Proceedings of the 7th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/2027.39050

Cultural Cognition and Creative Practice of Traditional Music Recontextualization in Human AI Co-Creation Environments

Binru Zou

Conservatory of Music, Shanghai University, Shanghai, China
wellington589125@gmail.com

Abstract. Conventional music is being increasingly embedded into generative artificial intelligence contexts, where fast melodic transformations would help recontextualize while losing modal integrity, motivic unity, and culturally grounded meanings. A symbolic human-AI collaborative music-making system was built based on an efficient Transformer model, culturally interpretative reranking of cultural features, and interactive recording. A within-subject counterbalanced experiment with seventy-two participants covered unaided creation, general AI generation, and culturally constrained AI generation. The assessment included a 15-question cultural-cognition test, five-expert cultural appropriateness ratings, symbolic feature analysis, MIDI edit distance, completion time, cognitive load, and perceived agency. Simulation results suggested that culturally constrained AI led to greater cognition gains, better modal and motivic preservation, more culturally appropriate rating, and shorter completion time while having greater creative agency compared to general AI. Cultural cognition gain was 2.31 ± 1.26 points, expert cultural appropriateness ratings were 5.62 ± 0.58, and completion time was 29.6 ± 4.8 min. These trends reflect how culturally constrained AI reranking can link computation of music generation with culturally informed cognition and keep meaningful human participation.

Keywords: traditional music, human AI co-creation, cultural cognition, music generation, recontextualization

1. Introduction

Traditional music recontextualization is taking place in a variety of formats such as games, short video clips, contemporary dance, and more, leading to increased need for tools that can transform melodic content while maintaining culturally relevant structures. Symbolic generation literature proves that Transformers can learn event sequences and long dependencies of music but have difficulties in linking formal coherence to human creative judgement [1]. Moreover, the research done regarding the intersection between technology and traditional music proves that innovation can stand in contrast with cultural values when musical practice gets stripped down to data [2]. The creative uncertainty plays a role in the workflow of musicians but the generative systems aim at

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2027.39050

controllability and efficiency [3]. This indicates that cultural knowledge needs to get included into generation process through an interpretable constraint included in the candidate selection process.

2. Literature review

2.1. Cognitive mechanisms of traditional music recontextualization

Traditional music's cultural cognition entails modal arrangement, structural pitches, rhythmic conventions, phrasing tendencies, performance practice, and social function. Chinese music information retrieval studies have brought about increased accessibility to curated datasets and computational procedures, thereby fostering reproducibility of culturally specific analysis [4]. Situated approaches further reveal the fact that the meaning of music arises out of embodied, situated, and technological mediation, apart from the acoustic features [5]. The recontextualization process thus presents a contradiction between continuity on the one hand and contextual change on the other. A melody can maintain its pentatonic scale distribution while lacking in culturally significant phrase ending, or veer away from the original context while maintaining its motivic logic.

2.2. Co-creative logic of generative music systems

Generative musical interfaces offer the potential for prompting, continuing, varying, and candidate selection processes and move creativity towards one of iterative comparison and modification. In value-centered design of music-AI interfaces, creative differentiation and continued engagement are important considerations [6]. Research from the domain of AI-supported composition also demonstrates that generated output can reveal new creative potentials while transforming notions of creative control. Increased independence and productivity along with fears of replacement and exploitation have been found in interviews with musicians using AI systems [7]. These results challenge the notion of efficiency as an end goal of co-creative interaction. Interactivity is contingent upon acceptance, rejection, modification, and understanding by users. Thus, recording these metrics can inform us of whether the AI expands creative potential or simply expedites established stylistic production.

2.3. Multidimensional evaluation of cultural appropriateness

Cultural appropriateness is a multi-dimensional assessment since a high degree of similarity between the melody under study and the source melody could suggest imitative behavior. Musical-creativity studies have established that there are tasks in which generation and evaluation could be assessed in structured sequence form in order to distinguish between the two processes [8]. Computational measures can be used to evaluate such factors as modal consistency, retention of central motif, distribution of rhythm, and transformation distance; and expert judgments can help in assessing such dimensions as contextual appropriateness, instrumental significance, and creative evaluation. The two approaches might differ in interesting ways since a candidate who demonstrates significant feature similarity but is awarded a poor expert judgment rating because the key gestures appear in an inappropriate context could be transformed but still preserve structural identity [9].

