Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

How Learning Reshapes Aesthetic Evaluation of AI-generated and Traditional Xilankapu Patterns: Evidence from a Longitudinal Intervention Study

Shiyuan Zhu
Changsha University of Science and Technology
Jinxia Hu
Changsha University of Science and Technology
Ziyue Cheng
Changsha University of Science and Technology
Yanyou Hu
Changsha University of Science and Technology
Jiahao Sun
sunjiahao@csust.edu.cn
Changsha University of Science and Technology

Article

Keywords: aesthetic evaluation, learning intervention, artificial intelligence, human–AI collaboration, creator attribution, cultural cognition, Xilankapu

Posted Date: October 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10520041/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Cultural knowledge may alter how viewers evaluate traditional and AI-generated cultural artifacts, yet most studies of AI-generated art use cross-sectional designs. This study examined whether structured learning changes aesthetic evaluation and creator attribution for traditional Xilankapu patterns and AI-generated patterns.

Methods

A 30-day single-cohort pre–post intervention study was conducted with 118 recruited participants who had no prior experience with Xilankapu; after quality screening, 100 valid participants were included in the final analysis. Participants evaluated traditional, human-out-of-the-loop (HOTL), and human-in-the-loop (HITL) patterns across 15 aesthetic dimensions and completed creator-attribution judgments before and after the intervention. Analyses included descriptive statistics, analyses of variance, linear mixed-effects models, one-sample t tests, and SHapley Additive exPlanations (SHAP) based on an XGBoost model. Before beginning the questionnaire, participants viewed an electronic informed consent form and were required to indicate their agreement before proceeding. Participation was voluntary, and no directly identifying personal information was retained.

Results

After the intervention, the weighting of aesthetic evaluation shifted from perceptual features such as expressiveness and vividness toward conceptual dimensions such as typicality and meaning. HITL patterns received the highest beauty ratings at both time points, whereas HOTL patterns received the lowest ratings. Attribution accuracy for traditional patterns increased significantly from 51.8% to 69.2%.

Conclusions

Structured cultural learning was associated with a shift from sensory-driven toward concept-driven aesthetic judgment and with improved recognition of traditional cultural patterns. The findings also suggest that human involvement remains important when AI is used to generate culturally meaningful visual artifacts.

Clinical trial registration

Not applicable. This study was not a clinical trial. It did not involve a medical intervention, clinical treatment, diagnosis, prevention, or health-related outcome. The intervention consisted of structured learning materials on Xilankapu culture, semiotic interpretation, and AI-based pattern generation, and the outcomes were aesthetic ratings and creator-attribution judgments. Therefore, clinical trial registration and CONSORT reporting are not applicable.

Background

Generative artificial intelligence (AIGC) has attracted increasing attention in visual arts and cultural heritage design. Researchers have examined how these technologies expand formal innovation and reshape cultural production [1]. At the same time, AI-generated representations of traditional culture raise questions in cognitive psychology and neuroaesthetics about how participants perceive and evaluate non-human-generated cultural artifacts [2].

Existing empirical studies on AI-generated art have typically compared AIGC works with human-created works by concealing the creator's identity and asking participants to make attribution judgments and aesthetic evaluations [3, 4]. These studies have also examined algorithm aversion by linking creator attribution with aesthetic evaluation [5]. However, most of this work has relied on static, cross-sectional comparisons between AI-generated and human-created works. Aesthetic cognition, by contrast, is a dynamic and plastic process shaped by cultural context and prior knowledge. As individuals acquire relevant knowledge, their aesthetic judgments may shift from surface-level perceptual features toward deeper conceptual and semantic dimensions [6, 7]. It therefore remains unclear how structured learning interventions reshape both attribution judgments and aesthetic evaluation over time.

This study extends previous research on AI-generated art to the context of traditional Chinese ethnic cultural symbols. By integrating Leder's aesthetic processing model with Peirce's triadic theory of signs, the study conducted a 30-day pre–post intervention experiment to compare participants' responses before and after a cultural learning intervention across three types of patterns: purely AI-generated patterns (HOTL), human–AI collaborative patterns (HITL), and traditionally created patterns. SHAP analysis was used to examine how the intervention changed the weighting of aesthetic perception dimensions and how these changes were associated with creator attribution and algorithm aversion. The findings provide quantitative evidence for aesthetic cognitive plasticity and offer practical implications for human–AI collaborative creation of traditional cultural patterns.

Aesthetic evaluation of AI-generated art

Recent advances in artificial intelligence have enabled algorithms to generate images, music, and prose based on large-scale training data and prior examples [8, 9]. The application of AI in design has changed design production methods and raised new questions about aesthetic evaluation. Historically,

shifts in design tools have often been accompanied by changes in aesthetic norms and creative practices. Accordingly, scholarly attention has moved beyond the technical novelty of AI-generated works toward how audiences perceive, interpret, and evaluate machine-generated artworks [1].

From an aesthetic and philosophical perspective, some scholars argue that originality, creativity, and emotion are typically associated with human artistic creation [10, 11]. Because AI-generated artworks are produced through computational models and algorithmic procedures, they are often regarded as relying on the recombination of existing patterns and rules. From this perspective, AI-generated works may lack the emotional depth, subjective intention, and implicit meaning commonly associated with traditional human-created artworks [12–15]. Empirical studies have further shown that although viewers may find it difficult to distinguish AI-generated images from human-created ones, they often assign higher aesthetic value to works believed to be created by humans [8, 9]. These findings suggest that audiences may hold implicit biases against AI-generated artworks. At the same time, other studies indicate that collaboration between human creativity and AI automation can improve both the efficiency of artistic production and the perceived aesthetic value of creative works [5, 8, 16], highlighting the potential of human–AI collaboration in artistic practice.

Existing findings suggest that participants often show aversion to AI-generated works and a preference for human-made ones in aesthetic evaluations. However, current experimental paradigms have largely overlooked a key cognitive dimension: the dynamic plasticity of the aesthetic cognitive system [7, 17]. Most studies capture instantaneous and subjective judgments of AI- and human-created works at a single point in time. Because cognition is rarely static, this approach cannot fully explain whether aesthetic preferences change systematically as individuals accumulate knowledge and experience. This study therefore examines how aesthetic evaluation shifts over time in the context of artificial intelligence.

Leder's aesthetic processing model

Aesthetics has long been concerned with sensory perception and aesthetic experience. Alexander Baumgarten coined the term “aesthetics” to denote a science of perception, defining aesthetic activity as sensory experience or the pleasure and satisfaction it produces. In contemporary research, the integration of the humanities and the natural sciences has encouraged empirical approaches to aesthetic questions. Within this interdisciplinary context, neuroaesthetics emerged as a field that examines aesthetic experience through cognitive and neuroscientific methods. In the late twentieth century, Semir Zeki investigated the neural connections between visual art and visual cortical regions and helped establish the concept of neuroaesthetics [18]. As the field developed, Pearce and colleagues defined neuroaesthetics as the cognitive neuroscience of aesthetic experience [7]. One important branch of this research, represented by Leder and colleagues, combines psychological theory with questionnaires, interviews, quantitative experiments, and behavioral analysis to explain how individuals judge and appreciate aesthetic objects, including works of art [6, 19].

The core contribution of Leder’s framework lies in identifying key variables that influence the multi-stage process of aesthetic experience. In 2013, Leder further formalized this framework as an aesthetic information-processing model, which describes aesthetic experience as a sequence of processing stages. This model has since become an important framework for empirical research on artistic perception and aesthetic experience. In the present study, the stages of the model are organized as perceptual analysis, implicit memory integration, explicit classification, cognitive mastery, emotional state, and aesthetic evaluation, culminating in final aesthetic perception and cognitive judgment.

Based on this theoretical framework, the present study uses Tujia Xilankapu patterns as cultural visual stimuli. Through a learning intervention, it examines participants’ aesthetic processing pathways from two dimensions: creator attribution and aesthetic perception. The study further explores how participants evaluate traditional ethnic patterns and AI-generated patterns, as well as the cognitive mechanisms that drive these evaluations, as shown in Figure 1.

Semiotic framework and learning intervention

Charles Sanders Peirce’s triadic theory of signs is a core theory in semiotics [20]. Peirce argued that signs and meaning are inseparable: signs do not possess inherent meaning, but acquire meaning through interpretation. He divided the structure of a sign into three elements: the representamen, the object, and the interpretant. The representamen refers to the external form of the sign, including attributes such as shape and structure. The object refers to the entity or concept represented by the sign, which may be concrete or abstract. The interpretant refers to the meaning derived from the relation between the sign’s object and representamen within a specific cultural context. Semiotic methods can therefore help participants recognize the multiple possibilities of signs and better understand their narrative content.

Based on Peirce’s triadic theory of signs, this study established a three-dimensional teaching framework comprising the sign, the object, and the interpretant. Tujia Xilankapu patterns are graphic expressions of ethnic spirit and important symbols for disseminating Chinese ethnic culture. The learning intervention was organized around this three-dimensional framework. The “object” dimension covered the historical origins of Xilankapu patterns, common techniques for transforming natural objects into geometric forms, and their applications in Tujia daily life, thereby linking abstract patterns to traceable cultural objects. The “representamen” dimension analyzed formal characteristics such as color patterns and compositional principles to decode the visual language of the patterns. The “interpretant” dimension examined Tujia religious beliefs and common interpretations of pattern meanings, as shown in Figure 2.

Research questions

As an emerging field [1, 21], generative AI art has raised ethical, social, and legal questions [4] and has also challenged long-held assumptions about art, creativity, and emotional expression. Consequently, understanding how people perceive, interpret, and appreciate machine-generated art has become an important research topic [1, 22].

Neuroaesthetics views aesthetic experience as a psychological process involving the cognitive and perceptual evaluation of stimuli [6, 23, 24]. This study divides the experiment into two parts: aesthetic perception and creator attribution. It adopts a longitudinal intervention paradigm with the same group of participants to examine how aesthetic cognition changes as participants acquire domain-specific cultural knowledge.

In the aesthetic perception module, aesthetic activity was treated as a multidimensional process rather than a single judgment of "beauty." Drawing on previous research in art and psychology [3, 25, 26] and on the expressive characteristics of Xilankapu patterns [27], we identified 15 dimensions that may influence participants' aesthetic perception. Participants completed two rounds of subjective ratings, before and after the learning intervention, based on prompts for these evaluation dimensions. Given the dynamic and plastic nature of aesthetic cognition, the learning intervention was expected to alter the mechanisms through which each aesthetic dimension influenced the overall "beauty" rating. Therefore:

Research Question 1: How will aesthetic perception change before and after the learning intervention?

Research Question 2: How will aesthetic perception dimensions change before and after the learning intervention?

In the creator attribution module, prior research has found that people's preferences for human-created artworks are inconsistent [3, 28], while their ability to correctly distinguish the origin of artworks (human or machine) ranges from approximately 42.5% to 63.8% [4]. However, these findings need to be re-examined for several reasons. First, early studies were conducted when AIGC technology was less mature and often visibly different from human-created works. Second, many studies included broad participant samples, such as experts in image processing or artificial intelligence together with general participants, which may have reduced the specificity of the results. Therefore, although this study investigates phenomena similar to those in previous research, it introduces differences in experimental design, participant sampling, stimulus materials, and measurement methods to better capture the factors that distinguish human-created from machine-generated artworks [4]. By using current generative technologies and a learning-intervention design with the same cohort of participants who had no prior exposure to the study content, this study re-examines the following questions:

Research Question 3: How does creator-attribution accuracy change before and after the learning intervention across traditional, HOTL, and HITL patterns?

