (*Mean* = 22.23, *SD* = 2.70). All participants were regular music listeners without professional film or music study backgrounds. The first four participants were excluded as part of a preliminary test, leaving 18 for the formal experiment. Before the experiment, participants were given a detailed explanation of the procedure but were not informed of the specific research objective to minimize expectancy bias. Written informed consent was obtained from all participants, and the study was approved by the Ethics Committee of the relevant institution, ensuring adherence to ethical guidelines.

Procedure

The experiment follows a 2 (music type: human-created, AI-generated) × 3 (playback principle: ISO, reverse ISO, random) design, consisting of six blocks within participants (as shown in Figure 1, top right). The block order was counterbalanced across participants using the Latin Square design.

Before inducing negative emotions using a film, participants were instructed to remain quiet for two minutes to record EEG data from the resting state as a baseline (Figure 1 C). The SAM scale will be used for pre- and post-tests: the pre-test will be conducted following the resting baseline period, and the post-test will occur after the participant has listened to the final music track. After the pre-test, the participants will watch a video clip to induce negative emotions, followed by completing the SAM scale to assess whether the intended emotional response has been elicited. Subsequently, participants will listen to three 60-second music pieces according to block instructions, with a SAM test conducted after each track to evaluate their emotional state. This procedure will be repeated six times per participant, corresponding to each condition, for a total of 18 music tracks — 9 AIGM and 9 HMCM tracks.

EEG-Data Acquisition

The EEG signals for this experiment were acquired using Curry9 and Eprime software, with a 64-channel Quik-Cap headset (compatible with SynAmps 2/RT and Neuvo amplifiers). The electrodes were arranged following the international 10-20 system, and standardized procedures were applied to ensure data reliability. Certain electrode channels, such as CB1, CB2, EKG, EMG, and TRIGGER, were removed as they provided no directly relevant brain activity information. CB1 and CB2 are reference channels, EKG and EMG record heart and muscle signals (used for marking movement artifacts), and TRIGGER synchronizes external devices, none of which were needed for this analysis.

EEG data were processed by first importing raw signals and standardizing electrode locations (VEO, HEO, G). The data were then re-referenced to M1 and M2 for noise

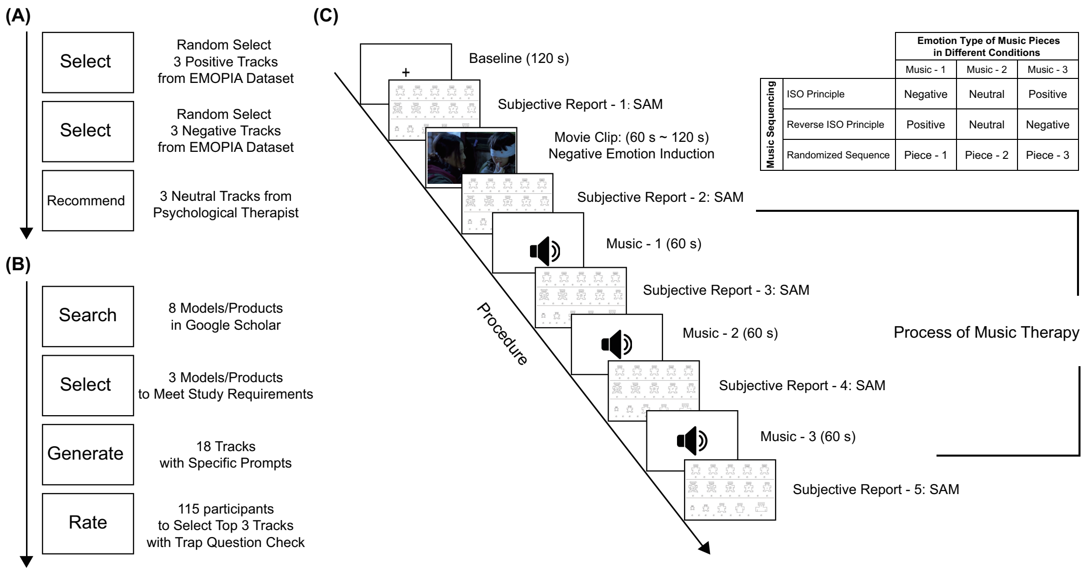

Figure 1: (A) The procedure of selecting HMCMs as materials. (B) The steps to generate and filter AIGMs. (C) The procedure of the experiment in one block. The table in the top-right illustrates the music playing order for each block under different ISO conditions.

