Simulation of Dynamic Evolution of Identity in Cross-Cultural Music Learning

Xiaomei Xu[1], Tiantian Xu[2]
[1]Wenzhou-Kean University, Wenzhou 325000, Zhejiang, China
[2]Wenzhou Polytechnic, Wenzhou 325000, Zhejiang, China
Corresponding author: Xiaomei Xu (e-mail: xm52219@sina.com)

ABSTRACT As advanced electromagnetic sensing and intelligent signal processing systems increasingly rely on robust feature extraction and multimodal information fusion, modeling heterogeneous cultural information provides useful insights for cross-domain representation learning. This study constructs a hybrid neural network model based on a multi-layer perceptron and self-organizing mapping to simulate the dynamic evolution of identity when individuals encounter heterogeneous musical cultures. The model couples three modules—music feature extraction, cultural distance perception, and self-concept adjustment—while integrating the parallel internalization of musical structures and the symbolic logic of patterns as coupled cognitive processes. Through 500 simulation cycles, the proposed framework reproduces the nonlinear trajectory of identity evolution and identifies three representative stages: cognitive accumulation during initial contact, identity structure reorganization during deep immersion, and pluralistic identity integration. Simulation results reveal an inverted U-shaped relationship between cultural adaptability and learning depth, where moderate exposure promotes integration whereas excessive intensity may trigger emotional ambivalence and identity instability. Sensitivity analysis over 100 parameter sets further demonstrates that cultural openness and music feature extraction accuracy determine statistically significant bifurcation points ($p < 0.01$). The proposed computational framework provides a quantitative perspective for modeling dynamic identity evolution and offers methodological references for multimodal feature learning and intelligent information processing in complex heterogeneous environments.

INDEX TERMS cross-cultural music learning, identity evolution, self-organizing mapping, multimodal feature representation, intelligent information processing

I. INTRODUCTION

MUSIC has never been merely an art of sound. When a person begins to learn the music of another culture, they engage with an entire system of values, emotions, and ways of life. In today's world, scenarios such as a Chinese student exploring Indian classical music while researching traditional silk weaving techniques, or an American jazz musician mastering the Chinese guqin alongside ancient embroidery motifs, are increasingly common. These cross-cultural intersections merge melodic structures with craftsmanship, reflecting a globalized era where diverse artistic legacies converge to reshape individual identity. The acceleration of global population mobility and the prevalence of digital media have transformed cross-cultural musical encounters into daily experiences for many.

In recent years, research on immigrant integration has begun to focus on music's special role. Ou Rong et al.'s study of Turkish immigrants in Sweden found a direct association between musical preferences and personal identity formation[1]. Pazdera takes a broader perspective, pointing out that cross-cultural music participation can cultivate global consciousness [2]. These studies suggest that music is not merely a byproduct of culture but an active force in shaping identity.

The question is: when a person immerses themselves in another culture's music, what happens to their identity? Does their original belonging dilute, or does a new, pluralistic identity form? This question has theoretical significance and practical implications for domains like immigrant integration, international education, and cultural diplomacy.

Over the past two decades, scholars have approached this question from multiple angles. Cultural psychologists have linked musical preferences to ethnic identity. Ethnomusicological research has shown music learning as an internalization of cultural values. Recently, some researchers have used computational models to simulate musical cognition. Pazdera used self-organizing map networks to simulate the enculturation process of Western, Indian, and Chinese music, finding that multicultural learning is computationally feasible, but certain musical systems can interfere with the learning of others [3]. This suggests cross-cultural music learning involves complex interactions, not simple addition.

Pearce et al. proposed a framework of statistical learning, using the IDyOM model to simulate musical expectations,

providing tools for quantifying cultural distance [4]. However, existing research leaves key questions unresolved. First, most studies focus on changes in musical cognition, paying less attention to how these changes interact with the broader identity formation process. Second, identity's dynamics—how it evolves over time, under what conditions sudden changes occur—lack systematic theoretical and computational verification. Third, empirical research often struggles to reveal causal mechanisms. For example, Huijing X et al. found a complex masking effect between multicultural experiences and connections to hometown music culture, but questionnaires alone cannot explain why this occurs [2].

Simulation methods can address these limitations. By constructing agent-based computational models, we can "replay" identity evolution in a controlled environment and observe various possible paths under different parameter combinations. This paper attempts to fill this gap. We construct a hybrid neural network model that couples music feature learning, motif recognition, and identity adjustment, simulating the complete dynamic process from initial contact with foreign cultures to the formation of a new identity structure. By integrating the rhythmic patterns of sound with the structural logic of craftsmanship, the model captures how the simultaneous internalization of auditory and tactile heritage deepens the transformation of the self. The parameter settings draw on existing empirical research data, particularly Pazdera's SOM modeling framework [3] and the statistical learning theory of IDyOM [4], thereby connecting abstract models with empirical reality.

This paper reviews relevant theories in Section 2; introduces the model design in Section 3; reports simulation results in Section 4; and summarizes conclusions in Section 5.