Research Question 4: How do human attitudes toward AI-generated patterns change after the learning intervention, as reflected in the relationship between attribution accuracy and aesthetic evaluation?

Specifically, based on Leder's aesthetic processing model and Peirce's semiotic framework, this study addresses two parallel dimensions. First, from the aesthetic perception dimension, it analyzes participants' evaluations before and after the learning intervention to determine whether a relatively brief intervention makes their evaluation criteria more closely aligned with those of experts and which dimensions change significantly. Second, from the creator attribution dimension, it investigates whether the learning intervention influences the criteria used for discrimination. By analyzing the results across these two dimensions and their interactions, the study provides empirical evidence for the effectiveness of learning interventions in the evaluation of generative cultural works.

Methods

Study design and setting

This study used a 30-day, single-cohort pre–post intervention design. The same participants completed the aesthetic-rating and creator-attribution tasks before and after a structured learning intervention on Xilankapu culture, semiotic interpretation, and AI-based pattern generation. The study compared responses to traditional, HOTL, and HITL pattern stimuli at both time points.

Participants

Prior knowledge can significantly influence the evaluation of artworks [29]. To reduce this influence, the study recruited 118 volunteers. All participants had a certain level of artistic training and aesthetic sensitivity, primarily as students or professionals in visual communication, digital media art, or related design fields. This ensured that they had basic analytical skills for evaluating visual elements. None of the participants had prior exposure to Xilankapu culture or AI-generated pattern design before the learning intervention.

Participants completed a pre-intervention questionnaire through in-person interviews. After a 30-day cognitive training program, they completed the same questionnaire again using the same procedure. Eighteen responses were excluded because participants failed the attention check or completed the full evaluation task in less than 3,000 s, a threshold set to exclude responses unlikely to reflect careful evaluation of all stimulus samples. After these exclusions, the final analytic sample comprised 100 participants, representing 84.7% of the recruited sample. The final sample consisted of 58 male and 42 female participants, with a mean age of 32.55 years (SD = 5.44). All participants were informed of the study's purpose, procedures, potential risks, duration, voluntary nature of participation, right to withdraw, and data handling procedures.

Ethics approval and informed consent

The study protocol was reviewed and approved by the Ethics Committee of Changsha University of Science and Technology, as documented in the ethical review approval form dated 1 January 2025. The ethics approval document states that the project complied with ethical principles and was approved for conduct. All procedures were performed in accordance with the Declaration of Helsinki and relevant institutional guidelines. Before accessing the questionnaire, all participants were presented with an electronic informed consent form describing the purpose and procedures of the study, the voluntary

nature of participation, their right to withdraw, potential risks, and the handling of their data. Only participants who provided electronic informed consent could proceed to the questionnaire. No directly identifying personal information was retained, and all data were de-identified before analysis.

Stimulus preparation and AI-generation procedure

Pattern database and model training

A total of 275 original Xilankapu tapestries were collected through literature reviews and field surveys. Clear vector patterns were traced using digital technology to establish a primary database. Among currently popular generative models, Stable Diffusion can preserve key features of traditional patterns and meet the artistic design requirements of this study [30]. Compared with traditional GAN models, Stable Diffusion provides greater stability and image quality. Its latent space also improves computational efficiency during generation, offering advantages in computational cost and speed [31]. Therefore, this model was selected as the main source for generating AI-based pattern materials.

LoRA technology enables efficient fine-tuning of large models through low-rank matrix decomposition. It reduces computational costs and memory requirements while maintaining model stability and improving output quality. Applied to Stable Diffusion, LoRA can reduce the risk of overfitting under limited-resource conditions while adapting the model to the specific task of generating Xilankapu-style images. This allowed the model to capture the distinctive visual features of Xilankapu art.

The network model was built using PyCharm, Python 3.6, and the PyTorch deep learning framework. Training and validation were primarily performed using an NVIDIA GeForce RTX 4080 GPU. Based on the database established above, the uniform training resolution was set to 512×512. The Stable Diffusion 1.5 model was selected as the base model, with hyperparameters set as follows: learning rate = 4e-4, batch size = 4, epochs = 10. The AdamW-8bit optimizer was used to update model parameters, with L2 regularization applied. A Cosine Annealing learning rate scheduler was employed to dynamically adjust the learning rate.

The LoRA fine-tuning process consisted of three stages. First, parameter injection embedded the low-rank matrix ΔW into the cross-attention layer of the pre-trained U-Net. Second, noise diffusion and denoising were performed iteratively in the latent space to approximate the Xilankapu style distribution. Finally, weight fusion linearly superimposed the LoRA matrix onto the original weights to control style intensity. With weights ranging from 0.5 to 1.0, the model demonstrated strong pattern-based style representation.

Given the importance of image stimuli [32], 10 pattern designers were invited to generate 30 AI-generated patterns each based on the provided keywords: traditional Chinese patterns, beauty, rich colors, vivid contrasts, geometric compositions, natural elements, profound meanings, exquisite craftsmanship, and complex patterns. Five traditional ethnic patterns of different content types were randomly selected from the original pattern library. To reduce the influence of subjective stimulus selection on the ecological validity of the results, five pattern experts and five cultural custodians were invited to standardize and select the final research stimuli based on four criteria: form, color, characteristics, and innovation. Ten AI-generated patterns were ultimately selected for the experiment.

Based on the creation method, the AI-generated patterns were categorized into two types: HOTL patterns, which were autonomously generated by AI without human intervention, and HITL patterns, which were co-created by human designers and AI. The five highest-scoring generated patterns were selected, and pattern designers were invited to refine their lines, shapes, and color tones to create HITL pattern materials. The remaining five patterns were designated as HOTL pattern materials. To compare traditional ethnic patterns with AI-generated patterns, the five highest-scoring traditional ethnic patterns from the original primitives database were selected as the reference group, as shown in Figure 3.

Measures

Aesthetic perception. Although many factors influence aesthetic perception, no evaluation method has been developed specifically for traditional Chinese patterns. Given the complexity and multifactorial nature of aesthetic perception in AI-generated content, a single evaluation criterion cannot fully capture the key elements that elicit aesthetic experiences. Rooted in Leder's aesthetic processing model, this study integrates research findings in art and psychology [3, 25, 26] with the expressive characteristics of Xilankapu patterns [27] to construct an aesthetic perception evaluation system. The 15 dimensions in the system were mapped onto three processing stages following participants' exposure to image stimuli. "Familiarity," "typicality," and "peak-shifting effect" were treated as indicators of implicit memory integration. "Style," "content," "color," "composition," "vividness," "refinement," "harmony," and "expressiveness" corresponded to explicit classification and captured participants' perception of visual form. "Resonance" and "meaning" represented cognitive mastery and required participants to use semantic knowledge to understand the work's connotations. By tracking participants' progression from implicit memory integration to cognitive mastery, the study obtained an overall evaluation of "beauty" and changes in emotional responses after viewing the patterns, as shown in Table 1.

Table 1. Aesthetic dimensions

Stage	Dimension	Item wording
Perceptual Analysis	Color	The pattern features beautiful colors
	Expressiveness	I think the pattern exhibits a rich visual expression
	Vividness	The pattern is vivid and vibrant
	Refinement	The pattern is exquisite
	Composition	The composition of this pattern is harmonious
	Harmony	Feel the overall harmony of the pattern
Implicit Memory Integration	Peak-shift effect	There are features in the works that can represent the type of pattern
	Familiarity	It feels familiar, as if I've seen it before
	Typicality	I sense the typical characteristics of ethnic patterns
Explicit classification	Style	I sense a unique style
	Content	I feel like it seems to be expressing something
Cognitive Mastery	Resonance	I can establish some emotional connections with it
	Meaning	I believe that patterns express the depth and essence of a nation's culture, history, or spiritual signific
Emotional State	Emotional response	Upon seeing this pattern, my aesthetic appreciation has increased
Aesthetic Evaluation	Beauty	This pattern is full of aesthetic beauty

Participants were asked to subjectively rate 15 stimulus samples across the aforementioned 15 evaluation dimensions. For clarity, each dimension was described in detail; higher scale scores indicate stronger agreement with the dimension (1 = strongly disagree, 6 = strongly agree).

Creator attribution. This module used attribution experiments to examine whether participants could accurately identify the actual creators of pattern works and to explore attitudes toward AI-generated patterns through the relationship between discrimination rates and aesthetic perception [24, 28]. To examine the role of human–AI collaboration, the creation method of HITL patterns was concealed; participants were informed only that the experiment involved two types of patterns: AI-generated patterns and original traditional ethnic patterns. The attribution task asked: "Do you believe this pattern was created by a human artist or a machine (AI) artist?" To provide participants with sufficient criteria for judgment, the questionnaire combined the artistic characteristics of Xilankapu with general features of human-created works and included 16 guiding evaluation dimensions: rhythm, consistency, regularity, complexity, depth of meaning, abstractness, intentionality, uniqueness, creativity, expressiveness, emotional connection, value, aesthetic appeal, harmony, impact, and refinement. Participants selected one or more criteria from this list or provided an open-ended explanation of how they judged the creator of each pattern.

Learning intervention and procedure

Learning intervention. According to Peirce's triadic semiotics, the object refers to the archetype or entity represented in the pattern. In this study, it includes the Tujia people's artistic techniques for designing Xilankapu and the ethnic characteristics used to depict natural imagery. These features correspond to implicit memory integration in the aesthetic process. The representamen refers to the explicit visual features of the pattern, including colors, composition, and style as analyzed by participants after receiving the pattern stimulus; these features correspond to explicit classification. The interpretant concerns the interpretation of pattern meaning and corresponds to cognitive mastery. Based on this framework, the study established a one-to-one correspondence between semiotic processes and aesthetic cognitive stages, forming the correspondence shown in Figure 4. After completing the first questionnaire, participants underwent 30 days of cognitive training covering the three dimensions of semiotics and the basic principles of AI-generated patterns. To ensure experimental validity, participants were instructed not to seek additional information about Xilankapu patterns outside the intervention materials during the study period. They then completed the second questionnaire to observe changes in each process before and after the learning intervention. We hypothesized that this structured knowledge input would help participants engage in deeper aesthetic reflection and alter the weighting of their aesthetic judgments, as shown in Figure 4.

Statistical analysis

Data were screened using box plots and z scores; no missing values were identified. Shapiro–Wilk and Levene tests were used to assess distributional assumptions and homogeneity of variance. Descriptive statistics and one-way repeated-measures analyses of variance were calculated in SPSS version 29.0.2.0.

Separate linear mixed-effects models were fitted with the lme4 package in R [33] to estimate the association of each of the 14 dimensions other than beauty with beauty ratings. In each model, one aesthetic dimension was entered as the focal fixed effect, pattern category was included as a covariate, and participant ID and pattern ID were included as random intercepts. Creator-attribution accuracy was calculated separately for HOTL, HITL, and traditional patterns as the proportion of correctly classified stimuli within each pattern type for each participant. Because the attribution task involved two possible creator categories, chance performance was defined as 0.50. Two-sided one-sample t tests were used to compare participant-level attribution accuracy with chance at each time point. Paired-samples t tests were additionally used to examine pre–post changes in attribution accuracy within each pattern type.