reduction, bandpass filtered (0.1-40 Hz) to remove drift and noise, and sampled at 500 Hz for high temporal resolution. Bad channels (e.g., P3) were repaired or interpolated, and Independent Component Analysis (ICA) was used to remove artifacts Homan et al. (1987). After ICA, the data were reduced to 61 components, and artifact components, such as eye muscle artifacts, were identified and removed based on a correlation coefficient range of 0.9-1.

Data Analysis

Subjective Report To ensure data quality, we applied a strict data filtering process to the final dataset. First, we excluded data from participants who failed to show proper arousal responses. The specific criteria for failure were that both valence and arousal values were greater than 0. Emotion valence and arousal data beyond three standard deviations from a participant's mean are excluded. This ensures consistency in emotion induction within the participant. After filtering, 18 experimental sessions were removed.

The results of the sphericity and homogeneity of variance tests confirm that our data meet the ANOVA test prerequisites. We conducted a two-way repeated-measures ANOVA with a 2 (music type) × 3 (playback principle) design to examine the effects of HMCM and AIGM, as well as different ISO music playback principles on emotion-focused therapy. This analysis focused on the overall emotion-focused therapy process, as measured by valence and arousal levels after listening to three musical pieces, along with main effect testing. Subsequently, to explore the influence of music type and playback principles at different stages of the listening process, we performed the same two-way ANOVA at each time point, including analyses of main effects, interactions, and post-hoc tests.

EEG Data For the quantitative analysis of EEG data, six brain regions (prefrontal, frontal, parietal, occipital, temporal, and central) were selected according to the standard of 10-20 system (Homan et al., 1987). Besides the different brain regions, we analyzed the different frequency bands (average 2 Hz delta band, 6 Hz theta band, 12 Hz alpha band, 22 Hz beta band, 40 Hz gamma band) of the EEG result to figure out the potential brain activity and response model. Descriptive analysis and normality test were firstly examined. To focus on the target hypothesis in multiple variables, the paired-sample t-tests, with effect size (ES) analysis, were then conducted to represent the different brain activity across regions and frequency bands.

Results

Consistency Check of Self-Report and EEG Data

This study designs a classification task to investigate the relationship between participants' emotional valence and arousal levels with their electroencephalogram (EEG) signals. Emotional states are categorized into three labels based on self-reported data: high (value > 0), neutral (value = 0), and low (value < 0). The EEG data undergoes preprocessing with baseline correction, followed by feature extraction across all channels using a sliding window approach (4-second window with 2-second step). Key features include PSD, FD, and DE.

Four machine learning models—Random Forest (RF), XGBoost (XG), Support Vector Machine (SVM), and LightGBM (LG)—are employed, with the dataset divided into 70% training and 30% testing sets. Results