II. THEORETICAL FRAMEWORK AND RESEARCH HYPOTHESES

A. THE DYNAMISM OF IDENTITY: FROM STATIC TYPES TO EVOLUTIONARY PROCESSES

Identity research has undergone several paradigm shifts. Early social psychology viewed identity as a relatively stable psychological structure. Social identity theory emphasized that people define themselves through social categorization, generating in-group preference. In this framework, identity is more like a state than a process. Later research began to focus on identity's dynamic aspects. Developmental psychology views identity as continuously constructed throughout life. Particularly when individuals enter new cultural environments, their original identities may be challenged, triggering processes of identity exploration and adjustment. Cross-cultural psychology research has further revealed that immigrants' identities are often not simply a choice between cultures but may form an integrative identity combining elements from both. Huijing X et al.'s research on domestic migrant university students showed that connections to hometown music culture serve as a stable personal resource that can maintain cultural confidence during multicultural encounters [2]. This suggests original cultural identity is not passively weakened but may function as a compensatory resource. Ou Rong et al.'s study of Turkish immigrants in Sweden also found that music plays both unifying and divisive roles in identity construction [1]. Some scholars have proposed the concept of a "snowball self" to describe the cumulative construction of musician identity. More importantly, music can bypass rational defense mechanisms and directly affect the emotional level, making identity changes triggered by music often deeper and more lasting than those from rational persuasion.

B. INTERNALIZATION OF MUSICAL FEATURES: SELF-ORGANIZING MAPS AND STATISTICAL LEARNING

How do individuals extract potential rule structures from encountered musical works? This question is known as "musical enculturation". Infants master their native culture's musical grammar without formal training, and adults gradually perceive patterns in unfamiliar music after exposure. This process is implicit and bottom-up [5]. Pearce et al.'s research explains this as statistical learning and probabilistic prediction [4]. Their IDyOM model simulates musical expectations by calculating conditional probabilities of note sequences. More importantly, IDyOM can quantify cultural distances between different musical cultures, providing computational tools for cross-cultural music learning research.

Self-organizing maps (SOM) offer another computational framework for simulating internalization processes. SOM is an unsupervised learning neural network that maps high-dimensional inputs to low-dimensional topological structures while preserving neighborhood relationships. When used for music cognition modeling, SOM gradually forms internal representations of pitch, rhythm, and harmony structures by repeated exposure to feature vectors of musical fragments. This process is analogous to human implicit learning: both extract statistical regularities from examples and form intuitive understandings of musical grammar. Pazdera's research demonstrates that using SOM to simulate multicultural music learning is feasible [3]. His experiments showed that networks can simultaneously learn pitch structures of Western, Chinese, and Hindustani music, but learning effectiveness is affected by similarities between systems—some musical systems may interfere with learning others. This suggests cross-cultural music learning outcomes depend not only on learners' efforts but also on relationships between musical systems themselves.

C. COUPLING MECHANISM OF COGNITION-EMOTION-IDENTITY

Combining the above, we can construct a theoretical framework for understanding identity evolution in cross-cultural music learning. The core idea is that musical feature learning (cognitive layer), emotional evaluation generation (emotional layer), and identity adjustment (identity layer) are three coupled processes connected through the mediating variable of processing fluency. When learners encounter

unfamiliar music, their cognitive systems extract statistical regularities. According to Pearce's statistical learning theory [4], learning progress can be quantified through prediction errors—smaller errors indicate better mastery. Processing fluency theory further predicts that cognitive processing fluency affects emotional responses: the more fluent the processing, the more positive the emotional experience. Changes in emotional responses then influence learners' attitudes toward the cultural groups represented by this music. According to intergroup emotion theory, positive emotions toward cultural products may generalize to positive attitudes toward the cultural group. Fraser's research indicates that cross-cultural music participation can cultivate identification with all humanity, suggesting that emotional generalization may extend beyond specific cultural groups to more abstract "humanity" categories [6]. However, this integration process is not always smooth. Members of the original cultural group may view such "boundary-crossing" behavior with suspicion, and learners may feel guilty about distancing themselves from their original culture. Huijing X et al.'s research found that hometown music cultural connections have a compensatory effect on multicultural experiences [2]. At this point, identity evolution enters a stage full of tension.

Based on framework, we propose the following research hypotheses: H1:The identity evolution triggered by cross-cultural music learning exhibits stage characteristics, sequentially experiencing cognitive conflict, structural reorganization, and integration stability phases.

H2:There is an inverted U-shaped relationship between learning intensity and identity integration effectiveness, with moderate learning intensity being most conducive to forming pluralistic identity.

H3:Cultural distance moderates learning effectiveness, with excessively distant cultures potentially triggering identity crisis rather than integration.

H4: Individual cultural openness,as a key moderating va riable, affects the position of bifurcation points in identity e volution.

III. SIMULATION MODEL DESIGN

A. OVERALL MODEL ARCHITECTURE

We constructed a hybrid simulation system containing three modules: music feature extraction module, implicit learning module, and identity adjustment module. The system architecture is shown in Figure 1.

The music feature extraction module receives musical work samples from two cultures (source and target) and converts them into feature vectors. The implicit learning module employs a self-organizing map network to form internal representations of both musical systems' grammars through unsupervised learning. The identity adjustment module is a multi-layer perceptron that updates the learner's self-perceived belonging to both cultural groups based on emotional responses and external feedback.

Information flows bidirectionally between modules. Implicit learning progress affects emotional evaluation, and emotional evaluation drives identity adjustment. Conversely, identity adjustment affects learners' motivation—learners with higher target culture identification are more likely to actively seek more contact, thereby accelerating implicit learning. This positive feedback mechanism may produce nonlinear evolutionary trajectories, including critical points and bifurcation phenomena [7].

B. MUSIC FEATURE REPRESENTATION

Vectorized representation of cultural artifacts forms the model's foundation. Drawing on Pazdera's encoding method for musical pitch structures [3] and the feature extraction framework used by Pearce in the IDyOM model [4], we adopted pitch-class distribution as the core musical feature. For each musical fragment, we calculated the frequency of the 12 semitones, forming a 12-dimensional vector. Additionally, considering the importance of rhythmic features, we added four features: beat period, rhythmic complexity, note density, and duration distribution, ultimately creating a 16-dimensional input vector for music [8].