Effect sizes were reported as Cohen's d for one-sample tests and Cohen's dz for paired comparisons. XGBoost models and SHAP values were used to examine the contribution of each aesthetic dimension to predicted beauty ratings. All tests were two-sided, with $p < .05$ indicating statistical significance.

Use of generative AI in manuscript preparation

ChatGPT (OpenAI) was used to assist with English-language editing and manuscript formatting. The authors reviewed and revised all generated text and take full responsibility for the final content.

Results

Aesthetic ratings before and after the intervention

For Research Question 1, which examines changes in aesthetic perception ratings before and after the learning intervention, we analyzed the mean and standard deviation of the 15 aesthetic perception items and the ANOVA results for each pattern type under the traditional ethnic, HOTL, and HITL conditions. Tables 2 and 3 present descriptive statistics, significance tests, and post hoc comparisons across pattern types.

The results show significant differences in beauty ratings among traditional ethnic patterns, HOTL patterns, and HITL patterns both before and after the learning intervention (pre-intervention: $F = 8.611$, $p < 0.001$; post-intervention: $F = 11.907$, $p < 0.001$). Descriptively, the distribution of aesthetic ratings differed across the three stimulus types following the intervention. Ratings increased on several culturally relevant dimensions, particularly typicality, meaning, and emotional response for traditional patterns, whereas some perceptual ratings showed little change or modest decreases. HOTL patterns generally remained the lowest-rated group.

Before the intervention, HITL patterns had the highest descriptive mean scores on six dimensions, including vividness ($M = 4.33$) and harmony ($M = 4.46$). At the post-intervention assessment, traditional patterns showed numerically higher mean ratings than HITL patterns on 12 of the 15 evaluation dimensions, including typicality ($M = 4.65$, $\Delta = +0.37$), content ($M = 4.31$, $\Delta = +0.19$), and emotional response ($M = 4.41$, $\Delta = +0.27$). HITL patterns remained close behind, whereas HOTL patterns continued to receive the lowest scores.

Table 2. Descriptive statistics of aesthetic ratings before the intervention

	Mean/SD			*F*	*P*	Post-hoc Tests
Evaluation Dimension	HOTL	HITL	Traditional Ethnic Patterns			
Beauty	3.75	4.05	3.90	8.611	<.001**	HOTL < Traditional < HITL
	(1.235)	(1.178)	(0.997)			
Expressive	4.04	4.23	4.07	4.029	0.018*	HOTL < Traditional < HITL
	(1.094)	(1.060)	(1.086)			
Style	4.19	4.34	4.25	2.580	0.076	No significant pairwise differences
	(1.254)	(0.966)	(0.918)			
Composition	3.85	3.98	3.85	1.813	0.163	No significant pairwise differences
	(1.185)	(1.126)	(1.030)			
Vividness	3.85	4.33	4.12	25.768	<.001**	HOTL < Traditional < HITL
	(1.182)	(0.981)	(0.996)			
Refinement	3.81	4.26	4.08	17.380	<.001**	HOTL < Traditional < HITL
	(1.324)	(1.089)	(1.205)			
Typicality	3.98	4.19	4.28	10.145	<.001**	HOTL < HITL < Traditional
	(1.139)	(1.071)	(1.025)			
Harmony	4.14	4.46	4.23	12.073	<.001**	HOTL < Traditional < HITL
	(1.254)	(0.981)	(0.947)			
Color	4.21	4.35	4.31	2.016	0.134	No significant pairwise differences
	(1.147)	(1.117)	(1.053)			
Peak-shift effect	4.12	4.37	4.16	8.529	<.001**	HOTL < Traditional < HITL
	(1.108)	(0.983)	(0.967)			
Content	3.89	4.17	4.12	8.788	<.001**	HOTL < Traditional < HITL
	(1.216)	(1.034)	(1.041)			
Meaning	3.94	4.12	4.17	6.016	0.003**	HOTL < HITL < Traditional
	(1.214)	(1.029)	(1.054)			
Resonance	3.69	4.07	4.17	25.953	<.001**	HOTL < HITL < Traditional
	(1.155)	(1.056)	(1.077)			
Emotional response	3.88	4.16	4.14	8.707	<.001**	HOTL < Traditional < HITL
	(1.229)	(1.162)	(1.095)			
Familiarity	3.65	3.93	3.80	6.670	<.001**	HOTL < Traditional < HITL
	(1.268)	(1.236)	(1.093)			

*Note: P > 0.*05 indicates no significant difference; 0.01 < P < 0.05 indicates a significant difference, marked with *; P < 0.01 indicates a highly significant difference, marked with **

Table 3. Descriptive statistics of aesthetic ratings after the intervention

	Mean/SD			*F*	*P*	Post-hoc Tests
Evaluation Dimension	HOTL	HITL	Traditional Ethnic Patterns			
Beauty	3.83	4.23	4.14	11.907	<.001**	HOTL < Traditional < HITL
	(1.308)	(1.454)	(1.312)			
Expressive	4.16	4.47	4.36	7.435	<.001**	HOTL < Traditional < HITL
	(1.314)	(1.413)	(1.182)			
Style	4.26	4.34	4.55	6.621	<.001**	HOTL < HITL < Traditional
	(1.419)	(1.329)	(1.134)			
Composition	3.93	4.29	4.14	9.674	<.001**	HOTL < Traditional < HITL
	(1.236)	(1.385)	(1.204)			
Vividness	3.90	4.18	4.33	15.267	<.001**	HOTL < HITL < Traditional
	(1.313)	(1.370)	(1.069)			
Refinement	3.82	4.16	4.24	14.855	<.001**	HOTL < HITL < Traditional
	(1.437)	(1.343)	(1.128)			
Typicality	3.95	4.40	4.65	43.431	<.001**	HOTL < HITL < Traditional
	(1.315)	(1.252)	(1.057)			
Harmony	4.08	4.42	4.47	14.960	<.001**	HOTL < HITL < Traditional
	(1.272)	(1.239)	(1.126)			
Color	4.25	4.49	4.50	7.353	<.001**	HOTL < HITL < Traditional
	(1.212)	(1.301)	(1.060)			
Peak-shift effect	4.18	4.34	4.39	4.693	0.009**	HOTL < HITL < Traditional
	(1.243)	(1.206)	(0.938)			
Content	3.79	4.27	4.31	26.522	<.001**	HOTL < HITL < Traditional
	(1.374)	(1.212)	(1.156)			
Meaning	4.00	4.39	4.52	26.786	<.001**	HOTL < HITL < Traditional
	(1.325)	(1.133)	(0.988)			
Resonance	3.70	4.22	4.20	26.962	<.001**	HOTL < Traditional < HITL
	(1.320)	(1.313)	(1.151)			
Emotional response	4.10	4.38	4.41	10.192	<.001**	HOTL < HITL < Traditional
	(1.226)	(1.208)	(1.103)			
Familiarity	3.70	3.90	3.86	2.789	0.062	No significant pairwise differences
	(1.434)	(1.455)	(1.297)			

*Note: P > 0.*05 indicates no significant difference; 0.01 < P < 0.05 indicates a significant difference, marked with *; P < 0.01 indicates a highly significant difference, marked with **

The descriptive comparisons indicate that, although participants were unaware of the origin of the works, HITL patterns received higher beauty ratings than purely AI-generated HOTL patterns. This suggests that human involvement may improve the perceived aesthetic quality of AI-generated cultural patterns. Before the learning intervention, participants' evaluations relied more on aesthetic intuition than on cultural knowledge, yet they were still sensitive to the meanings conveyed by the works. These results indicate that human participation provides advantages in pattern creation that are difficult for algorithms to fully replicate. They also suggest that current AI still has limitations in conveying cultural spirit and semantic depth. For innovative cultural patterns, human designers remain important for improving the acceptance and aesthetic quality of AI-generated outputs.

The distribution of beauty ratings across the three pattern-generation modes before and after the learning intervention is shown in Figure 5.

Changes in aesthetic evaluation dimensions

For Research Question 2, separate linear mixed-effects models were fitted to examine the association between each of the 14 aesthetic dimensions and beauty ratings. In each model, beauty was entered as the dependent variable, one aesthetic dimension was entered as the focal fixed effect, and pattern

category was included as a covariate. Participant ID and pattern ID were included as random intercepts to account for repeated observations nested within participants and stimuli. Separate models were fitted for the pre-intervention and post-intervention assessments, each based on 1,500 observations from 100 participants evaluating 15 patterns. The results are presented in Tables 4 and 5. Before the intervention, expressiveness, composition, vividness, typicality, color, emotional response, familiarity, and meaning were significantly associated with beauty ratings. After the intervention, expressiveness, composition, vividness, typicality, harmony, color, peak-shift effect, meaning, and emotional response were significantly associated with beauty ratings. After the learning intervention, harmony and the peak-shift effect became statistically significant, whereas familiarity was no longer statistically significant. These findings indicate that short-term knowledge-based learning can alter participants' aesthetic evaluation patterns for specific art forms.

The change in the familiarity dimension is particularly noteworthy. Before the intervention, familiarity was significantly associated with beauty ratings ($F = 46.868$, $p < .001$), whereas after the intervention this association was no longer statistically significant ($F = 1.000$, $p = .318$). This result is consistent with a possible cognitive-decoupling account: untrained observers may interpret AI-generated visual collages as familiar cultural symbols because they rely on superficial form. This pattern suggests that cultural knowledge may help participants distinguish superficial visual familiarity from culturally grounded typicality and meaning.

Table 4. Linear mixed-effects model before the intervention

Fixed effect	F	p	df_2
Expressiveness	29.149	< .001**	1405.077
Style	0.399	.528	1452.771
Composition	103.666	< .001**	1465.725
Vividness	16.363	< .001**	1456.346
Refinement	0.109	.742	1465.460
Typicality	9.275	.002**	1457.063
Harmony	1.256	.263	1456.972
Color	12.597	< .001**	1460.747
Peak-shift effect	1.883	.170	1452.761
Content	3.180	.075	1421.866
Meaning	5.456	.020*	1462.991
Resonance	.715	.398	1766.793
Emotional response	19.587	< .001**	1467.576
Familiarity	46.868	< .001**	1377.380

Note: *p < .05; **p < .01.

Table 5. Linear mixed-effects model after the intervention

Fixed effect	F	p	df_2
Expressiveness	67.46933	< .001**	1482.127
Style	3.344	.067	1393.792
Composition	14.043	< .001**	1441.811
Vividness	11.473	.001**	1467.278
Refinement	0.587	.444	1401.639
Typicality	4.669	.031*	1355.063
Harmony	15.013	< .001**	1467.422
Color	17.893	< .001**	1469.977
Peak-shift effect	8.843	.003**	1453.019
Content	0.029	.865	1078.241
Meaning	9.864	.002**	620.622
Resonance	0.348	.555	938.515
Emotional response	23.328	< .001**	1460.386
Familiarity	1.000	.318	1101.831

Note: *p < .05; **p < .01.