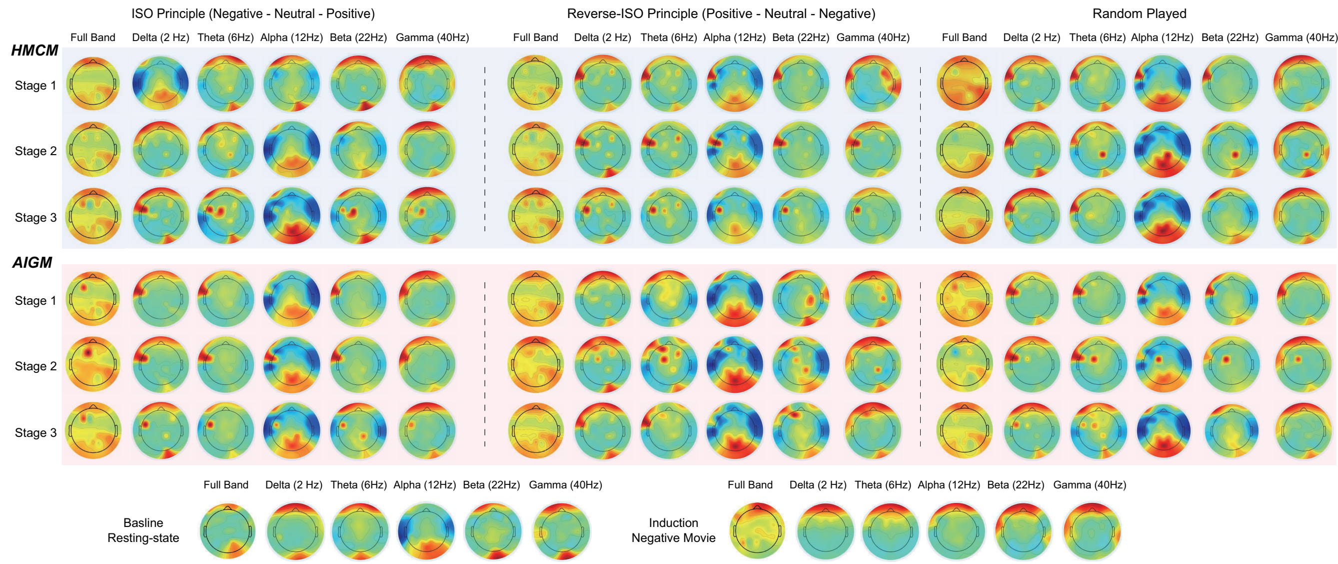

Figure 2: EEG topography of the Experiment. The full band column is the average result for all bands in different brain regions.

demonstrate acceptable classification performance across models according to the state-of-the-art models. For valence classification, LG achieves the highest accuracy of 65.23% (RF: 64.78%, XG: 63.81%, SVM: 64.71%), while XG outperforms others in arousal classification with 72.70% accuracy (RF: 71.78%, SVM: 72.00%, LG: 71.71%). The maximal accuracy difference between the two classification tasks remains below 1.5 percentage points.

Different Intervention Effect in HMCM and AIGM

Main effects of music type ($F(1, 17) = 25.5$, $p < .001$, $\eta^2 = 0.600$) and playback principle ($F(2, 17) = 20.314$, $p < .001$, $\eta^2 = 0.705$) were significant. Post-hoc tests of these main effects revealed several key findings. In contrast, the interaction effect between music type and playback principle was non-significant, with no significant effects for Valence ($F(2, 34) = 0.614$, $p = 0.547$) or Arousal ($F(2, 34) = 0.952$, $p = 0.396$). To further compare the difference in HMCM and AIGM, we selected only the first stage of music therapy, as this stage was aligned after inducing negative emotions, allowing for a direct comparison. Furthermore, due to the design of playback sequence conditions, where Stage 1 involved negative music in the ISO principle condition and positive music in the reverse ISO principle condition, we can conduct additional analyses on emotional responses to negative and positive music respectively.

Key Finding 1: All types of music can produce positive emotional therapy effects. The results showed that all types of music, regardless of their emotional tone, elicited positive emotional experiences (Valence: $t(17) = 6.309$, $p < .001$, $ES = 1.487$; Arousal: $t(17) = 6.254$, $p < .001$, $ES = 1.474$). This finding was further supported by EEG activity in the prefrontal cortex, where significant differences were observed across various musical conditions. Specifically, human-created negative music ($t(17) = 7.760$, $p < .001$, $ES = 1.829$), generative negative music ($t(17) = 6.635$, $p < .001$, $ES = 1.564$), positive HMCM ($t(17) = 7.789$, $p < .001$, $ES = 1.836$), and positive AIGM ($t(17) = 3.691$, $p < .01$, $ES = 0.870$) all elicited robust prefrontal cortex activity. Even with random music segments, both groups demonstrated notable effects: HMCM ($t(17) = 6.907$, $p < .001$, $ES = 1.628$) and AIGM ($t(17) = 5.228$, $p < .001$, $ES = 1.268$). These results indicate that participants were actively engaging in the coordination and processing of emotions related to the music (Dixon et al., 2017). Furthermore, the undifferentiated effect of different types of music on negative emotional states is consistent with previous research (Taruffi & Koelsch, 2014).