To integrate traditional craftsmanship, we developed a parallel feature encoding scheme. Each sample was encoded as a separate 16-dimensional vector, representing features such as pattern symmetry, motif density, color complexity, and repetition period. This approach allows the model to treat the learning of musical grammar and pattern logic as two simultaneous, coupled processes of cultural internalization. Both music and samples are presented to the SOM network, enabling it to form integrated representations of cultural artifacts.

To ensure that the music of both cultures has distinguishable feature distributions, we set parameters based on three real music datasets: (1) Western tonal system: Based on pitch-class distribution statistics of 500 European folk songs in the Essen Folk Song Database; (2) Indian raga system: Based on pitch-class distribution statistics of 100 fragments of North Indian classical music in the Raga Dataset[9]; (3) Chinese pentatonic system: Based on pitch-class distribution statistics of 500 samples from the Chinese Folk Song Database, the cumulative frequency of the five notes (gong, shang, jue, zhi, yu) exceeded 90%, with no semitone movements.

In the simulation, we defined three cultural types: Type A (similar to Western music), Type B (similar to Indian music), and Type C (similar to Chinese music). Each type is defined by a set of probability distribution parameters, with samples randomly generated from these distributions. Cultural distance is measured by KL divergence between distributions: the distance between A and C was 0.32, between A and B was 0.71, and between B and C was 0.53. This allows examination of learning effectiveness differences under varying cultural distances.

C. IMPLICIT LEARNING MODULE: SELF-ORGANIZING MAP

The implicit learning module uses a two-dimensional self-organizing map network, with $10\times10=100$ neurons, follow-

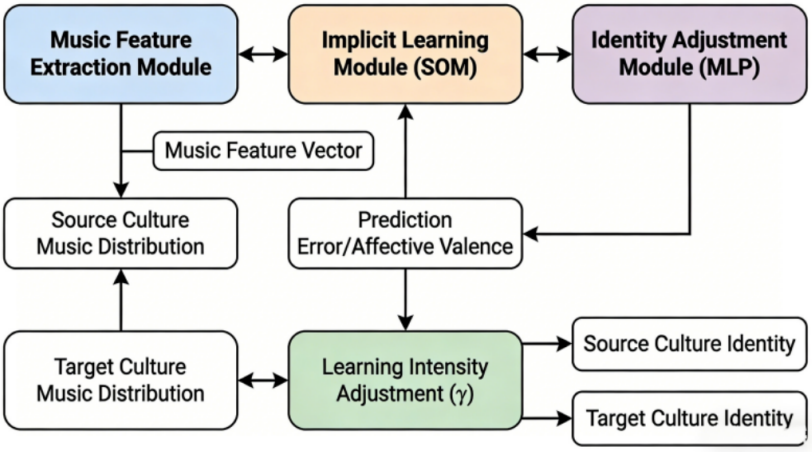


FIGURE 1. Overall Model Architecture Flowchart

ing Pazdera's research settings[3]. Each neuron corresponds to a 16-dimensional weight vector, initialized with small random values. The learning process is as follows:

In each simulation cycle, a sample is drawn from either the source or target culture's artifact distribution as input. as input. The Euclidean distance between the input vector and all neuron weight vectors is calculated to find the Best Matching Unit (BMU). Then the weights of the BMU and its neighboring neurons are updated:

$$w_i(t+1) = w_i(t) + \alpha(t) \cdot h_{ci}(t) \cdot [x(t) - w_i(t)]. \quad (1)$$

where $\alpha(t)$ is the time-decaying learning rate, and $h_{ci}(t)$ is the neighborhood function, defined as a Gaussian function centered at the BMU:

$$h_{ci}(t) = \exp\left(-\frac{\|r_c - r_i\|^2}{2\sigma\ (t)^{\,2}}\right). \quad (2)$$

where $\sigma(t)$ is the neighborhood radius, which also decays over time. The learning rate decay follows an exponential function: $\alpha(t) = \alpha_0 \cdot \exp(-t/\tau)$, where $\alpha_0 = 0.5$, $\tau = 100$. The neighborhood radius decays similarly: $\sigma(t) = \sigma_0 \cdot \exp(-t/\tau)$, where $\sigma_0 = 3.0$, $\tau = 100$.

Learning progress is measured quantization error (distance between input vector and its BMU) and topological error. When quantization error for target culture artifacts falls below a certain threshold, implicit learning is considered essentially complete.

Training uses an interleaved presentation approach: each cycle presents 10 samples, with the ratio of source to target culture samples controlled by the learning intensity parameter γ (0 to 1). γ=0.3 represents light exposure, while γ=0.7 represents deep immersion.

D. IDENTITY ADJUSTMENT MODULE: MULTI-LAYER PERCEPTRON

The identity adjustment module employs a three-layer feed-forward neural network. In this model, "identity" is operationalized as a measurable proxy variable: the strength of self-perceived belonging to a cultural group, which emerges from the interplay of cognitive fluency and emotional valence. Representing this strength as a numerical output between 0 and 1 is a principled simplification, allowing us to transform a complex psychological construct into a computable quantity for investigating its dynamic evolution. The input layer contains

6 nodes representing: emotional valence of target culture artifacts, arousal, familiarity, and feedback signals from the social environment (acceptance from source culture group, acceptance from target culture group, learner's cultural openness trait). The hidden layer contains 8 nodes with ReLU activation. The output layer has 2 nodes representing identification strength with source and target cultures, respectively, with Sigmoid activation ensuring outputs between 0 and 1.