To investigate the extent to which different aesthetic dimensions contributed to beauty ratings, SHAP values were used to quantify the contribution of each feature to the model's prediction for a specific instance. A positive SHAP value indicates that the feature increases the beauty score, whereas a negative value indicates that it decreases the score. SHAP summary plots [34] for the 14 aesthetic dimensions, excluding beauty, were calculated based on the XGBoost model [35] and are shown in Figure 6. Dimensions identified as significant factors are highlighted with a dark background. The results show that the learning intervention altered the contribution patterns of aesthetic dimensions to beauty ratings. Before the intervention, expressiveness and vividness were the dominant factors. After the intervention, typicality and meaning emerged as core predictive dimensions alongside vividness. The SHAP plots in Figure 6 also show a polarization pattern after the intervention, especially in evaluations of HOTL works.

After the intervention, the SHAP-value distributions appeared more dispersed for several features, particularly in the HOTL condition. This pattern suggests greater heterogeneity in the contribution of these features to predicted beauty ratings. However, the SHAP plots alone cannot determine whether this heterogeneity reflects increased tolerance, stronger rejection, or other participant-level response strategies.

Creator attribution accuracy

To address Research Question 3, creator-attribution accuracy was calculated at the participant level for each pattern type and compared with the chance level of 0.50. Before the intervention, mean attribution accuracy was 53.2% for HOTL patterns, 41.4% for HITL patterns, and 51.8% for traditional patterns. Accuracy for HOTL patterns did not differ significantly from chance, $t(99) = 1.85$, $p = .068$, 95% CI [49.8%, 56.6%]. Accuracy for traditional patterns also did not differ from chance, $t(99) = 0.68$, $p = .497$, 95% CI [46.6%, 57.0%]. By contrast, attribution accuracy for HITL patterns was significantly below chance, $t(99) = -3.61$, $p < .001$, 95% CI [36.7%, 46.1%], indicating that participants frequently misclassified human–AI collaborative patterns as traditionally created patterns.

After the intervention, attribution accuracy increased to 57.6% for HOTL patterns, 44.8% for HITL patterns, and 69.2% for traditional patterns. HOTL accuracy was significantly above chance, $t(99) = 3.89$, $p < .001$, 95% CI [53.7%, 61.5%], and traditional-pattern accuracy was also significantly above chance, $t(99) = 6.95$, $p < .001$, 95% CI [63.7%, 74.7%]. HITL accuracy remained significantly below chance, $t(99) = -2.85$, $p = .005$, 95% CI [41.2%, 48.4%].

Paired comparisons indicated that attribution accuracy for traditional patterns increased significantly from 51.8% before the intervention to 69.2% after the intervention, a mean increase of 17.4 percentage points, $t(99) = 4.83$, $p < .001$, 95% CI of the mean change [10.3%, 24.5%], Cohen's $dz = 0.48$. The increases for HOTL patterns, from 53.2% to 57.6%, $t(99) = 1.63$, $p = .107$, and HITL patterns, from 41.4% to 44.8%, $t(99) = 1.11$, $p = .270$, were not statistically significant. These findings suggest that the intervention primarily improved recognition of traditional cultural patterns rather than producing a general improvement in creator attribution across all generation modes.

Relationship between attribution accuracy and aesthetic evaluation

Aesthetic evaluations are not determined by beauty ratings alone. When evaluating AI-generated works, participants may be influenced by algorithm aversion. Addressing Research Question 4, which examines the relationship between attribution accuracy and aesthetic evaluation before and after the learning intervention, Figures 8 and 9 show that beauty scores for traditional ethnic patterns were positively correlated with attribution accuracy. This indicates that participants tended to identify patterns they perceived as beautiful as traditional ethnic patterns. By contrast, aesthetic evaluations of HITL and HOTL patterns were negatively correlated with attribution accuracy: the higher the beauty score for an AI-generated pattern, the lower the attribution accuracy. This suggests that participants' aesthetic preferences were shaped not only by the actual creator identity but also by their attribution of authorship. Participants generally tended to believe that aesthetically pleasing patterns were more likely to have been created by humans. However, this relationship became weaker after the learning intervention. Unlike previous studies [36], participants assigned higher beauty ratings to patterns attributed to human creation both before and after the intervention.

Discussion

Principal findings

With the development of artificial intelligence and generative art, research on how people experience machine-generated artworks has grown rapidly [4]. Most studies have focused on preferences for AI-generated works, biases toward AI, and participants' ability to identify authorship. Less attention has been paid to the dynamic nature of aesthetic cognition and to how cognitive shifts influence aesthetic judgment. This study used a learning intervention to investigate how participants evaluated AI-generated patterns and which cognitive factors shaped these evaluations, focusing on both aesthetic perception and creator attribution.

At the level of aesthetic perception, the final aesthetic ratings for all three types of patterns increased to varying degrees. After participants received learning interventions on Tujia brocade culture and AIGC principles, the weighting system of aesthetic evaluation dimensions changed, with increased importance of meaning, composition, and the peak-shift effect. This pattern can be explained through the dynamic processing mechanisms of aesthetic cognition: according to Leder's aesthetic processing model, learning strengthens top-down conceptual integration through cultural knowledge input and improves symbolic decoding ability. Although traditional Xilankapu patterns vary in composition and visual features, their forms follow recognizable cultural patterns, indicating that participants' aesthetic perception shifted from sensory experience toward interpretation of deeper meaning. The disappearance of statistical significance in the familiarity dimension (p = 0.318) is consistent with a possible cognitive-decoupling account [37]: before the intervention, intuition may

have produced a false sense of familiarity; after cultural knowledge was activated, participants re-anchored familiarity to symbolic and cultural identity. This mechanism helped them distinguish visual similarity from cultural authenticity. Based on the peak-shift effect [11], participants may have constructed a specialized category prototype for “Tujia brocade” through cultural knowledge input. This prototype included typical formal characteristics, such as specific motifs and compositional rules, as well as cultural connotations. These features helped participants perceive and recognize the stimuli and may explain the increase in emotional response scores in the aesthetic evaluations: when the anticipated emotional state of an aesthetic experience is positive, aesthetic processing is more likely to produce a positive evaluation.

By comparing SHAP values before and after the learning intervention, this study found changes in the plasticity of aesthetic cognition. Before the intervention, participants’ aesthetic judgments relied mainly on overall emotional resonance and features such as expressiveness and vividness. The evaluation system was therefore oriented toward identifying visual strengths. After the intervention, aesthetic evaluation placed greater emphasis on rational analysis and meaning interpretation. In particular, typicality and meaning became more important evaluation dimensions. The distribution of SHAP values also shifted from a mainly positive distribution before the intervention to a bidirectional distribution with both positive and negative values. This suggests that participants developed a more complex evaluative framework that integrated appreciation and criticism, enabling them to make more contextualized judgments based on cultural and artistic knowledge. This transformation is consistent with prototype theory [38] and with neuroaesthetic accounts of knowledge-based aesthetic processing.

At the level of creator attribution, participants did not always accurately distinguish human-created from AI-generated works. Attribution accuracy increased numerically for all three pattern types, although a statistically significant improvement was observed only for traditional patterns. Notably, attribution accuracy for HITL patterns remained significantly below the chance level both before and after the intervention. Rather than reflecting random guessing, this systematic below-chance performance indicates that participants consistently classified human–AI collaborative patterns as human-created. One possible explanation is that human refinement reduced perceptible algorithmic cues, making collaborative outputs appear more consistent with human craftsmanship.

Beauty scores for traditional patterns were positively associated with attribution accuracy, whereas HOTL and HITL patterns showed negative associations. This pattern is consistent with an algorithm-aversion account: participants tended to attribute aesthetically pleasing works to human rather than AI creators. However, the association between creator attribution and aesthetic evaluation appeared descriptively weaker after the intervention. In light of Boden’s discussion of authenticity in computer art [12], this pattern does not necessarily indicate that participants recognized AI as an independent creative agent. Rather, structured learning may have helped participants understand aesthetic evaluation criteria and AIGC generation principles, making the algorithmic process less opaque and encouraging them to view AI more as a creative tool than as either an autonomous creator or a threat to cultural authenticity.

Interpretation of findings

Through a systematic learning intervention, this study revealed the plasticity of participants’ culture-specific aesthetic cognition and its underlying mechanisms. The preliminary results indicate that short-term interventions in semiotics and cultural context not only increased participants’ aesthetic acceptance of the patterns but also restructured the weighting system of aesthetic evaluation dimensions.

At the level of aesthetic perception, the learning intervention shifted participants’ aesthetic decision-making from sensory-driven to concept-driven evaluation. Before the intervention, participants mainly relied on superficial perceptual features such as expressiveness and vividness for intuitive judgments, and the aesthetic process showed strong emotional and holistic tendencies. After the intervention, dimensions closely related to cultural prototypes and symbolic connotations, such as typicality, meaning, and harmony, gained greater weight. This indicates that participants began to rely more on cultural knowledge and symbolic decoding when making aesthetic judgments. The shift is consistent with the theoretical expectation of enhanced top-down cognitive processing in Leder’s aesthetic processing model [6] and supports the central role of the interpretant in meaning construction within Peircean semiotics. Notably, familiarity became statistically insignificant after the intervention, suggesting that learning helped participants distinguish genuine cultural familiarity from superficial visual déjà vu, reflecting cognitive decoupling and enhanced aesthetic specialization.

HITL patterns received the highest beauty ratings before the intervention. After the intervention, however, traditional ethnic patterns outperformed HITL patterns on core dimensions such as style, typicality, and meaning. This dynamic reversal suggests that once participants acquired greater domain-specific cultural knowledge through the learning intervention, novelty in visual form became less central, while authenticity, including human history, religious beliefs, and emotional meaning, became more important. This indicates that although AI can reorganize visual features effectively, human creativity and historical accumulation remain important for conveying deeper cultural meanings. At the level of creator attribution, the association between creator attribution and aesthetic evaluation appeared descriptively weaker after the intervention, although this change was not formally tested for statistical significance. The remaining association suggests that aesthetic evaluation may still be influenced by both visual characteristics and assumptions about creator identity. Attribution accuracy for traditional patterns increased significantly from 51.8% to 69.2% after the intervention, representing a mean increase of 17.4 percentage points. By contrast, the modest increases observed for HOTL and HITL patterns did not reach statistical significance. Participants may have learned to identify common features of original patterns, and cultural knowledge improved their sensitivity to the formal language of traditional patterns. The descriptively weaker negative association between aesthetic evaluation and attribution accuracy for AI-generated patterns, particularly HITL patterns, is consistent with the possibility that educational interventions encourage participants to attend more closely to cultural and formal characteristics rather than relying primarily on assumptions about creator identity.

Theoretical contributions

At the theoretical level, this study advances the exploration of aesthetic mechanisms of cultural symbols in neuroaesthetics. First, it provides empirical evidence from a cultural context for the influence of top-down cognitive control on aesthetic experience. By introducing semiotics into Leder’s aesthetic

processing model, the study shows that aesthetic evaluation is not passive perception, but a dynamic interaction between perceptual input and internal knowledge frameworks.