Key Finding 2: AIGM generally produced a greater increase in arousal levels According to self-report data, the AIGM group showed a significant increase in arousal after listening to three pieces of music ($t(17) = 2.406$, $p < .05$, $ES = 0.567$, see Table 1, as shown in Intervention Evaluation stage), whereas the HMCM group did not ($t(17) = 1.256$, $p = 0.226$, $ES = 0.296$). EEG results revealed that AIGM led to broader activation, particularly in the Gamma wave band (Figure 2, column *Gamma (2 Hz)*). Specifically, significant increases in Gamma activity were observed in the prefrontal cortex ($t(17) = 2.266$, $p < .05$, $ES = 0.534$) and occipital cortex ($t(17) = 6.008$, $p < .001$, $ES = 1.416$). According to Tallon-Baudry and Bertrand (1999), high complexity or intensity stimuli can induce more pronounced Gamma wave activity in perceptual experiments.

Key Finding 3: Positive AIGM demonstrated superior efficacy in producing positive emotional therapeutic effects. When listening to the positive AIGM music ($t(17) = 2.015$, $p = .01$, $ES = 0.687$, see Table 1 - reverse ISO Principle

Table 1: Analysis of Subjective Emotional Responses (Valence/Arousal)

Music Playing Principle		**ISO Principle**		**Reverse ISO Principle**		**Random Order**	
Music Type		**HMCM**	**AIGM**	**HMCM**	**AIGM**	**HMCM**	**AIGM**
Induction	Mean Difference	-3.200 / -0.200	-2.000 / -0.889	-3.500 / -0.833	-3.333 / -0.667	-2.600 / -0.700	-3.125 / -1.625
	Cohen's d	0.810 / 0.010	0.429 / 0.145	0.751 / 0.100	0.696 / 0.080	0.571 / 0.091	0.636 / 0.298
	p-value	0.003 / 0.967	0.086 / 0.547	0.005 / 0.677	0.009 / 0.738	0.027 / 0.704	0.015 / 0.223
Stage 1	Mean Difference	2.600 / 0.200	1.000 / 0.222	2.500 / 0.833	4.000 / 2.500	2.200 / 0.400	3.125 / 1.000
	Cohen's d	0.690 / 0.005	0.158 / 0.009	0.354 / 0.123	0.687 / 0.455	0.324 / 0.031	0.636 / 0.100
	p-value	0.009 / 0.983	0.512 / 0.970	0.151 / 0.609	0.010 / 0.070	0.187 / 0.897	0.015 / 0.677
Stage 2	Mean Difference	0.400 / -0.400	2.111 / 1.556	2.000 / -1.833	-0.500 / -2.083	0.000 / 1.200	1.125 / -0.125
	Cohen's d	0.018 / 0.020	0.374 / 0.176	0.247 / 0.349	0.017 / 0.347	0.000 / 0.171	0.141 / 0.001
	p-value	0.940 / 0.933	0.131 / 0.465	0.309 / 0.157	0.943 / 0.159	1.000 / 0.478	0.558 / 0.997
Stage 3	Mean Difference	-0.600 / 1.800	0.000 / 0.889	-2.167 / 2.333	-0.750 / -0.833	0.900 / -1.700	-0.375 / 0.500
	Cohen's d	0.021 / 0.325	0.000 / 0.066	0.333 / 0.434	0.003 / 0.130	0.087 / 0.284	0.016 / 0.014
	p-value	0.930 / 0.186	1.000 / 0.783	0.176 / 0.083	0.990 / 0.588	0.717 / 0.245	0.947 / 0.953
Intervention Evaluation	Mean Difference	2.400 / 1.600	3.111 / 2.667	2.333 / 1.333	2.750 / -0.417	3.100 / -0.100	3.875 / 1.375
	Cohen's d	0.340 / 0.296	0.669 / 0.567	0.389 / 0.235	0.395 / 0.042	0.609 / 0.002	0.703 / 0.134
	p-value	0.167 / 0.226	0.011 / 0.028	0.117 / 0.333	0.112 / 0.861	0.019 / 0.993	0.008 / 0.577