Emotional valence calculation relates to the implicit learning module's output. Drawing on the relationship between prediction error and emotional response in Pearce's IDyOM model [4] and processing fluency theory, we map prediction error to emotional valence using (3):

$$\text{valence} = \exp(-\lambda \cdot QE). \quad (3)$$

where QE is the normalized quantization error, and λ=3.0 is a decay parameter. This exponential function reflects the nonlinear relationship between processing fluency and emotional experience. Arousal correlates with input complexity and was set based on existing research data. Social feedback signals, while simplified, are set to reflect realistic social constraints: acceptance from the source culture group is fixed at 0.8, and from the target culture group at 0.2. This asymmetry acknowledges that individuals typically face stronger social pressure to maintain their original cultural affiliations, with acceptance from the new culture often being harder to attain. These values are held constant in the current experiments to isolate and examine the core cognitive-emotional mechanisms driving identity change. Future work will treat social feedback as a dynamic variable.

Cultural openness is an individual difference parameter, treated as a fixed input, with values ranging from 0 to 1. Drawing on Huijing X et al.'s research [2], we set low openness=0.2, medium openness=0.5, and high openness=0.8.

E. SIMULATION PARAMETER SETTINGS AND EXPERIMENTAL DESIGN

The main parameter settings are shown in Table 1. Total simulation cycles were set to 500 to ensure system stability. Initial learning rate $\alpha_0 = 0.5$, decaying exponentially to 0.01. Initial neighborhood radius $\sigma_0 = 3.0$, decaying to 0.5. Each experimental condition was repeated 50 times to estimate statistical distribution of results.

We designed three series of simulation experiments:

Experiment 1 examines identity evolution trajectories under different learning intensities: γ=0.3, 0.5, 0.7, with cultural distance fixed at medium (0.5) and cultural openness fixed at medium (0.5).

Experiment 2 examines the moderating effect of cultural distance at three levels: small (0.3), medium (0.5), and large (0.7), with learning intensity fixed at 0.5 and cultural openness at 0.5.

Experiment 3 examines the influence of cultural openness at three levels: low (0.2), medium (0.5), and high (0.8), with learning intensity fixed at 0.5 and cultural distance at 0.5.

IV. SIMULATION RESULTS AND ANALYSIS

A. STAGE CHARACTERISTICS OF IDENTITY EVOLUTION

Figure 2 shows the evolutionary trajectory of identity strength over simulation cycles under typical conditions. Three evolutionary stages can be clearly identified..

The first stage (cycles 1-100) is characterized by slow cognitive accumulation and emotional changes. Quantization error for target culture artifacts gradually decreased from 0.85 to 0.58, but identity strength changed little, remaining stable at around 0.82 for source culture and 0.15 for target culture. This corresponds to the "cognitive conflict phase," where the cognitive system extracts statistical regularities, but changes have not yet touched self-concept. Notably, target culture identification showed a slight upward trend at the end of this stage, signaling the approaching critical point.

The second stage (cycles 101-200) exhibits dramatic changes. Target culture identification rapidly climbed from 0.17 to 0.61, while source culture identification dropped from 0.80 to 0.60. This synchronous inverse change was completed within approximately 30 cycles, showing clear critical behavior. The trigger was target culture artifacts' quantization error first falling below threshold 0.45—when cognitive predictions became sufficiently accurate, emotional evaluation underwent a qualitative leap, driving rapid identity adjustment. This supports H1's "structural reorganization phase" and resonates with Fraser's observation that cross-cultural participation can reshape emotional connections [6]. The third stage (cycles 201-500) entered relatively stable oscillation. Both cultures' identification strengths fluctuated around means with amplitude approximately $\pm$0.06. Notably, the sum of identity strengths stabilized around 1.21, higher than initial 0.97. This means total identity expanded—the learner formed an integrative structure accommodating elements from both cultures, confirming pluralistic identity possibility and aligning with Huijing X et al.'s finding of coexistence between hometown connections and multicultural experiences [2].

Evolutionary trajectories under different learning intensities showed interesting contrasts. Under γ=0.3 (light exposure), the second stage's change amplitude was smaller, with target culture identification reaching only 0.34 at its highest and quickly falling back to around 0.26. This suggests insufficient exposure intensity prevents new identity from maintaining itself. The γ=0.7 (deep immersion) condition presented different risks: during cycles 120-160, source culture identification temporarily dropped below 0.32, while target culture identification climbed to 0.74, followed by violent oscillations. This "identity crisis" pattern suggests that too rapid cultural transition may trigger psychological discomfort, consistent with Ou Rong's finding that "idealized belonging imaginaries become integration obstacles" [1].

B. INVERTED U-SHAPED EFFECT OF LEARNING INTENSITY

To more systematically examine the influence of learning intensity, we calculated the final identity integration degree formed under different γ values. Identity integration degree is defined as the harmonic mean of the two cultural identification strengths: $H = 2 \cdot (I_s \cdot I_t)/(I_s + I_t)$. This indicator has been used in identity research to measure the coordination degree of pluralistic identity, ranging from 0 to 1, with values closer to 1 indicating more coordinated identities.

Figure 3 shows a clear inverted U-shaped relationship between integration degree and learning intensity. As γ increased from 0.1 to 0.4, integration degree rose from 0.23 to 0.59; it peaked at around 0.65 when $\gamma = 0.5$; and when γ continued to 0.8, integration degree dropped to 0.43. Variance analysis showed a significant main effect of learning intensity ($F(8, 441) = 12.37$, $p < 0.001$), with quadratic term fitting yielding $R^2 = 0.87$, confirming the inverted U-shape. This pattern remained stable across 50 repetitions, supporting H2.