Second, the results reveal how aesthetic plasticity may operate in the domain of cultural cognition. After learning, participants' aesthetic judgments shifted from early visual features such as vividness and expressiveness toward higher-order cognitive features such as typicality and meaning, which depend on cultural knowledge. This shift from sensory-driven to concept-driven evaluation may reflect greater reliance on higher-order conceptual processing and meaning integration. However, because no neural measures were collected, the present findings cannot determine which neural systems underlie this cognitive change.

Practical implications

Beyond its theoretical contributions, this study offers practical implications for cultural heritage preservation, public education, and intelligent design workflows in the AIGC era. First, for the digital dissemination of traditional culture and public education, the findings suggest that improving AI literacy and cultural decoding ability can help reduce technology-related bias and support cultural understanding. Simply improving the visual quality of algorithmic generation may not be sufficient. Cultural institutions may need structured semiotic knowledge interventions to guide audiences in evaluating generated content through a rational and critical lens. Second, for intelligent design workflows, the findings help clarify the roles of humans and AI in cultural creation. When audiences understand cultural meaning, they place greater emphasis on authenticity. Therefore, in the digital generation of cultural heritage, AI may be better positioned as a tool for handling representational elements such as formal variation, color derivation, and compositional extrapolation, whereas human designers should retain control over interpretive elements such as cultural significance, typological features, and emotional connections. This division of roles may help prevent generated content from becoming culturally hollow [39].

Limitations and future research

The current study has several limitations and suggests directions for future research. Due to cost constraints, the sample size was limited. Based on affordance theory [40], aesthetic experience emerges from the interaction between the perceiver's abilities and the object's attributes. Therefore, participants from different backgrounds may respond to the same intervention in different ways. Future research could include wider age groups and professional backgrounds to account for varying levels of learning ability and exposure to the arts.

This study also relied on explicit self-report scales, which may introduce social desirability bias. After completing a 30-day course on ethnic culture, participants may have assigned higher emotional ratings to traditional ethnic patterns because they wanted to protect cultural heritage or conform to perceived experimental expectations. Future research should incorporate implicit measurement techniques, such as eye-tracking, pupil-change analysis, or the Implicit Association Test, to cross-validate the scale results.

In addition, although changes in the weighting of aesthetic dimensions after the learning intervention reveal psychological mechanisms related to cognitive framework shifts and aesthetic expertise, the study could not distinguish the relative contributions of cultural background learning and generative-principle learning because participants received both interventions. Future research could draw on dual-process theory [41], which distinguishes between rapid, intuitive heuristic processing (System 1) and slower analytical processing (System 2). This framework could help examine whether cultural learning and generative-principle learning influence different processing systems and thereby clarify the cognitive mechanisms underlying aesthetic change.

Future studies could employ neuroimaging techniques such as fMRI and EEG to directly examine the neural mechanisms underlying the shift from sensory-driven to concept-driven processing. In particular, researchers could investigate changes in activation patterns and functional connectivity among the occipital visual cortex, prefrontal cortex, and default mode network before and after learning interventions, thereby providing physiological evidence for the reshaping of aesthetic cognition.

Longitudinal research is also needed to examine the long-term stability and transferability of such changes in aesthetic cognition and neural patterns. For example, delayed tests could be conducted several months after the intervention, or the cultural schemas established by participants could be applied to other symbol types within the same system to assess the persistence and generalizability of learning effects. Such studies would deepen understanding of culture-specific aesthetic learning at the cognitive and neural levels, support the development of an integrated culture–cognition–neuroscience model, and provide evidence for the contemporary transformation and dissemination of cultural heritage.

Conclusions

Artificial intelligence is increasingly capable of producing outputs that resemble human creative activity [39]. This has prompted researchers to investigate how people perceive machine-generated artworks [3, 26, 28, 36]. Building on this research, the present study integrated aesthetic processing theory and semiotics to examine participants' judgments of traditional ethnic patterns and AI-generated patterns before and after a learning intervention. It focused on two dimensions, aesthetic perception and creator attribution, to clarify how aesthetic evaluation changes in the context of artificial intelligence. The results reveal the dynamic plasticity of participants' aesthetic cognition regarding Tujia brocade patterns and address the four research questions.

First, the post-intervention assessment showed higher ratings on several aesthetic dimensions together with a different weighting pattern of aesthetic evaluation, indicating a shift from predominantly sensory-driven toward more concept-driven judgments. Second, participants evaluated HOTL patterns more critically after the intervention when these patterns lacked culturally meaningful characteristics. This finding challenges the assumption that learning simply increases tolerance toward AI-generated art. Third, the relationship between creator attribution and aesthetic evaluation became descriptively weaker following the intervention, suggesting that increased cultural knowledge may reduce reliance on creator identity when evaluating AI-generated patterns. Fourth, among the generative modes involving AI, human–AI collaboration (HITL) consistently outperformed human-out-of-the-loop generation (HOTL).

However, after the learning intervention, traditional ethnic patterns regained the highest ratings in core cultural dimensions such as typicality and meaning. These findings suggest that, in cultural innovation, human cultural inheritance and semantic interpretation remain central.

These conclusions deepen understanding of how top-down processing influences aesthetic experience in neuroaesthetics and offer a practical path for supporting the creative transformation and contemporary development of traditional cultural symbols in the AI era.

Abbreviations

AI: artificial intelligence

AIGC: artificial intelligence-generated content

HITL: human-in-the-loop

HOTL: human-out-of-the-loop

LoRA: Low-Rank Adaptation

SHAP: SHapley Additive exPlanations

Declarations

Clinical trial registration

Not applicable. This study was not a clinical trial. It did not involve a medical intervention, clinical treatment, diagnosis, prevention, or health-related outcome. The intervention consisted of structured learning materials on Xilankapu culture, semiotic interpretation, and AI-based pattern generation, and the outcomes were aesthetic ratings and creator-attribution judgments. Therefore, clinical trial registration and CONSORT reporting are not applicable.

Ethics approval and consent to participate

The study protocol was reviewed and approved by the Ethics Committee of Changsha University of Science and Technology, as documented in the ethical review approval form dated 1 January 2025. The ethics approval document states that the project complied with ethical principles and was approved for conduct. All procedures were performed in accordance with the Declaration of Helsinki and relevant institutional guidelines. All participants provided electronic informed consent before participating. Participation was voluntary, no directly identifying personal information was retained, and all data were de-identified before analysis.

Consent for publication

Not applicable.

Data availability

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

The authors received no specific funding for this work.

Authors' contributions

SZ was responsible for conceptualization, methodology, investigation, and drafting the manuscript. JH contributed to data collection, data curation, and validation. ZC contributed to data collection, visualization, and formal analysis. YH contributed to formal analysis, visualization, and manuscript revision. JS supervised the study, contributed to project administration and interpretation of the results, and reviewed and edited the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

References

1. Cetinic, E. & She, J. Understanding and creating art with AI: review and outlook. arXiv [Preprint]. (2021). 10.48550/arXiv.2102.09109
2. Daniele, A. & Song, Y. Z. AI + art = human. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. New York: ACM; pp. 155–61. (2019). 10.1145/3306618.3314233
3. Gangadharbatla, H. The role of AI attribution knowledge in the evaluation of artwork. *Empir. Stud. Arts*. **40** (2), 125–142. 10.1177/0276237421994697 (2022).
4. Samo, A. & Highhouse, S. Artificial intelligence and art: identifying the aesthetic judgment factors that distinguish human- and machine-generated artwork. *Psychol. Aesthet. Creat Arts*. 10.1037/aca0000570 (2023).
5. Hitsuwari, J., Ueda, Y., Yun, W. & Nomura, M. Does human–AI collaboration lead to more creative art? Aesthetic evaluation of human-made and AI-generated haiku poetry. *Comput. Hum. Behav.* **139**, 107502. 10.1016/j.chb.2022.107502 (2023).
6. Leder, H., Belke, B., Oeberst, A. & Augustin, D. A model of aesthetic appreciation and aesthetic judgments. *Br. J. Psychol.* **95** (4), 489–508. 10.1348/0007126042369811 (2004).
7. Pearce, M. T. et al. Neuroaesthetics: the cognitive neuroscience of aesthetic experience. *Perspect. Psychol. Sci.* **11** (2), 265–279. 10.1177/1745691615621274 (2016).
8. Köbis, N. & Mossink, L. D. Artificial intelligence versus Maya Angelou: experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Comput. Hum. Behav.* **114**, 106553. 10.1016/j.chb.2020.106553 (2021).
9. Saharia, C. et al. Photorealistic text-to-image diffusion models with deep language understanding. *Adv. Neural. Inf. Process. Syst.* **35**, 36479–36494 (2022).
10. Coeckelbergh, M. Can machines create art? *Philos. Technol.* **30** (3), 285–303. 10.1007/s13347-016-0231-5 (2017).
11. Ramachandran, V. S. & Hirstein, W. The science of art: a neurological theory of aesthetic experience. *J. Conscious. Stud.* **6** (6–7), 15–51 (1999).
12. Boden, M. A. Authenticity and computer art. *Digit. Creat*. **18** (1), 3–10. 10.1080/14626260701252285 (2007).
13. Llano, M. T. et al. Explainable computational creativity. arXiv [Preprint]. (2022). 10.48550/arXiv.2205.05682
14. Mazzone, M. & Elgammal, A. Art, creativity, and the potential of artificial intelligence. *Arts* **8** (1), 26. 10.3390/arts8010026 (2019).
15. McCormack, J., Gifford, T. & Hutchings, P. Autonomy, authenticity, authorship and intention in computer-generated art. *arXiv [Preprint]*. 10.48550/arXiv.1903.02166 (2019).
16. Bellaiche, L. et al. Humans versus AI: whether and why we prefer human-created compared to AI-created artwork. *Cogn. Res. Princ Implic*. **8**, 42. 10.1186/s41235-023-00499-6 (2023).
17. Leder, H. Next steps in neuroaesthetics: which processes and processing stages to study? *Psychol. Aesthet. Creat Arts*. **7** (1), 27–37. 10.1037/a0031585 (2013).
18. Chatterjee, A. Neuroaesthetics: a coming of age story. *J. Cogn. Neurosci.* **23** (1), 53–62. 10.1162/jocn.2010.21457 (2011).
19. Leder, H. & Nadal, M. Ten years of a model of aesthetic appreciation and aesthetic judgments: the aesthetic episode—developments and challenges in empirical aesthetics. *Br. J. Psychol.* **105** (4), 443–464. 10.1111/bjop.12084 (2014).
20. Deni, M. & Zingale, S. Semiotics in design education: semiotics by design. *Des. J.* **20** (Suppl 1), S1293–S1303. 10.1080/14606925.2017.1352658 (2017).
21. Bommasani, R. et al. On the opportunities and risks of foundation models. *arXiv [Preprint]*. 10.48550/arXiv.2108.07258 (2022).
22. Colton, S. & Wiggins, G. A. Computational creativity: the final frontier? In: ECAI 2012. Amsterdam: IOS; 21–26. doi:10.3233/978-1-61499-098-7-21. (2012).
23. Lindell, A. K. & Mueller, J. Can science account for taste? Psychological insights into art appreciation. *J. Cogn. Psychol.* **23** (4), 453–475. 10.1080/20445911.2011.539556 (2011).
24. Pelowski, M., Markey, P. S., Lauring, J. O. & Leder, H. Visualizing the impact of art: an update and comparison of current psychological models of art experience. *Front. Hum. Neurosci.* **10**, 160. 10.3389/fnhum.2016.00160 (2016).
25. Brielmann, A. A., Nuzzo, A. & Pelli, D. G. Beauty, the feeling. *Acta Psychol.* **219**, 103365. 10.1016/j.actpsy.2021.103365 (2021).
26. Ragot, M., Martin, N. & Cojean, S. AI-generated vs. human artworks: a perception bias towards artificial intelligence? In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. New York: ACM; pp. 1–10. (2020). 10.1145/3334480.3382892
27. Luo, X. D., Yang, L. L. & Zhang, X. Y. Extraction and regeneration of perceptual factors in Tujia brocade patterns. *Wool. Text. Sci. Technol.* **51** (8), 54–60 (2023).
28. Chamberlain, R., Mullin, C., Scheerlinck, B. & Wagemans, J. Putting the art in artificial: aesthetic responses to computer-generated art. *Psychol. Aesthet. Creat Arts*. **12** (2), 177–192. 10.1037/aca0000136 (2018).
29. Sato, T. The experienced readers' liking for haiku and their personality. *Bull. Hachinohe Inst. Technol.* **26**, 139–147 (2007).
30. Ho, J., Jain, A. & Abbeel, P. Denoising diffusion probabilistic models. *Adv. Neural. Inf. Process. Syst.* **33**, 6840–6851 (2020).
31. Rombach, R., Blattmann, A., Lorenz, D., Esser, P. & Ommer, B. High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10684–95. (2022). 10.1109/CVPR52688.2022.01042