with positive emotional music in ths stage), participants showed a more significant improvement in emotional valence compared to HMCM ($t(17) = 1.502$, $p = 0.151$, $ES = 0.354$), with the difference between the two being substantial ($t(17) = 7.675$, $p < .001$, $ES = 1.809$). The EEG findings offer further insight into these results (Figure 2). In response to positive music, the temporal lobe activation induced by (*Mean Difference* [MD] = 21.472) was significantly greater than that triggered by HMCM ($MD = 19.828$), with the difference being statistically significant ($t(17) = 4.696$, $p < .001$). The increased temporal lobe activity associated with positive AIGM may suggest that is more effective in eliciting positive emotion (Adolphs et al., 2002).

Key Finding 4: Negative HMCM showed better positive emotional therapeutic effects. When paired with negative emotions, HMCM ($t(17) = 2.927$, $p = 0.009$, $ES = 0.690$, see Table 1 - the ISO Principle with negative music in the first stage) resulted in a greater increase in emotional valence compared to AIGM ($t(17) = 0.670$, $p = 0.512$, $ES = 0.158$), with the difference being statistically significant ($t(17) = 5.897$, $p < .001$, $ES = 1.390$). Interestingly, the EEG effects observed in positive music were not present in negative music; instead, the pattern was reversed ($t(17) = 1.207$, $p = 0.240$). These findings suggest an interaction between the intended emotional expression of the music and the music type (i.e., HMCM and AIGM) on the therapeutic effect.

Difference between ISO Principle Adherence and Non-adherence

Key Finding 5: AIGM, like HMCM, demonstrated superior emotional therapeutic effects when adhering to the ISO principle. For both AIGM and HMCM, when adhering to the ISO principle, participants' arousal was significantly improved than the non-adherence conditions (while no significant differences were observed in valence). At the third stage, the main effect of the ISO principle was highly significant ($F(2,17) = 29.788$, $p < .001$, $\eta^2 = 0.778$). Post-hoc analyses showed that the ISO condition resulted in significantly greater effects compared to both the reverse-ISO condition ($MD = 1.171$, $p < .01$) and the random condition ($MD = 0.806$, $p < .05$). Across the entire intervention process, a marginal main effect of the ISO principle was identified ($F(2,17) = 3.033$, $p = 0.075$, $\eta^2 = 0.263$). Post-hoc comparisons indicated that the ISO condition showed a moderate difference compared to the reverse-ISO ($MD = 1.318$, $p < .001$) and random condition ($MD = 0.694$, $p < .05$).

Discussion

Underlying Neural Mechanisms of Music Therapy

This study is the first to explore the neural mechanisms of AIGM and HMCM in emotion-focused therapy. Consistent with previous studies, we observed that both AIGM and HMCM effectively improved emotional states, accompanied by significant activation in the prefrontal cortex, aligning with their role in emotion-focused therapy and cognitive control (Dixon et al., 2017). In contrast, positive AIGM induced stronger temporal lobe activation than human-created positive music, correlating with higher emotional valence enhancement (Adolphs et al., 2002). Additionally, across the overall therapeutic process, AIGM significantly increased arousal levels, as evidenced by heightened Gamma-band activity in the prefrontal and occipital regions, indicating enhanced perceptual engagement (Tallon-Baudry & Bertrand, 1999). Additionally, applying the ISO principle with AIGM and HMCM both demonstrated a significant main effect on arousal levels, particularly in stage 3 and across the entire therapeutic process. This suggests that the ISO principle, regardless of whether the music was AIGM or HMCM, outperformed both the reverse ISO and random conditions in enhancing arousal, highlighting its superior efficacy in regulating emotional states.

Acoustic Characteristics Differences Between AIGM and HMCM

To interpret the therapeutic effect difference between AIGM and HMCM, we further analyze the acoustic characteristics of two kinds of music.