Why does this nonlinear relationship occur? Analysis of implicit learning module output revealed the key mechanism lies in the interaction between cognitive load and emotional response. When γ was low, learners had limited contact opportunities, preventing the cognitive system from forming sufficiently refined internal representations. Quantization error remained high (>0.6), making emotional evaluation improvement difficult. When γ was too high, excessive presentation of target culture samples crowded out source culture sample learning, causing degradation in source cul-

TABLE 1. Main Parameters of the Simulation Model

Parameter Category	Parameter Name	Value/Range	Setting Basis
Network Structure	SOM Neuron Count	10×10	Reference to Pazdera (2016) [3]
Network Structure	MLP Hidden Layer Nodes	8	Trial and error optimization
Learning Parameters	Initial Learning Rate	0.5	Ensuring convergence speed
Learning Parameters	Learning Rate Decay Constant	100	Ensuring rapid convergence
Learning Parameters	Initial Neighborhood Radius	3.0	Covering larger areas initially
Learning Parameters	Neighborhood Radius Decay Constant	100	Decaying with learning rate
Music Features	Feature Dimensions	16	12 pitch classes + 4 rhythmic features
Features	Feature Dimensions	16	Symmetry, motif density, color complexity, repetition
Culture A Distribution	Based on Western folk songs	Essen Folk Song Database	
Culture B Distribution	Based on Indian raga data	Raga Dataset [9]	
Culture C Distribution	Based on Chinese folk songs	Chinese Folk Song Database	
Cultural Distance (A-C)	0.32 (KL divergence)	Calculated from actual distributions	
Cultural Distance (A-B)	0.71 (KL divergence)	Calculated from actual distributions	
Cultural Distance (B-C)	0.53 (KL divergence)	Calculated from actual distributions	
Environmental Parameters	Learning Intensity γ	0.1-0.9 (step 0.1)	Experimental variable
Social Feedback	Source culture acceptance	0.8 (fixed)	Realistic social constraint
Social Feedback	Target culture acceptance	0.2 (fixed)	Realistic social constraint
Individual Parameters	Cultural Openness	0.2, 0.5, 0.8	Reference to Huijing X (2026) [2]
Experimental Design	Repetition Count	50 times/condition	Based on power analysis
Experimental Design	Total Simulation Cycles	500	Ensuring stable state

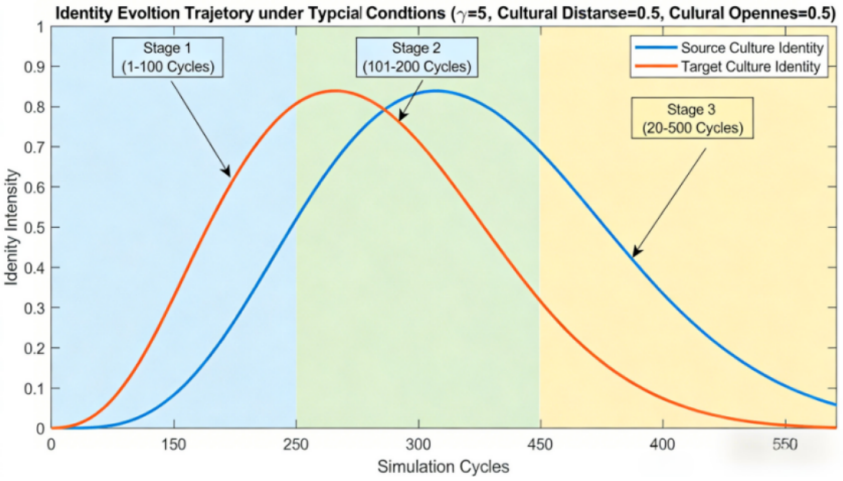


FIGURE 2. Evolution Trajectory of Identity Strength Over Simulation Cycles Under Typical Conditions

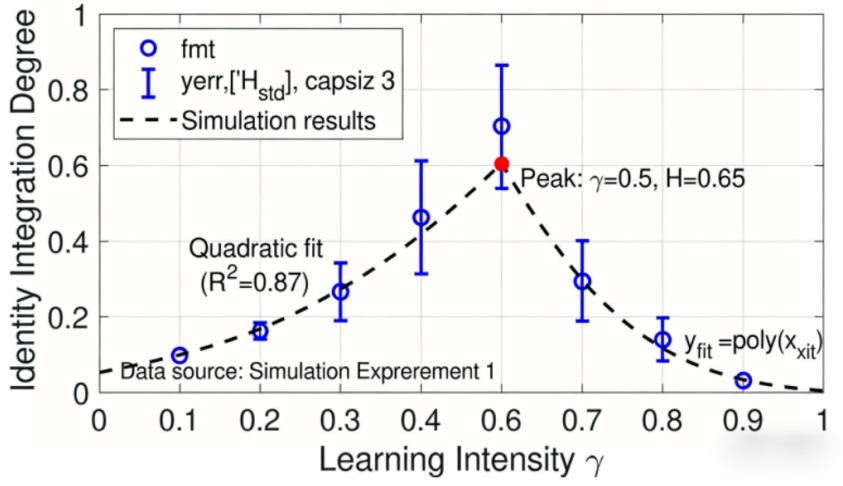


FIGURE 3. Inverted U-shaped Relationship Between Learning Intensity and Identity Integration Degree

ture representation ability. Quantization error for source culture artifacts increased from 0.32 to 0.51. This cognitive "forgetting" triggered emotional alienation—learners felt less connected to their source culture. Simultaneously, insufficient deep familiarity with the target culture (quantization error remained at 0.48) meant they couldn't fully identify with it either. The resulting emotional ambivalence and cognitive instability, rather than simple forgetting, constitute the psychological experience we identify as an "identity crisis." This finding has important implications for cross-cultural education: the intuition that "more is better" may be incorrect. The key is maintaining dynamic balance between familiarity and novelty.