DOI: 10.54254/2755-2721/2027.39050

3. Experimental methods

3.1. Symbolic music co-creation platform

The system was developed using Python 3.11, PyTorch 2.1 for model training, music21 for symbolic processing and feature extraction, and Gradio for interactive task presentation. Melodies used symbolically were extracted from the Chinese part of the Essen Folksong Collection through parsing, duplicate elimination, length filtering, and event normalization. Original melodies were divided into training, validation, and test sets and augmented so that transposed instances would not end up in different datasets. 12-key transposition increased melodic variety while keeping pitch relationships intact. A four- to six-layer Transformer predicted eight-bar continuation using REMI-style event tokens corresponding to bar number, beat, pitch, and duration, in line with modern symbolic music generation literature [10]. Given an eight-bar melody as the standardized input, the model generated 12 continuations by sampling with predefined decoding parameters. Then 4 continuations were presented via the user interface. The platform maintained information about source MIDI files, generated candidates, listening events, candidate selection, generation again, edits, and exported compositions. This information could be associated with participants and tasks to enable straightforward integration with cognitive tests, expert ratings, and feature analysis.

3.2. Cultural feature-based candidate reranking

All AI conditions use the same Transformer and identical twelve-candidate pools, while candidate presentation differs only in ranking strategy. Modal similarity evaluates agreement between source and generated melodies in pentatonic pitch-class distribution, tonal center, and phrase-ending tone, which are computationally measurable characteristics of Chinese traditional melodic organization [11]. The modal score is calculated as shown in Equation (1).

$$S_{mode}=\alpha\cos(p_s,p_c)+\beta I(t_s=t_c)+\gamma I(e_s=e_c) \tag{1}$$

Where p_s and p_c denote normalized pitch-class vectors, t represents tonal center, e represents phrase-ending tone, $I(\cdot)$ is an indicator function, and $\alpha+\beta+\gamma=1$. Motif similarity measures weighted overlap among two-to-five-note interval n-grams, while rhythmic similarity is derived from one minus the Jensen-Shannon divergence between duration-pattern distributions. Pilot testing determines the corresponding feature weights. The final cultural score is calculated as shown in Equation (2).

$$S_{cult}=w_mS_{mode}+w_oS_{motif}+w_rS_{rhythm}$$

$$w_m+w_o+w_r=1 \tag{2}$$

General AI presents the four candidates with the highest generation probabilities, whereas culturally constrained AI presents the four candidates with the highest S_{cult}, thereby isolating the influence of cultural reranking.

DOI: 10.54254/2755-2721/2027.39050

3.3. Counterbalanced creative experiment

A total of 72 subjects having a minimum of two years of experience in instrumental performance, composition, and/or digital music are used for the within-subject Latin-square experiment, involving unaided composition, general AI, and culturally constrained AI. The three tasks of composing music for games, short videos, and contemporary dance last for 40 minutes each, based on standardized eight-bar folk melodies to create 60 seconds of music adaptation. Before and after the tasks, a cultural cognition test, which consists of 15 items, will be given. There are five experts in traditional music who will assess cultural suitability and creativity on a seven-point scale. MIDI editing intensity is quantified as shown in Equation (3).

$$D_{MIDI}=\frac{N_{ins}+N_{del}+N_{sub}}{\max(N_c,N_f)} \tag{3}$$

Where N_{ins}, N_{del}, and N_{sub} represent inserted, deleted, and substituted MIDI events, while N_c and N_f denote candidate and final-event counts. Condition effects are estimated using the linear mixed-effects model shown in Equation (4).

$$Y_{ijkm}=\beta_0+\beta_1 G_i+\beta_2 C_i+\beta_3 O_j+u_k+v_m+\varepsilon_{ijkm} \tag{4}$$

Here, G_i and C_i represent the two AI conditions, O_j denotes task order, u_k+ and v_m are participant and melody random intercepts, and Y_{ijkm} represents each cognitive or creative outcome.