32. Specker, E. et al. The Vienna Art Interest and Art Knowledge Questionnaire (VAIAK): a unified and validated measure of art interest and art knowledge. *Psychol. Aesthet. Creat Arts*. **14** (2), 172–185. 10.1037/aca0000205 (2020).
33. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67** (1), 1–48. 10.18637/jss.v067.i01 (2015).
34. Lundberg, S. M. & Lee, S. I. A unified approach to interpreting model predictions. *Adv. Neural. Inf. Process. Syst.* **30**, 4765–4774 (2017).
35. Chen, T. & Guestrin, C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM; pp. 785–94. (2016). 10.1145/2939672.2939785
36. Hong, J. W. & Curran, N. M. Artificial intelligence, artists, and art: attitudes toward artwork produced by humans vs. artificial intelligence. *ACM Trans. Multimed Comput. Commun. Appl.* **15** (2s), 1–16. 10.1145/3326337 (2019).
37. Stanovich, K. E. & West, R. F. Individual differences in reasoning: implications for the rationality debate? *Behav. Brain Sci.* **23** (5), 645–665. 10.1017/S0140525X00003435 (2000).
38. Rosch, E. Cognitive representations of semantic categories. *J. Exp. Psychol. Gen.* **104** (3), 192–233. 10.1037/0096-3445.104.3.192 (1975).
39. Rahwan, I. et al. *Mach. Behav. Nat.* ;**568**(7753):477–486. doi:10.1038/s41586-019-1138-y. (2019).
40. Gibson, J. J. *The ecological approach to visual perception: classic edition* (Psychology, 2014). 10.4324/9781315740218
41. Kahneman, D. *Thinking, fast and slow* (Farrar, Straus and Giroux, 2011).