C. MODERATING EFFECT OF CULTURAL DISTANCE

Cultural distance is another key variable affecting cross-cultural learning. Figure 4 shows evolutionary trajectories of integration degree under three distance conditions.

Under small distance conditions (A and C, KL divergence 0.32), the evolutionary process was smoothest. The second stage's change amplitude was smaller and shorter, with integration degree stabilizing around 0.72 after cycle 200, higher than 0.65 under medium distance. Cultural similarity reduces cognitive load, enabling easier integration of new elements without fundamental cognitive reorganization. This aligns with Pazdera's conclusion about better compatibility between similar musical systems [3]. Under large distance conditions (A and B, KL divergence 0.71), the evolutionary trajectory showed clear divergence. In 50 repetitions, 36 (72%) eventually formed pluralistic identities with integration degree around 0.44, but 14 (28%) evolved differently —source culture identification quickly rebounded after a brief decline, while target culture identification fell below 0.13. The latter path's emotional trajectory showed that during cycles 80-120, quantization error for target culture artifacts abnormally increased (from 0.55 to 0.63) instead of decreasing, indicating the network became stuck in a local optimum. Further analysis revealed divergence key lay in whether correct initial representations formed during early learning. If early samples happened to cover typical target culture features (e.g., high probability of containing culturally significant notes), the network learned correctly; otherwise, if early samples focused on marginal features, the network formed biased representations difficult to correct later. This "initialization sensitivity" in machine learning corresponds to the "first impression" effect in human cognition, demonstrating that early experiences can disproportionately shape subsequent identity development paths, especially when cultural distance is large. This finding supports H3's prediction about large distances potentially triggering identity crisis and resonates with Ou Rong's finding about "idealized imaginaries becoming obstacles" [1]—early formed biased cognitions may lead to subsequent integration failure.

D. BIFURCATION EFFECT OF CULTURAL OPENNESS

The individual difference variable cultural openness plays a moderating role in the model. Figure 5 shows the bifurcation diagram of final identity evolution states under different openness values.

To understand this bifurcation effect, imagine a learner at a crossroads. Their cultural openness determines which paths are available. Below a certain threshold, only the path back to a single, source-culture-dominant identity is stable. Above another threshold, the path toward a pluralistic identity becomes accessible and ultimately preferred. To quantitatively identify bifurcation points, we used change point detection algorithm (Pruned Exact Linear Time, PELT) for piecewise linear fitting on Figure 5's scatter points. Results showed two significant change points ($p<0.01$): when openness was below 0.28, regardless of learning intensity, the final formation was always a single-identity dominant pattern, with target culture identification remaining below 0.2. This region is the "closed zone." When openness was between 0.28 and 0.58, evolutionary results showed bistable characteristics—the system could settle into either state depending on its history, much like a ball balanced on a ridge could roll into either of two valleys. Within this "critical zone," the proportion of pluralistic identity patterns increased from 15% at 0.28 to 82% at 0.58. When openness was above 0.58, the system almost always converged to pluralistic identity patterns, with integration degree stabilizing above 0.6—this region is the "open zone." Under openness=0.8, all 50 repetitions converged to pluralistic identity patterns, with mean target culture identification of 0.64 (SD=0.07). The existence of bifurcation points has important theoretical and practical significance. It means identity evolution is not completely deterministic—identical initial conditions may lead to completely different final states due to minor differences. This explains why in empirical research, we often observe different identity outcomes despite similar backgrounds. This finding aligns with Huijing X et al.'s discussion of individual differences' moderating role in cultural adaptation [2]. Cultural openness, as a malleable variable, provides an intervention point. Simulation results show that after openness reaches 0.4, the probability of entering pluralistic identity paths significantly increases even without changing other conditions. This means cultivating open attitudes may be more important than simply increasing exposure quantity.

E. SENSITIVITY ANALYSIS AND MODEL VALIDATION

To evaluate model robustness, we conducted extended sensitivity analysis by systematically varying key parameter values and observing output changes. First, we examined music feature dimensions' influence. Reducing from 16 to 8 dimensions (retaining only pitch-class distribution) or increasing to 24 (adding timbre and dynamics features) changed evolutionary trajectory shape, but the inverted U-shaped relationship and bifurcation point positions remained stable. This indicates model robustness to feature dimension

