

Appendix

Appendix A: Experimental Procedures

1) Sampling Method

Stratified random sampling was employed:

Sample Size: $n = 481$

Age Distribution: 18-45 years ($M = 26.66$)

Gender Ratio: 220 males (45.7%), 261 females (54.3%)

Geographic Coverage: All 31 provinces, municipalities, and autonomous regions nationwide

Sampling Strategy: Stratified by age, gender, and region to ensure sample representativeness

2) Research Instruments

Drawing on the MakeBronze interactive system research , a four-dimensional assessment questionnaire was designed using a screening approach, where only participants who answered "Yes" to Q1 continued with subsequent questions.

Four-Dimensional Assessment Questionnaire:

Q1: Basic Cognitive Assessment (Binary choice)

"Are you familiar with Cizhou kiln?" (Yes/No)

Q2: Communication Channel Analysis (Multiple choice)

"Through which means did you learn about Cizhou kiln?"

Video resources

Print materials

Exhibitions

Lectures

Interactive participation

Q3: Content Understanding Depth (Multiple choice)

Assessing mastery of various aspects of Cizhou kiln:

Historical background

Cultural value

Type characteristics

Representative works

Q4: Sustained Engagement Intention (5-point Likert scale)

Assessing willingness to learn more

Q5: Is there anything else you would like to say about Cizhou kiln? (Open-ended question)

3) Data Collection Process

Time Period: October 2024 to November 2024

Collection Method: Online questionnaire

Valid Questionnaires: 481

Quality Control: Attention check questions were included, and samples with abnormal completion times were excluded

A1. Recruitment Channels

- **Online channels (68.2%):** Recruitment announcements distributed through mainstream Chinese social media platforms (WeChat, Weibo, Xiaohongshu/Little Red Book), employing snowball sampling to expand coverage
- **Offline channels (31.8%):** On-site recruitment at 5 universities in different regions (Beijing, Shanghai, Guangzhou, Chengdu, Xi'an), supplemented by recruitment posters at community cultural centers
- **Partner organizations:** Collaboration with 3 youth cultural organizations to promote survey links

A2. Stratified Sampling Design

- Stratification by **geographic region** (North China, East China, South China, Southwest, Northwest, and 2 additional regions), ensuring coverage of all 31 provincial-level administrative divisions in China
- Stratification by **age group** (18-24, 25-34, 35-45 years), with quota proportions of 4:3:3 respectively
- **Gender balance** quota (target ratio 1:1, actual achievement 45.7%:54.3%)

A3. Inclusion Criteria

1. Chinese citizens aged 18-45 years
2. Adequate Chinese reading comprehension ability
3. Voluntary participation with signed informed consent

A4. Exclusion Criteria

1. Professionals in ceramic arts or cultural heritage conservation (to avoid professional bias)
2. Participants who had engaged in similar cultural heritage research within the past 6 months
3. Data with questionnaire completion rates below 80% or exhibiting obvious careless response patterns

A5. Incentive Measures

To enhance participation rates and data quality, this research implemented the following incentive mechanisms:

Economic incentives:

- Participants completing the questionnaire received a **10 RMB electronic red packet** (approximately 1.5 USD)
- Twenty participants engaging in follow-up semi-structured interviews received an additional **50 RMB compensation** (approximately 7.5 USD)
- Total incentive expenditure: 5,810 RMB

Non-economic incentives:

- All participants received an electronic handbook titled "Introduction to Cizhou Kiln Culture" to foster cultural learning interest

Incentive effectiveness evaluation:

- Questionnaire completion rate: **94.3%** (510 initially distributed, 481 valid responses collected)
- Data quality verification: **96.7%** passed attention check questions

A6. Experimental Timeline

- Recruitment period: September-October 2024(6 weeks)
- Data collection: October-November 2024 (8 weeks)
- Data cleaning and analysis: November-December 2024 (6 weeks)

Appendix B: Complete Assessment Tools

B1. System Usability Scale (SUS)

This scale contains 10 five-point Likert questions using a mixed design of positive and negative items to eliminate response bias. Scoring rules: positive items (score = raw value - 1), negative items (score = 5 - raw value), total score conversion formula: $\Sigma \text{score} \times 2.5$, range 0-100. Scores above 70 indicate good system usability.