Figures

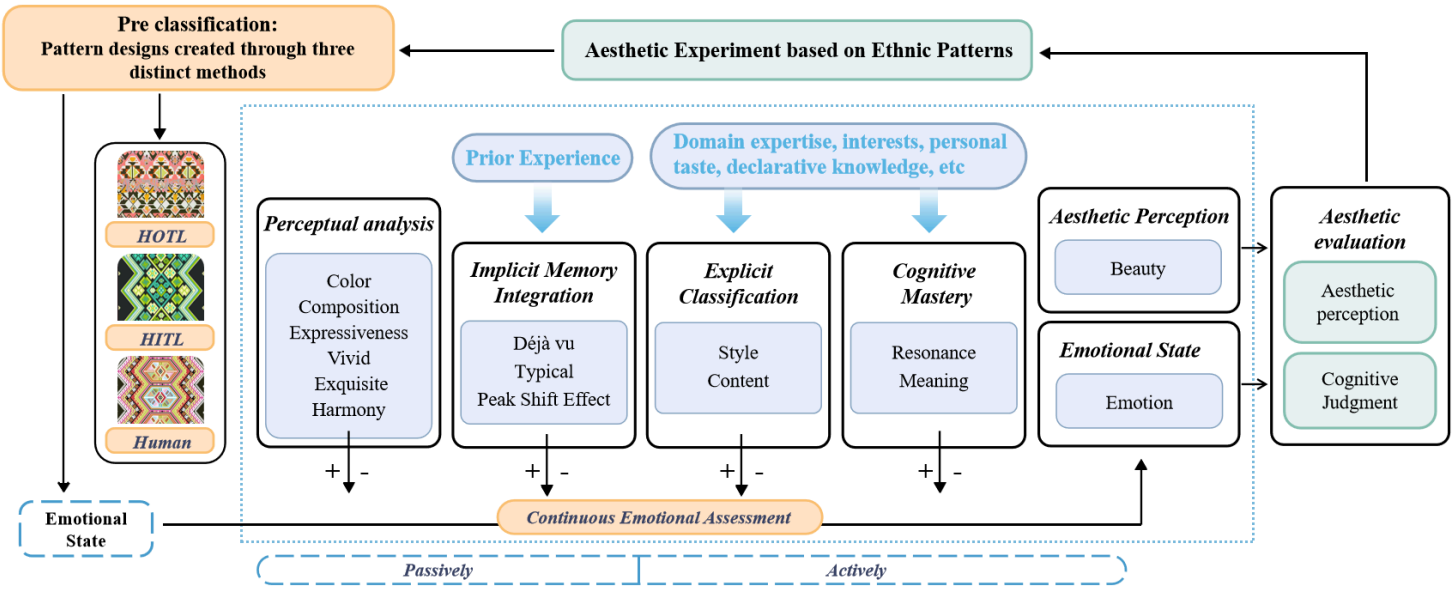


Figure 1

Model of the Aesthetic Cognitive Process

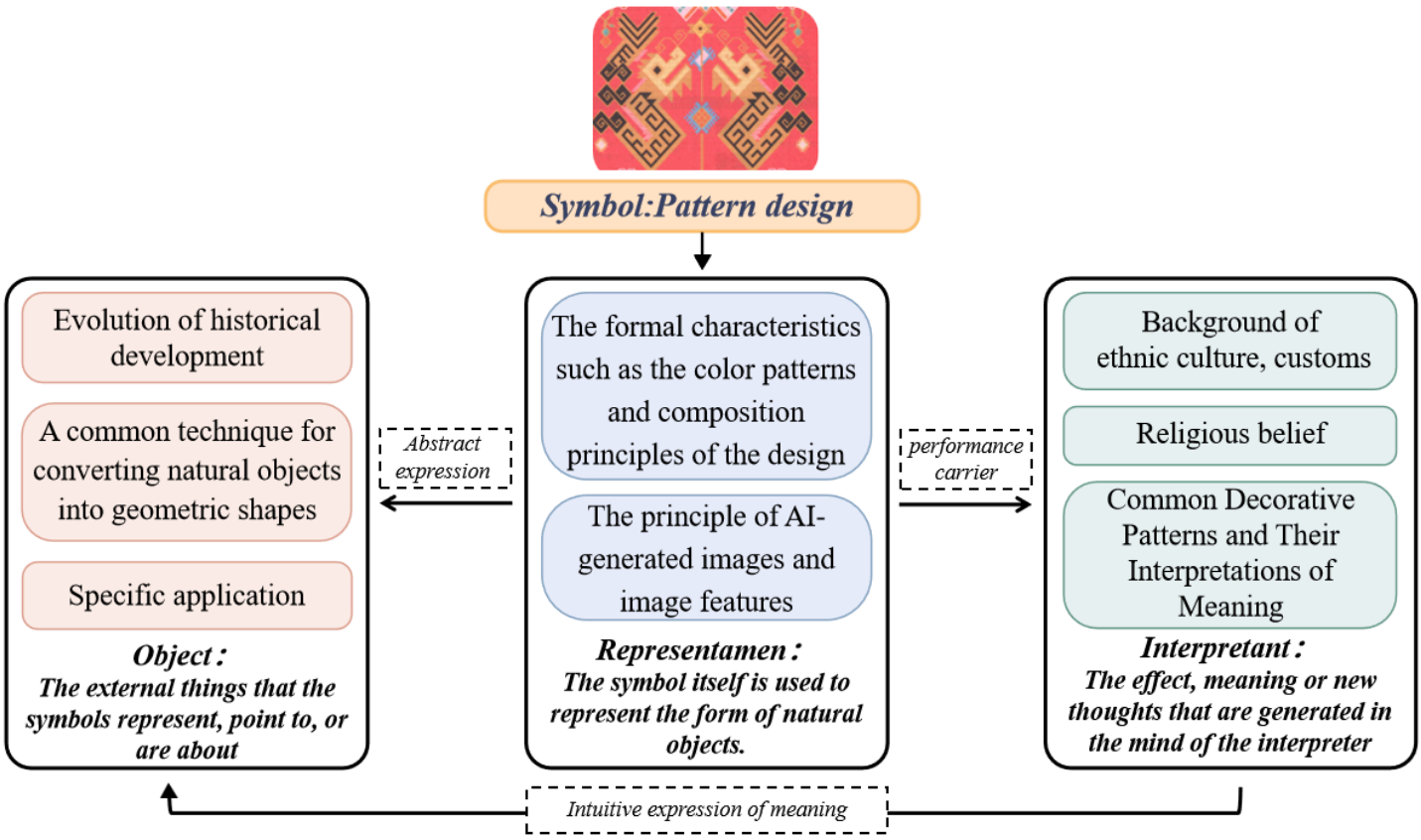


Figure 2

Learning intervention model

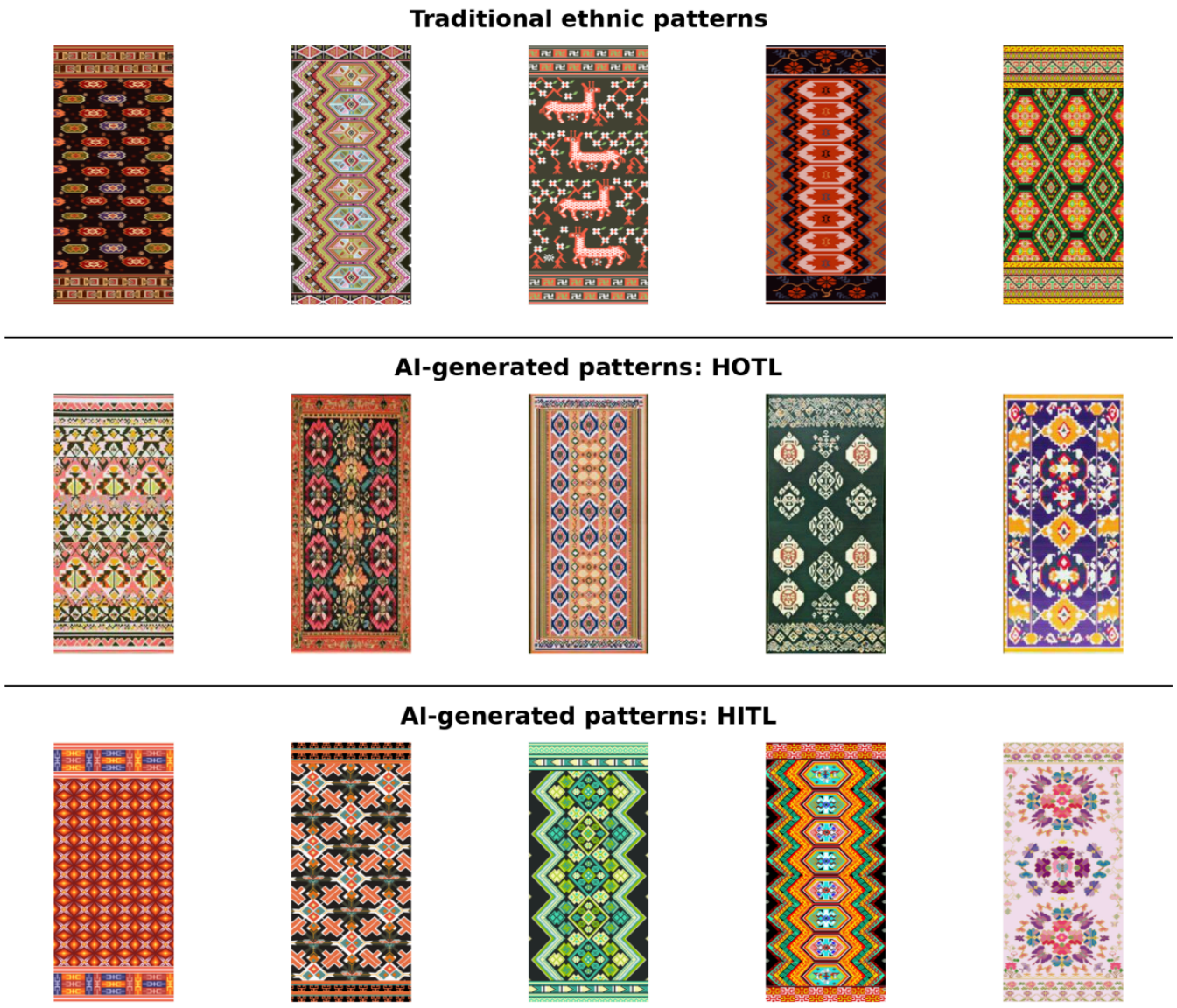


Figure 3

Examples of traditional, HOTL, and HITL pattern stimuli

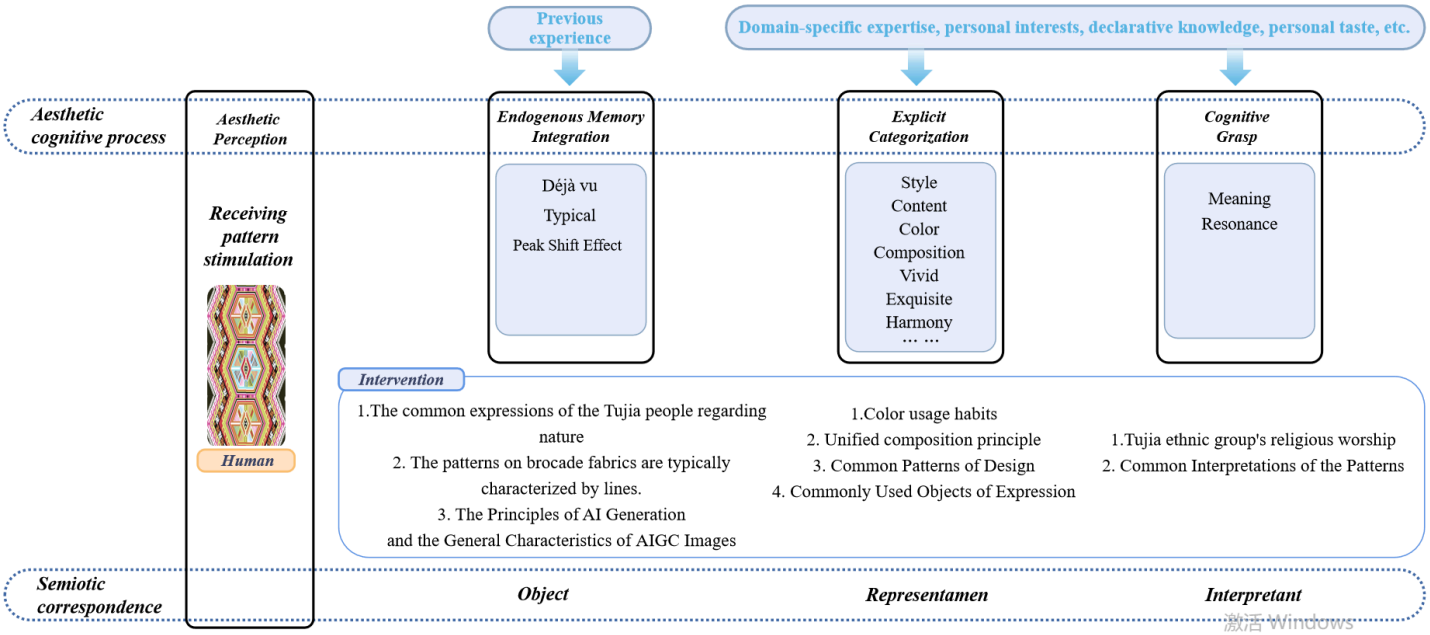


Figure 4

Correspondence between the aesthetic cognitive process and Peirce's triadic theory of signs

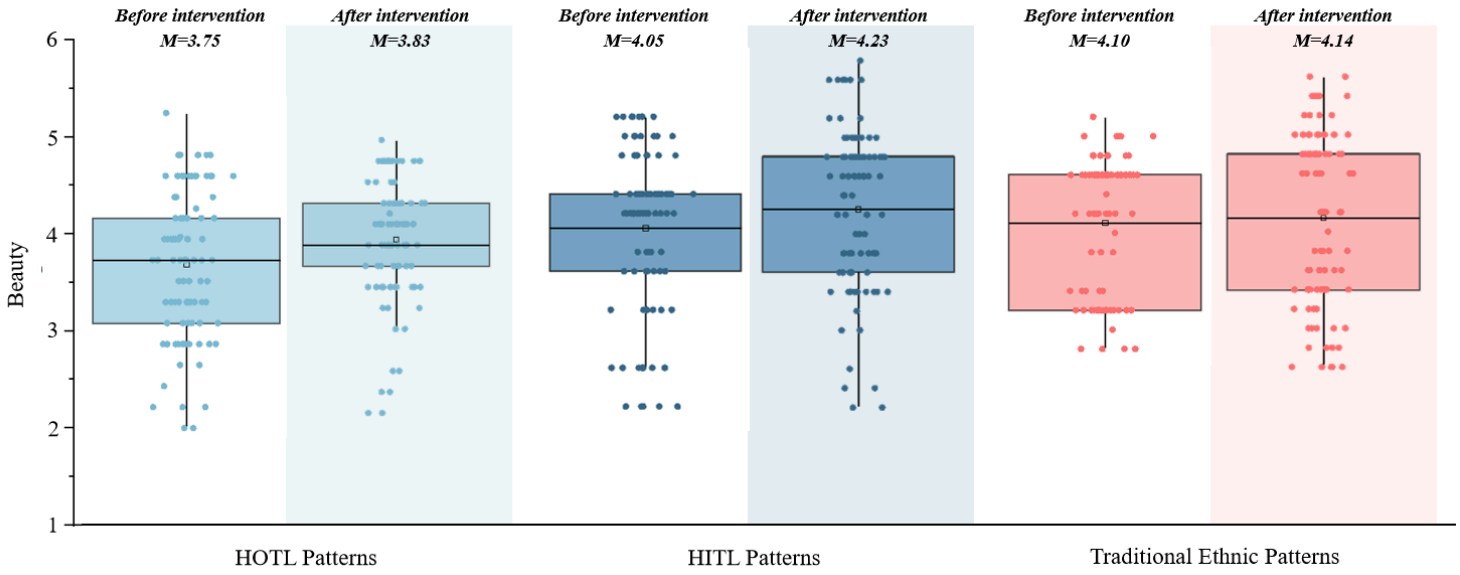


Figure 5

Aesthetic ratings by pattern-generation mode before and after the intervention

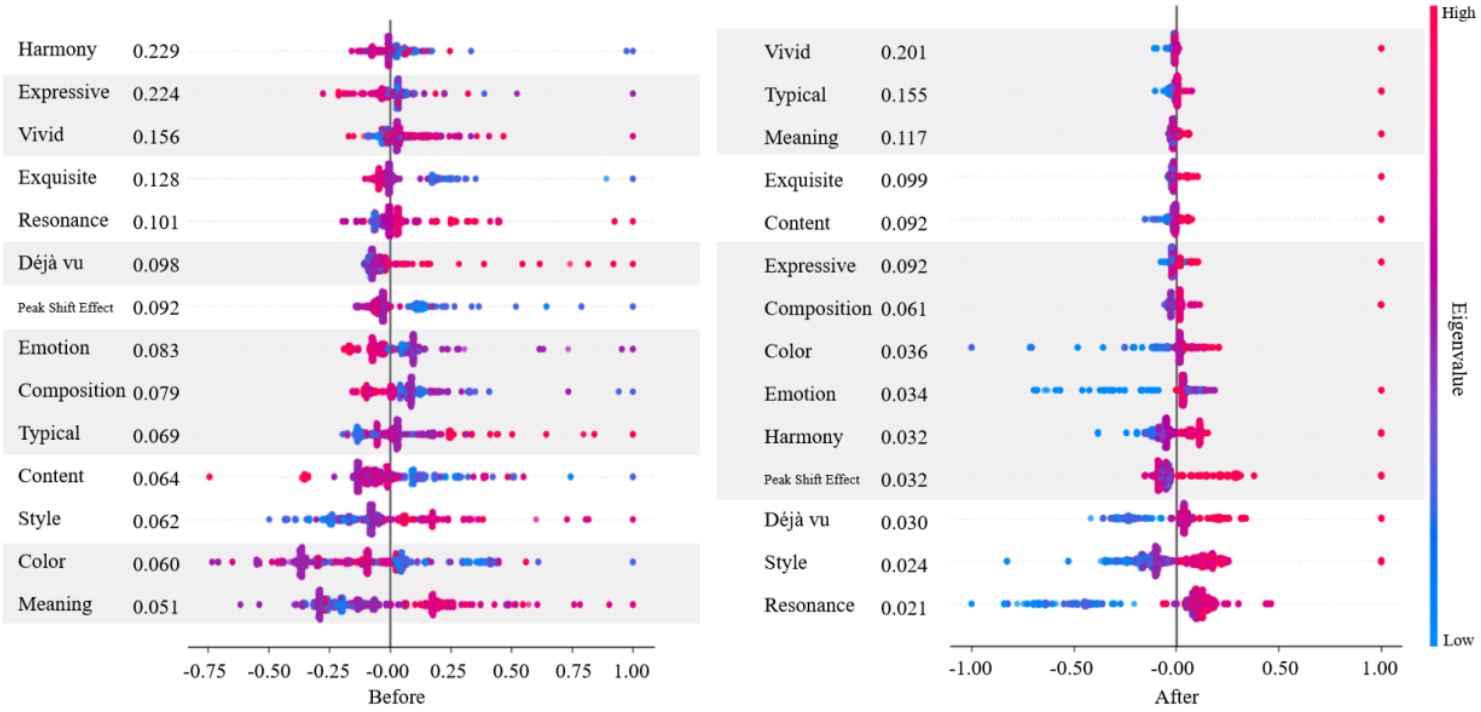


Figure 6

SHAP summary plots of aesthetic dimensions: (a) before and (b) after the intervention