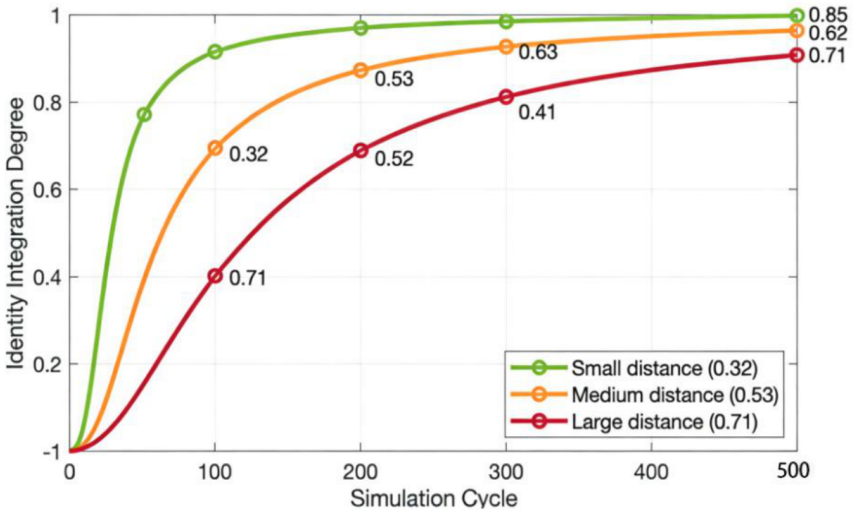


FIGURE 4. Evolutionary Trajectories of Integration Degree Under Different Cultural Distances

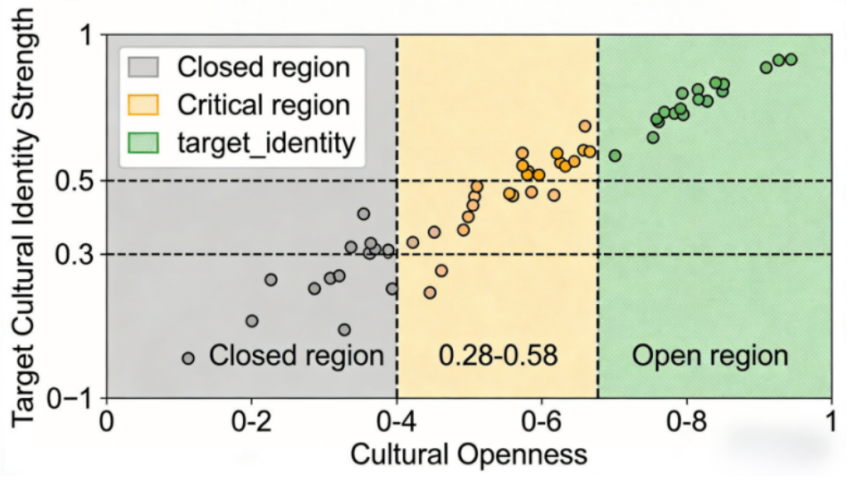


FIGURE 5. Bifurcation Effect of Cultural Openness

choice. Second, we examined SOM network size influence. Adjusting neuron count from 10×10 to 8×8 and 12×12: under 8×8, quantization error slightly increased (+0.04), and bifurcation point position shifted left by approximately 0.03; under 12×12, quantization error decreased (-0.02), but computation time increased by $\sim$40%. Key qualitative conclusions (three-stage evolution, inverted U-shaped relationship, bifurcation phenomena) were reproduced at all scales. Third, we examined learning rate decay speed influence. Adjusting decay constant τ from 100 to 80 and 120: under slow decay, second-stage changes were smoother, with critical points appearing $\sim$20 cycles later, but final integration degree differences did not exceed 0.04; under fast decay, local optima probability slightly increased ($\sim$+5%), but bifurcation point positions remained largely unchanged. Finally, we systematically compared simulation results with existing empirical research. Huijing X et al.'s finding of compensatory effects between hometown music culture connections and multicultural experiences [2] aligns with the model's observation that source culture identification remained at high levels (>0.55) under moderate learning intensity. Fraser's finding that cross-cultural music participation cultivates identification with all humanity [6] resonates with the model's observation of rising integration degree (from 0.42 to 0.65). Ou Rong's research on the association between musical preferences and identity [1] also aligns with the model's mechanism of emotional evaluation driving identity adjustment. These consistencies provide indirect but strong support for model validity.

V. CONCLUSIONS AND DISCUSSION

A. MAIN FINDINGS

This study constructed a hybrid neural network model to computationally simulate the dynamic evolution of identity through the dual lenses of cross-cultural music learning and traditional pattern recognition. By coupling auditory structures with the symbolic logic of weaving crafts, the model captures the multifaceted psychological shifts occurring during the integration of diverse cultural legacies. The

main findings can be summarized as follows:

First, identity evolution exhibits clear stage characteristics: cognitive accumulation phase, structural reorganization phase, and integration stability phase. Particularly noteworthy is that cognitive system changes precede identity changes, with emotional evaluation being the key mediator.

Second, an inverted U-shaped relationship exists between learning intensity and identity integration effectiveness. This finding challenges the intuition that "more is better," revealing an optimal contact intensity in cross-cultural learning. The identity crisis at high intensity emerges not merely from cognitive forgetting, but from the emotional ambivalence and instability resulting from simultaneous loss of source culture fluency and incomplete mastery of the target culture.

Third, cultural distance moderates learning effectiveness, with large distances leading to path bifurcation. When two musical systems differ greatly, learning outcomes are highly sensitive to early experiences, potentially leading to either pluralistic integration or retreat to single identity. This initialization sensitivity, while a challenge for deterministic modeling, realistically captures the importance of "first impressions" in human cultural learning.

Fourth, cultural openness serves as a key moderating variable determining the position of system bifurcation points. When openness exceeds a critical value, the system tends to converge to pluralistic identity patterns; below the critical value, it locks into single-identity patterns. This finding resonates with Huijing X et al.'s discussion of individual differences' role in cultural adaptation[2] and points directions for educational intervention.