4. Results

4.1. Cultural cognition performance

Simulation results for the proposed 72-subjects data set are presented in Table 1. Culture-cognition gain went up from 0.42 ± 1.08 points in case of unaided creation and 0.86 ± 1.15 points in general AI to 2.31 ± 1.26 points in case of culturally constrained AI. In a mixed model, the advantage was estimated at 1.89 points, 95% CI [1.55, 2.23], $p < .001$. The expert ratings of cultural-appropriateness were at 5.62 ± 0.58, while those of unaided creation and general AI were at 4.61 ± 0.66 and 4.78 ± 0.63, respectively. Modal stability rose to 0.86 ± 0.07, motif preservation came to 0.79 ± 0.08, and the rhythmic distribution distance dropped to 0.122 ± 0.045. The expert creativity scores stayed at 5.35 ± 0.62. As seen from Table 1, cognition, expert evaluation, and structural similarity all grew simultaneously. Such a convergence fits well with culturally guided reranking of music affecting both selection and cultural evaluation.

Table 1. Cultural cognition and musical-structure outcomes across experimental conditions

Measure	Unaided Creation	General AI	Culturally Constrained AI	Overall p
Cultural-cognition gain, points	0.42 ± 1.08	0.86 ± 1.15	2.31 ± 1.26	<.001
Cultural appropriateness, 1–7	4.61 ± 0.66	4.78 ± 0.63	5.62 ± 0.58	<.001
Modal stability, 0–1	0.74 ± 0.09	0.71 ± 0.10	0.86 ± 0.07	<.001
Motif preservation, 0–1	0.63 ± 0.11	0.58 ± 0.12	0.79 ± 0.08	<.001

DOI: 10.54254/2755-2721/2027.39050

Table 1. (continued)

Rhythmic JS distance	0.198 ± 0.061	0.214 ± 0.067	0.122 ± 0.045	<.001
Expert creativity, 1–7	4.82 ± 0.69	5.21 ± 0.65	5.35 ± 0.62	.004

4.2. Creative practice performance

The findings related to creative practice distinguished between the two types of AI assistance even more. The average duration of the generation process was cut from 37.8 ± 5.6 min for unaided creative practice to 31.4 ± 5.1 min for general AI and to 29.6 ± 4.8 min for culturally constrained AI. The acceptance rate of generated candidates was 63.1% ± 14.2% when general AI was used. However, using cultural reranking, this rate was lowered to 56.7% ± 13.8% but the normalized MIDI editing distance was increased to 0.41 ± 0.13 from 0.34 ± 0.12 (see Figure 1). Thus, using culturally constrained AI means moving towards selecting less candidates and performing more symbolic editing. Moreover, the level of agency was increased to 5.71 ± 0.67 from 5.08 ± 0.74 and the level of expertise creativity was increased to 5.35 ± 0.62 from 5.21 ± 0.65. The cognitive load was lowered to 3.21 ± 0.76 from 3.68 ± 0.81. The frequency of generation was 3.8 ± 1.5 in the first case and 3.3 ± 1.4 in the second one.

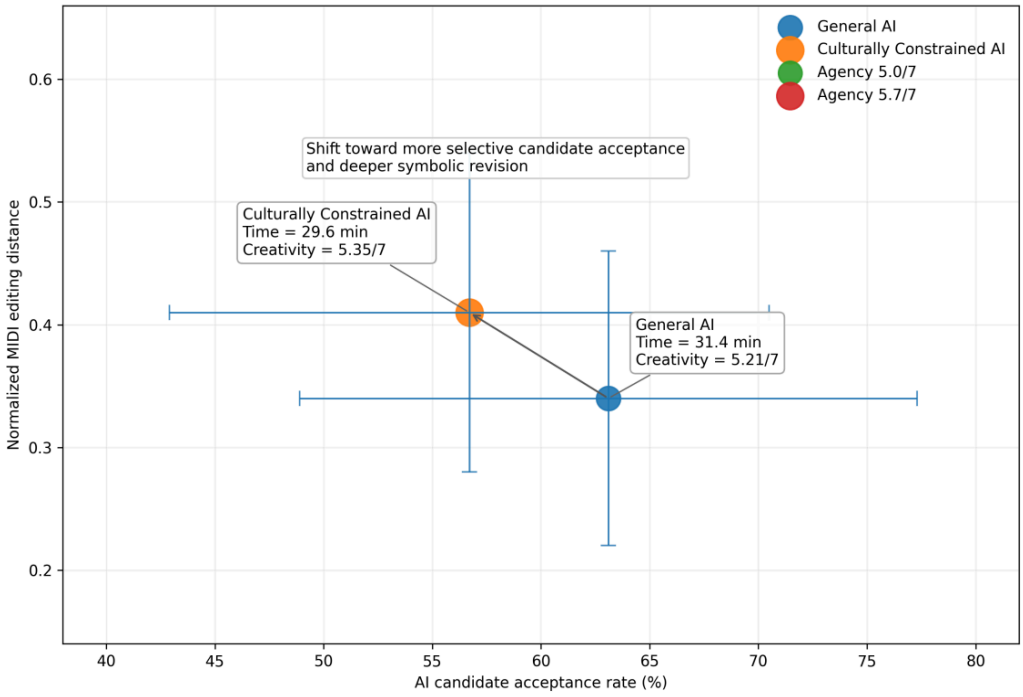


Figure 1. Human AI musical co-creation behavior under alternative candidate ranking