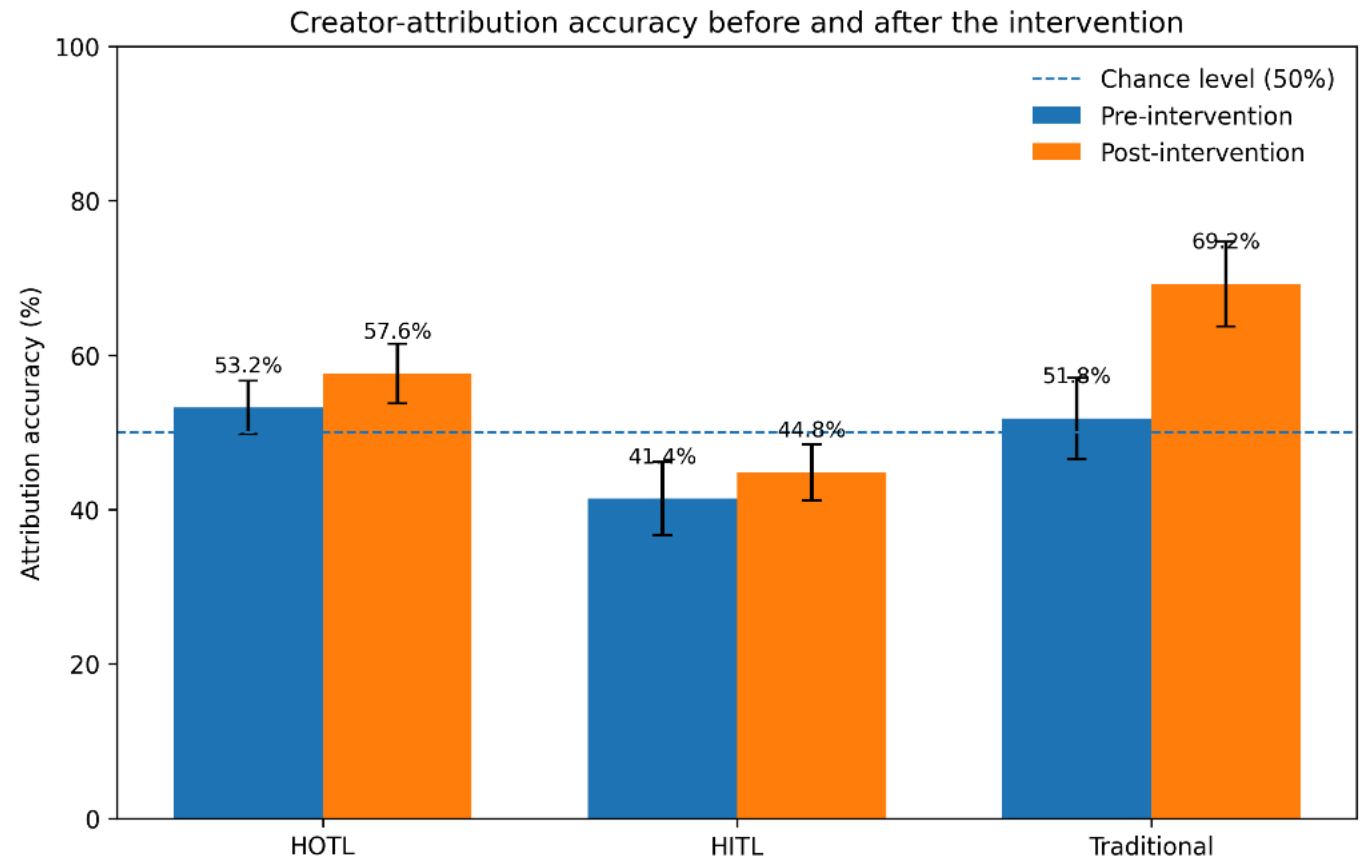


Figure 7

Creator-attribution accuracy before and after the learning intervention. Error bars represent 95% confidence intervals, and the dashed horizontal line indicates the chance level (50%).

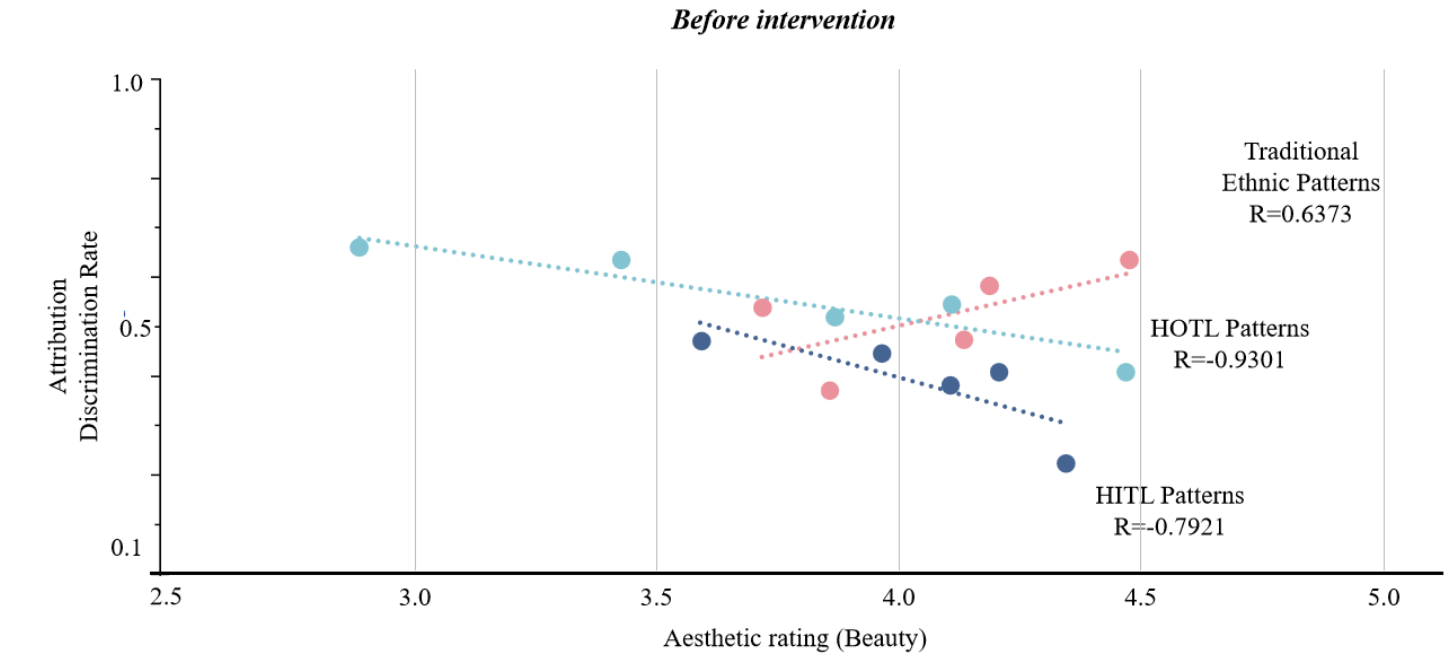


Figure 8

Relationship between aesthetic ratings and attribution accuracy (pre-intervention)

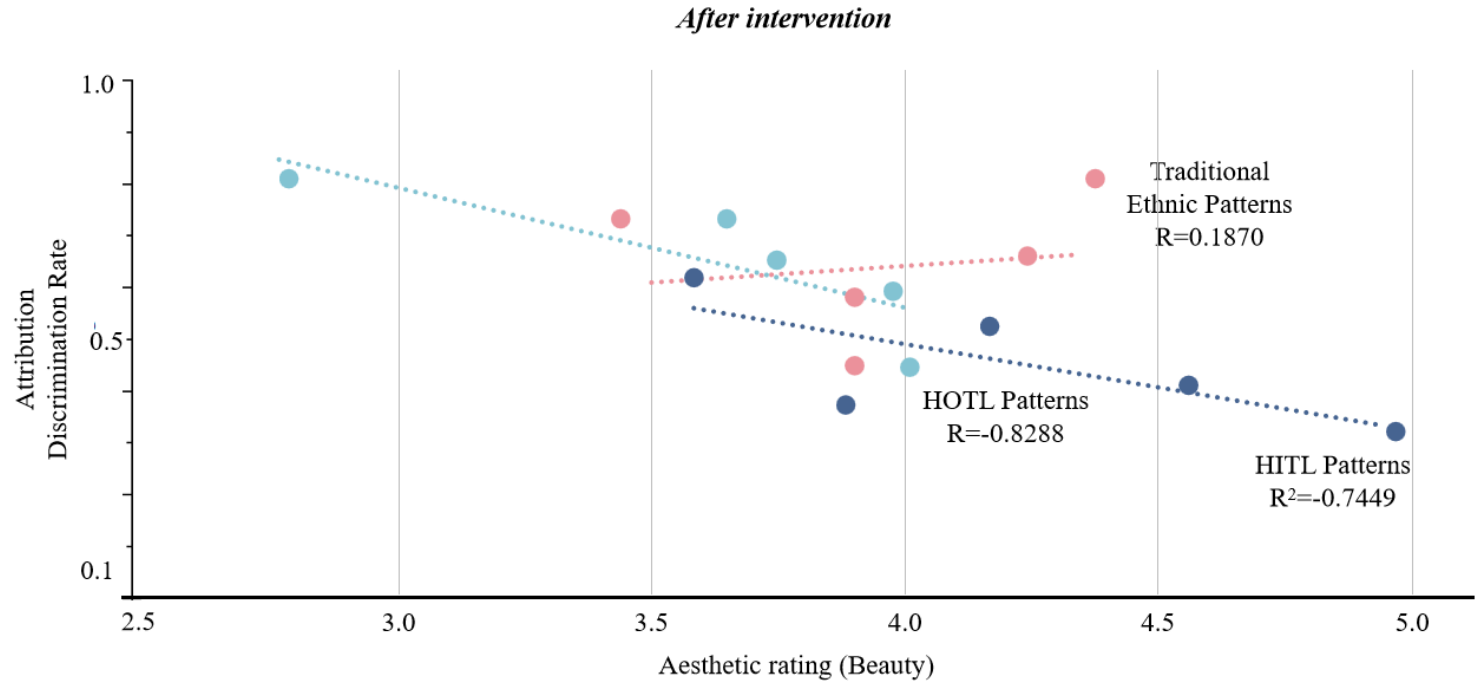


Figure 9

Relationship between aesthetic ratings and attribution accuracy (post-intervention)