1. I would like to use this system frequently
2. I found this system unnecessarily complex
3. I thought this system was easy to use
4. I think I would need support from a technical person to use this system
5. I found the various functions in this system were well integrated
6. I thought there was too much inconsistency in this system
7. I imagine most people would learn to use this system very quickly
8. I found this system very cumbersome to use
9. I felt very confident using this system
10. I needed to learn a lot of things before I could get going with this system

B2. Algorithm Performance Assessment

Includes both quantitative and qualitative indicators:

Quantitative Indicators:

FID Score (Fr chet Inception Distance): Evaluates quality and authenticity of generated content

PSNR (Peak Signal-to-Noise Ratio): Measures reconstruction quality

Training Time: Records model training efficiency

Memory Usage: Evaluates resource consumption

Qualitative Indicators:

Style Consistency: Evaluates conformity with Cizhou kiln characteristics

Detail Fidelity: Evaluates accuracy of pattern detail reproduction

Innovation Level: Evaluates design innovation

Overall Satisfaction: Evaluates overall satisfaction with the design

B3. Knowledge and Motivation Assessment

Questionnaire 1 contains 5 structured questions assessing knowledge mastery level. Questionnaire 2 uses a revised Cultural Participation Motivation Scale [43-44], with Cronbach's $\alpha = 0.919$.

Questionnaire 1:

SQ1: Recognition of Cizhou kiln's cultural heritage status

SQ2: Identification of Cizhou kiln ceramic product types

SQ3: Understanding of Cizhou kiln artistic characteristics

SQ4: Recognition of Cizhou kiln user groups

SQ5: Understanding of Cizhou kiln pattern symbolism

Questionnaire 2:

Learning Enjoyment: "This method makes learning about Cizhou kiln more enjoyable"

In-depth Interest: "This innovative method impressed me deeply, and I'm interested in learning more about Cizhou kiln culture"

Learning Motivation: "This method stimulated my motivation to learn about Cizhou kiln culture"

Comparative Preference: "Compared to other methods, this approach is more interesting to me"

MICSI Scale

The MICSI scale contains 18 seven-point rating items for evaluating human-machine creative collaboration, covering 5 core dimensions (enjoyment, exploration, expression, immersion, outcome value) and 4 interaction dimensions (communication, alignment, agency, partnership) plus an exploratory scale. Scores of 5 and above indicate positive experiences, with the agency dimension using reverse scoring. Note: In the original MICSI questions 15-18, we replaced the word "sketch" with "work" to better fit our research context.

Core dimensions assess users' subjective experiences

Interaction dimensions assess system collaboration effectiveness

Semi-structured interviews explore user needs in depth

Table 2: MICSI Scale

No.	Scale/Subscale	Question(s)
1	Enjoyment	"I would be happy to use this system or tool on a regular basis."
2		"I enjoyed using the system or tool."

3	Exploration	"It was easy for me to explore many different ideas, options, designs, or outcomes, using this system or tool."
4		"The system or tool was helpful in allowing me to track different ideas, outcomes, or possibilities"
5	Expresiveness	"I was able to be very creative while doing the activity inside this system or tool."
6		"The system or tool allowed me to be very expressive."
7	Immersiveness	"My attention was fully tuned to the activity, and I forgot about the system or tool that I was using."
8		"I became so absorbed in the activity that I forgot about the system or tool that I was using."
9	Worth	"I was satisfied with what I got out of the system or tool."
10		"What I was able to produce was worth the effort I had to exert to produce it."
11	Communication	"I was able to effectively communicate what I wanted to the system."
12	Alignment	"I was able to steer the system toward output that was aligned with my goals."
13	Agency	"At times, I felt that the system was steering me toward its own goals."
14	Partnership	"At times, it felt like the system and I were collaborating as equals."
15	Contribution	"I made the work" vs "The system made the work."
16	Satisfaction	"I'm very unsatisfied with the work" vs "I'm very satisfied with the work."
17	Surprise	"The work was what I was aiming for" vs "The work outcome was unexpected."
18	Novelty	"The work is very typical" vs "The work is very novel."