5. Discussion

The pattern produced by simulation indicates that culture-imposed constraints can alter interaction quality even when the generative process does not change at all. Selective visibility plays a key role in this process since the re-ranking operation changes which music options will be visible to the creator before being chosen actively. Higher values of modal and motif retention and increased distance between MIDI edits are especially important here because cultural continuity and human creativity proceed side-by-side. The lower candidate acceptance probability in case of culture-

DOI: 10.54254/2755-2721/2027.39050

imposed reranking puts into question the validity of the hypothesis that a good recommender must strive to increase immediate acceptance probability. Selective rejection and editing can rather signal the presence of stronger authorship and deliberate assessment of musical content. Fixed weights of features used to generate recommendations will continue to privilege those aspects that are easier to formalize and disregard ornamentation, regional specifics, timing of performance, and social functions. Thus, expert assessments and cultural cognition advantages are still needed for assessing the results of computational scoring. Cultural reranking needs to be viewed more as a scaffold for human decisions than a method itself.

6. Conclusion

A unifying approach has been formulated to study the recontextualization of traditional music through symbolic generation, cultural features re-ranking, and human-AI co-creation under control. The compact Transformer generates candidate completions, but general AI and culturally constrained AI differ in candidate ranking only, hence allowing for isolation of the effect caused by cultural constraints. The Latin-square experiment involved 72 participants, and consisted of cultural cognition tests, expert judgments, log of human-computer interactions, MIDI editing distance, completion time, cognitive load, and perceived agency. The unifying approach studies the possibility of preserving cultural structure with active human revision and creation at the same time. Numerical results obtained within the simulation so far are just simulation-driven place-holders. The chosen corpus is not representative of Chinese traditional music in its entirety.

References

[1] Ji, S., Yang, X., & Luo, J. (2023). A survey on deep learning for symbolic music generation: Representations, algorithms, evaluations, and challenges. *ACM Computing Surveys, 56*(1), 1–39.
[2] Kanhov, E., Kaila, A.-K., & Sturm, B. L. T. (2024). Innovation, data colonialism and ethics: Critical reflections on the impacts of AI on Irish traditional music. *Journal of New Music Research, 53*(1–2), 47–63.
[3] Loor Paredes, M. (2025). Emerging paradigms in music technology: Valuing mistakes, glitches and uncertainty in the age of generative AI and automation. *AI & Society, 40*(6), 4701–4712.
[4] Bown, O. (2024). From genies performing magic to sages imparting wisdom: A value-centred survey of music AI user interfaces, creative affordances and artist objectives. *Journal of New Music Research, 53*(1–2), 5–18.
[5] Weaver, A. (2026). Sisters, not twins: Exploring artistic control and anthropomorphism through composing with a bespoke generative AI. *AI & Society, 41*(2), 1609–1621.
[6] Zhou, M., et al. (2025). CCMusic: An open and diverse database for Chinese music information retrieval research. *Transactions of the International Society for Music Information Retrieval, 8*(1), 22–38.
[7] de Aguiar, V. (2024). Music and affectivity in the age of artificial intelligence. *Topoi, 43*(3), 725–735.
[8] Loui, P., et al. (2026). Sequence generation and evaluation: A novel assessment of musical creativity. *Creativity Research Journal, 38*(1), 210–224.
[9] Herington, J., et al. (2026). Musicians' ethical concerns about AI: An interview study. *AI & Society, 41*(2), 1075–1088.
[10] Mycka, J., & Mańdziuk, J. (2025). Artificial intelligence in music: Recent trends and challenges. *Neural Computing and Applications, 37*(2), 801–839.
[11] Xie, L., Wang, Y., & Gao, Y. (2024). Acoustical feature analysis and optimization for aesthetic recognition of Chinese traditional music. *EURASIP Journal on Audio, Speech, and Music Processing, 2024*(1), 7.