First, for chroma variance and rhythmic complexity, AIGM exhibited lower average values for negative, neutral, and positive music as follows: chroma variance of 0.3044, 0.2607, and 0.3170 (compared to 0.3612, 0.3021, and 0.3545 for HMCM), and rhythmic complexity of 0.7351, 0.5240, and 0.6985 (compared to 0.3560, 0.4099, and 0.4590 for HMCM). Pervious studies highlight this kind of harmonic stability in promoting relaxation (Bogatyrenko, 2024; Saarikallio, 2008). The consistent harmonic structure and predictable rhythms of AIGM likely contribute to its ability to evoke more immediate and positive emotional responses. This suggests that AIGM's emotional attributes are more direct and pronounced, making it particularly effective in therapeutic contexts that rely on sequential emotional transitions. These differences are closely related to AIGM's superior short-term arousal enhancement. Participants frequently described AIGM as "harmonious" and "calming", supporting its role in rapid stress reduction. In contrast, the complex rhythmic structures and timbral contrasts in HMCM are likely to evoke more complex or profound emotional reactions, which may also delay immediate emotional improvement, as the listener's cognitive processing of the music's intricate features requires more time.

Second, spectral centroid (SC), a key indicator of timbral brightness (Wun et al., 2014), further differentiated the emotional effects of the two music types. The SC values for AIGM and HMCM in negative, neutral, and positive music are as follows: Negative (AIGM 457.569, HMCM 1009.169), Neutral (627.527, 558.906), Positive (896.773, 1006.584). Negative HMCM exhibited significantly higher SC values than negative AIGM, aligning with participants' immediate valence elevation after listening to human-created negative tracks. This finding resonates with prior research showing that brighter timbres in HMCM are more effective in eliciting rapid emotional shifts (Ryczkowska, 2022). However, the higher SC in negative HMCM may also indicate that it does not fully conform to the typical characteristics of negative music, which are often associated with darker, more subdued timbres. This discrepancy could explain why negative HMCM elicited more positive emotional responses, as its brighter timbre may have inadvertently shifted the emotional tone toward a more neutral or positive valence.

Insight for Using AIGM in Emotion-focused Therapy

Based on the findings of this study, AIGM demonstrates therapeutic efficacy comparable to HMCM in music therapy. Notably, as indicated in Key Finding 2, AIGM elicits significantly stronger arousal responses. Moreover, regarding the therapeutic effects across different types of emotional music, positive AIGM produces more pronounced therapeutic outcomes (Key Finding 3). While negative AIGM exhibits less positive therapeutic effects compared to negative HMCM (Key Finding 4), this might inversely validate that negative AIGM evokes stronger negative emotional responses than HMCM. Therefore, our findings suggest that AIGM not only meets quality standards but also demonstrates more pronounced emotional attributes. Future music therapy practices could consider the appropriate incorporation of AIGM in suitable therapeutic contexts.

Limitation and Future work

While this study provides valuable insights into the neural and emotional effects of AIGM and HMCM in therapy, several limitations should be acknowledged. First, our findings may be limited to the selected human piano pieces and specific AI models used. Future studies should expand the music sample scope to verify the generalizability of these results. Second, as participants were East Asian, cultural influences may affect our findings. Research with diverse cultural backgrounds is needed to examine potential cultural effects. Third, our conclusions are constrained by the relatively small sample size and limited diversity of participants; future research should recruit larger cohorts and include participants from different cultures, age groups, and other demographic backgrounds to enhance external validity. Furthermore, the present study examined immediate therapeutic effects only following negative emotion induction. Future investigations should extend to examining emotional responses across diverse emotional baselines and validating longitudinal therapeutic efficacy.

Conclusion

This study explores the effects of AIGM and HMCM in emotional therapy. The results show that AIGM demonstrates a more significant effect on increasing arousal levels. In the context of positive emotional music, AIGM is more effective in promoting positive emotional therapeutic effects compared to HMCM, while HMCM shows better emotional regulation effects in the context of negative emotional music. Moreover, when music adheres to the ISO principle, both AIGM and HMCM exhibit significant improvements in therapeutic outcomes. In sum, for the first time, this work reveals the neural activities induced by AIGM through EEG analysis and validates its unique effects under the ISO principle. It lays a solid foundation for future in-depth research in this area.