Appendix C: Participant Selection Criteria

Expert Selection Criteria:

- Hold a doctoral degree in human-computer interaction, artificial intelligence, or design fields
- Have more than 5 years of relevant field work experience
- Have participated in the development or evaluation of at least 3 related systems

Young User Group Screening Criteria:

- Age Range: 18-24 years old
- Gender Balance: Equal male-female ratio
- Diverse Academic Backgrounds: Covering science/engineering, humanities/history, arts, and other fields
- Exclusion: Participants with existing professional knowledge of Cizhou kiln

Creative Experience Group Selection Criteria:

- From 5 different disciplines: Industrial Design (n=15), Environmental Design (n=15), Materials Science (n=10), New Energy Engineering (n=10), and Transportation & Logistics Engineering (n=10)
- Basic computer operation skills
- Interest in cultural and creative fields

Table 2: Demographic Characteristics of Experimental Participants (N=200)

Study	Sample Size	Mean Age	Gender Ratio (%)	Educational Background(%)	Professional Field Distribution (%)
	20	40.35±2.54	Female: 50	Doctorate: 75	Human-Computer

Study 1 (Expert Evaluation)			Male: 50	Master's: 25	Interaction: 45
					Design: 35
					Cultural Heritage: 20
Study 2 (Technical Validation)	20	38.45±3.21	Female: 55 Male: 45	Doctorate: 70 Master's: 30	Artificial Intelligence: 40
					Computer Vision: 35
					Art & Design: 25
Study 3 (Efficacy Assessment)	100	21.40±2.35	Female: 50 Male: 50	Undergraduate: 85 Graduate student: 15	Art & Design: 30
					Humanities & Social Sciences: 25
					Engineering & Technology: 25
					Other: 20
Study 4 (Co-creation Experience)	60	20.85±1.75	Female: 50 Male: 50	Undergraduate: 100	Industrial Design: 25
					Environmental Design: 25
					Materials Science: 17
					New Energy Engineering: 17
					Transportation & Logistics Engineering: 16

Notes:

1. Age data presented as "mean ± standard deviation"
2. All percentages rounded to one decimal place
3. Professional field distribution based on participants' primary research/study area
4. Educational background represents highest degree or current enrollment at time of experiment

Table 2 presents the demographic characteristics of participants across four studies. Study 1 (Expert Evaluation) and Study 2 (Technical Validation) participants were primarily senior experts in relevant fields with high educational backgrounds and extensive professional experience. Study 3 (Efficacy Assessment) participants were university students covering multiple disciplines, ensuring sample diversity and representativeness. Study 4 (Co-creation Experience) specifically selected undergraduates from design-related majors to evaluate the system's potential application in design education. All studies maintained basic gender balance, meeting fundamental scientific experimental requirements.

Appendix D: Subjective Quality Assessment Instrument

Semantic Differential Scale for Generated Pattern Evaluation

Instructions: Please rate each generated pattern on the following dimensions using a 7-point scale.

Visual Aesthetics Very Unaesthetic 1 - 2 - 3 - 4 - 5 - 6 - 7 Very Aesthetic

Cultural Element Recognition (Ability to identify Cizhou ware cultural elements) Completely Unrecognizable 1 - 2 - 3 - 4 - 5 - 6 - 7 Highly Recognizable

Style Consistency (Consistency with traditional Cizhou ware style) Completely Inconsistent 1 - 2 - 3 - 4 - 5 - 6 - 7 Completely Consistent