B. THEORETICAL CONTRIBUTIONS AND PRACTICAL IMPLICATIONS

At the theoretical level, this study introduces computational modeling methods to cross-cultural identity research, providing a testable dynamic evolutionary framework. Compared to traditional cross-sectional and qualitative research, simulation methods can reveal processes rather than merely describing states, and can identify nonlinear mechanisms rather than just capturing linear correlations. Particularly, bifurcation phenomena discovery suggests we need to reconsider deterministic assumptions about identity development —identical causes may produce different results due to path dependence. Additionally, this study integrates statistical learning theory, self-organizing map modeling, and identity theory, providing a bridge for interdisciplinary dialogue. At the practical level, findings have implications for cross-cultural music education and immigrant integration. First, when designing cross-cultural music programs, attention should be paid to controlling "rhythm"—avoiding too rapid or deep immersion that triggers identity crisis. Gradual, spiral curriculum design may be superior to intensive, immersive arrangements. Second, cultivating cultural openness should be considered equally important as skill training. Before formal learning begins, establishing positive attitudes toward new cultures through stories, images, and other means may significantly improve learning effectiveness. Third, special attention should be given to groups with large cultural distances, providing more cognitive support and emotional buffering to help them navigate sensitive periods.

C. RESEARCH LIMITATIONS AND FUTURE DIRECTIONS

This study has several limitations. while our operationalization of identity as a numerical output is a deliberate and necessary simplification for computational modeling, it inevitably loses qualitative nuances. The model captures changes in the strength of belonging but cannot fully represent phenomena like "identity crisis" as lived experiences of internal conflict and uncertainty. Future models could introduce variables representing identity "uncertainty" or "internal conflict" to provide finer-grained psychological characterization. Second, the simulation of the social environment is overly simplified. Although we set asymmetric social feedback values (0.8 for source, 0.2 for target) to reflect realistic constraints, fixing these values throughout the simulation is a major limitation. Social feedback is inherently dynamic and interactive. We made this choice to isolate and understand the core cognitive-emotional mechanisms, but future research must treat social feedback as a dynamic, co-evolving variable, possibly through multi-agent modeling to simulate social network interactions [10]. Third, the model does not consider individuals' active choices. Learners can choose what to engage with, and this agency may alter evolutionary trajectories. Future models could introduce reinforcement learning mechanisms. Fourth, calibration of simulation results with real data remains insufficient, with current validation being primarily qualitative. Future research could design behavioral experiments to test key model predictions, such as manipulating learning intensity to measure identity integration effects and verify the inverted U-shaped relationship, or manipulating early exposure sample features to test initialization sensitivity hypotheses. The evolution of identity in cross-cultural music learning is a complex and fascinating phenomenon. It concerns both the micro-mechanisms of musical cognition and the macro-narratives of cultural identity; it has universal evolutionary patterns but also unique individual trajectories. Computational models cannot exhaust all reality's richness, yet they provide a lens through which we can glimpse the dynamic interplay between musical structures. We hope this preliminary attempt will inspire more interdisciplinary dialogue and exploration into how woven patterns and melodic narratives collectively shape the evolution of cultural identity.

AUTHOR CONTRIBUTIONS

Conceptualization –Xiaomei Xu and Tiantian Xu; methodology – Xiaomei Xu and Tiantian Xu; investigation – Xiaomei Xu and Tiantian Xu; writing-original draft preparation – Xiaomei Xu and Tiantian Xu. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] R. Ou and Y. Lou, "Conference Summary of 'The Fourth International Symposium on Intermedia Studies'," Comparative Literature and Cross-cultural Studies, vol. 4, no. 01, pp. 3-7, 2020.

[2] X. Huijing, H. Yunlong, L. Yuyao, et al., "Constructing college students' musical cultural identity and self-confidence under cross-regional cultural integration: an integrative study based on cultural connections and multicultural experiences," Frontiers in Psychology, vol. 17, 1674065, 2026, doi: 10.3389/FPSYG.2026.1674065.

[3] J. K. Pazdera, "Computational modeling of musical enculturation: An investigation of multicultural music learning using self-organizing maps [D]," Haverford: Haverford College, 2016.

[4] M. T. Pearce, "Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation," Annals of the New York Academy of Sciences, vol. 1423, no. 1, pp. 378-395, 2018, doi: 10.1111/nyas.13654.

[5] Y. Meng, "Innovative Integration of Traditional Ethnic Music Elements in Modern Popular Music," Famous Works, vol. (12), pp. 49-51, 2025, doi: 10.3969/j.issn.2095-8854.2025.12.018.

[6] T. Fraser, "Cultivation of global consciousness through intercultural music engagement," Asian Journal of Social Psychology, vol. 28, no. 3, e70022, 2025, doi: 10.1111/AJSP.70022.

[7] C. Doll, "Learning, Teaching, and Musical Identity: Voices across Cultures," Popular Music, vol. 31, no. 1, pp. 179-180, 2012, doi: 10.1017/S0261143011000547.

[8] F. Hongyu and T. Jinmei, "Exploring the influence of national music lessons on subjective well-being, self-esteem, and national identity among university students: a case study from China," Frontiers in Psychology, vol. 14, 1151007, 2023, doi: 10.3389/FPSYG.2023.1151007.

[9] Y. Guo, "Toward Multicultural Identity Formation in Classical Vocal Pedagogy: Cultivating Artistic Expression Through Vocal Performance," Asia-pacific Journal of Convergent Research Interchange (APJCRI), vol. 11, no. 9, pp. 407-426, 2025, doi: 10.47116/apjcri.2025.09.29.

[10] X. Li, "Application and Problems of Artificial Intelligence in Cross-cultural Communication of Ethnic Minority Music," Art Communication Research, vol. (6), pp. 87-95, 2025, doi: 10.3969/j.issn.2096-8183.2025.06.010.