References

Adolphs, R., Tranel, D., & Damasio, A. R. (2002). The human amygdala in social judgment. *Psychological bulletin*, *372*(6507), 669–672.

Ahmad, N., & Rana, A. (2015). Impact of music on mood: Empirical investigation. *Research on Humanities and Social Sciences. ISSN (Paper)*, 2224–5766.

Altenmüller, E., Schürmann, K., Lim, V. K., & Parlitz, D. (2002). Hits to the left, flops to the right: Different emotions during listening to music are reflected in cortical lateralisation patterns. *Neuropsychologia*, *40*(13), 2242–2256.

Bogatyrenko, O. (2024). Meditative music as a method for stress reduction, relaxation, and focus enhancement.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, *25*(1), 49–59.

Davis, W. B. (2008). An introduction to music therapy: Theory and practice. *American Music Therapy Association*.

Deng, Y., Yang, M., & Zhou, R. (2017). A new standardized emotional film database for asian culture. *Frontiers in psychology*, *8*, 1941.

Dixon, M. L., Thiruchselvam, R., Todd, R., & Christoff, K. (2017). Emotion and the prefrontal cortex: An integrative review. *Psychological bulletin*, *143*(10), 1033.

Ge, Y., Zhao, G., Zhang, Y., Houston, R. J., & Song, J. (2019). A standardised database of chinese emotional film clips. *Cognition and Emotion*, *33*(5), 976–990.

Gross, J. J. (2002). Emotion regulation: Affective, cognitive, and social consequences. *Psychophysiology*, *39*(3), 281–291.

Heiderscheit, A., & Madson, A. (2015). Use of the iso principle as a central method in mood management: A music psychotherapy clinical case study. *Music therapy perspectives*, *33*(1), 45–52.

Homan, R. W., Herman, J., & Purdy, P. (1987). Cerebral location of international 10–20 system electrode placement. *Electroencephalography and clinical neurophysiology*, *66*(4), 376–382.

Hung, H.-T., Ching, J., Doh, S., Kim, N., Nam, J., & Yang, Y.-H. (2021). Emopia: A multi-modal pop piano dataset for emotion recognition and emotion-based music generation. *arXiv preprint arXiv:2108.01374*.

Lin, Y.-P., Wang, C.-H., Jung, T.-P., Wu, T.-L., Jeng, S.-K., Duann, J.-R., & Chen, J.-H. (2010). Eeg-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, *57*(7), 1798–1806.

Malchiodi, C. (2012). *Art therapy and health care*. Guilford Press.

Ryczkowska, A. (2022). Positive mood induction through music: The significance of listener age and musical timbre. *Psychology of Music*, *50*(6), 1961–1975.

Saarikallio, S. H. (2008). Music in mood regulation: Initial scale development. *Musicae scientiae*, *12*(2), 291–309.

Starcke, K., Mayr, J., & von Georgi, R. (2021). Emotion modulation through music after sadness induction—the iso principle in a controlled experimental study. *International journal of environmental research and public health*, *18*(23), 12486.

Sun, J., Yang, J., Zhou, G., Jin, Y., & Gong, J. (2024). Understanding human-ai collaboration in music therapy through co-design with therapists. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–21.

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in cognitive sciences*, *3*(4), 151–162.

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: An online survey. *PLOS ONE*, *9*(3), e110490.

Thaut, M., & Hoemberg, V. (2014). *Handbook of neurologic music therapy*. Oxford University Press, USA.

Wun, S., Horner, A., & Wu, B. (2014). Effect of spectral centroid manipulation on discrimination and identification of instrument timbres. *Journal of the Audio Engineering Society*, *62*(9), 575–583.

Zhang, G., Yu, M., Chen, G., Han, Y., Zhang, D., Zhao, G., & Liu, Y.-J. (2019). A review of eeg features for emotion recognition. *Scientia sinica informationis*, *49*(9), 1097–1118.