Detail Fidelity Very Poor Details 1 - 2 - 3 - 4 - 5 - 6 - 7 Excellent Details

Overall Satisfaction Very Dissatisfied 1 - 2 - 3 - 4 - 5 - 6 - 7 Very Satisfied

Evaluation Procedure:

1. 30 patterns presented in random order (generation methods blinded)
2. 30-second viewing time for each pattern
3. Rating immediately after viewing
4. Optional open-ended comments

Appendix E:Details of the algorithm code

Appendix E1:

To optimize computational efficiency, a multi-scale image perceptual compression algorithm was designed to reduce computational complexity while maintaining pattern details:

Algorithm 1 Multi-scale Image Compression

Input: Original image x_o , downsampling factor f

Output: Compressed representation z

- 1: Initialize VAE encoder E and decoder D
- 2: H, W = dimensions (x_o)
- 3: **for** each scale s in $[1 \dots \log_2 f]$ **do**
- 4: x_s = downsample (x_o , scale = 2^s)
- 5: $z_s = E(x_s)$
- 6: $\tilde{x}_s = D(z_s)$
- 7: $L_s = \text{MSE} \left(x_s, \tilde{x}_s \right) + \text{KL}(z_s \parallel N(0, 1))$
- 8: Update E, D by minimizing L_s

```
9:   end for
10:   $z = E(x_0)$ 
11:  return  $z$ 
```

Experiments show that when $f=8$, computational complexity is reduced to $1/64$ of the original space, with single inference time controlled within 12.3 ± 2.1 seconds on NVIDIA RTX 3090 GPU environment.

Appendix E2:

Algorithm 2 Template-guided Pattern Generation

Input: Image features z , composition template M , denoise strength α

Output: Generated pattern z'

```
1:  Initialize noise map  $n \sim N(0, 1)$ 
2:  for each position  $(i,j)$  do
3:      if  $M[i,j] == 1$  then
4:           $z'[i,j] = \alpha \cdot n[i,j] + (1 - \alpha) \cdot z[i,j]$ 
5:      else
6:           $z'[i,j] = z[i,j]$ 
7:      end if
8:  end for
9:  Apply structural constraints:
10:      Check "three-seven" composition rule
11:      Maintain symmetry/asymmetry as specified
12:  return  $z'$ 
```

Appendix E3:

Algorithm 3 Template-guided Pattern Generation

Input: Original image X_o , reference pattern X_c , composition template M , text prompt T , cultural constraints C

Output: Generated pattern X_g

```
1: // Feature Extraction
2:  $Z_o$  = VAE_encoder( $X_o$ ) //VAE encoder extracts porcelain features
3:  $z_c$  = VAE_encoder( $X_c$ ) // Extract pattern features
4:  $h_{text}$  = CLIP_encoder( $T$ ) // CLIP text encoding
5: // Cultural Feature Constraint
6: for each constraint  $c_i$  in  $C$  do
7:     if  $c_i.type == "composition"$  then
8:          $alidate\_three\_seven\_ratio(M)$  // Validate three-seven composition
9:     end if
10:    if  $c_i.type == "style"$  then
11:         $validate\_black\_white\_contrast(z_c)$  // Validate black-white contrast
12:    end if
13:    if  $c_i.type == "semantic"$  then
14:         $validate\_cultural\_meaning(h_{text})$  // Validate cultural symbolism
15:    end if
16:    // Multi-modal Feature Fusion
17:     $h_{image}$  = IP_Adapter( $z_c$ ) // Inject pattern features
18:     $h_{fused}$  = CrossAttention( [ $Z_o$ ,  $h_{text}$ ,  $h_{image}$ ])
19:    // Iterative Refinement
20:     $z = z_o$ 
21: end for
22: for  $t$  in range( $T\_steps$ ) do
23:     $noise\_p\_red$  = UNet( $z$ ,  $t$ ,  $h_{fused}$ )
24:     $z$  = denoise_step( $z$ ,  $noise\_p\_red$ ,  $t$ )
25:    if  $t \% check\_interval == 0$  then
26:         $validate\_cultural\_constraints(z, C)$ 
27:    end if
28: end for
29: // Final Generation
30:  $X_g$  = VAE_decoder( $z$ )
31: return  $X_g$ 
```

- Pattern elements: Decorative element library organized by symbolic meaning.

Appendix F: Randomization Procedures and Statistical Power Details

F1. Random Assignment Procedure

Participants were randomly assigned to groups following a standardized protocol:

Randomization Method:

- Random sequences generated using R software (version 4.3.1) sample() function
- Random seed set (set.seed=2023) to ensure reproducibility
- Stratified block randomization employed (block size=4)

Stratification Variables:

1. Gender (male/female)
2. Age group (18-24, 25-34, 35-45 years)
3. Prior cultural heritage exposure (none/minimal/moderate, based on screening questionnaire Q1-Q3)

Allocation Concealment:

- Group assignments sealed in sequentially numbered, opaque envelopes
- Envelopes prepared by research assistant not involved in recruitment or assessment
- Opened only after participants completed informed consent and baseline assessment

Implementation Integrity:

- Random sequence generator (Author 3) had no direct contact with participants
- Recruitment and assessment conducted by independent team members (Authors 1 and 2)
- Allocation concealment maintained throughout the study

F2. A Priori Statistical Power Analysis

Power analysis conducted using G*Power 3.1.9.7:

Primary Analysis (Independent Samples t-test):

- Expected effect size: Cohen's $d=0.70$ (based on pilot study $n=20$ and literature: Rather 2019, $d=0.65$; Yu et al. 2024, $d=0.58$)
- Significance level: $\alpha=0.05$ (two-tailed)
- Desired power: $1-\beta=0.85$
- Required minimum sample size: $n=72$ (36 per group)
- **Actual sample size: $n=100$ (50 per group)**

- **Achieved power: 0.95**

Secondary Analysis (Cultural Interest Scale):

- Expected effect size: $f=0.35$
- Required sample size: $n=52$ (26 per group)
- **Actual power: 0.98**

Conclusion: Our sample size exceeds minimum requirements by 39%, providing adequate power to detect anticipated effects and substantially reducing Type II error risk.

F3. Baseline Equivalence Verification

Statistical tests confirmed successful randomization achieving baseline balance between groups:

Table F1: Baseline Characteristics Comparison Between Groups

Characteristic	Experimental (n=50)	Control (n=50)	Statistic	p-value
Age (M $\pm$ SD)	21.38 $\pm$ 2.41	21.42 $\pm$ 2.30	t (98)=-0.08	.934
Gender (% female)	52.0%	56.0%	χ^2 (1)=0.16	.689
Education (undergrad/grad)	76%/24%	70%/30%	χ^2 (1)=1.24	.538
Cultural exposure (1-5)	2.34 $\pm$ 0.82	2.28 $\pm$ 0.79	t (98)=0.38	.703
Baseline cultural knowledge	1.98 $\pm$ 0.85	2.02 $\pm$ 0.88	t (98)=-0.23	.819

All p-values $>.05$ indicate no significant baseline differences between groups, confirming successful randomization.

F4. Sensitivity Analysis

To verify result robustness, the following analyses were conducted:

Intention-to-Treat (ITT) Analysis: All randomized participants analyzed in assigned groups; results consistent with per-protocol analysis ($d=1.01$ vs. $d=1.02$)

Covariate-Adjusted ANCOVA: Controlling for baseline knowledge, age, and gender: $F(1,95)=87.34$, $p<.001$, partial $\eta^2=0.48$

Non-parametric Validation: Mann-Whitney U tests confirmed significance for all primary outcomes (all $p<.001$)

F5. Limitations

While the randomization procedure ensured baseline equivalence, complete blinding of participants and facilitators was not feasible due to the intervention's interactive nature. However:

- Outcome assessors were blinded to group assignment during knowledge test scoring
- Objective measurement instruments were used (standardized tests, system log data)
- These measures effectively mitigated potential assessment bias