HUMAN BIOLOGY
2024, VOL. 94, ISSUE 1
ORIGINAL ARTICLE

Configurational Effects for Enhancing the User Acceptance of Cultural Heritage Virtual Humans

Jia Li[1,2], *KieSu Kim*[2]

ABSTRACT

Digital technology has changed information dissemination and social interaction, creating new media and models of cultural communication. In the field of human-computer interaction and digital exploration of cultural heritage objects, virtual human (VH) design is increasingly attracting the attention of users and playing an important role. Although extensive empirical research has been conducted on the user experience of cultural heritage VH design, most of these studies have used experiments or questionnaires to study the impact of VH appearance and interaction characteristics on user perception. Research on the complex relationship between cultural heritage VH design factors and user acceptance through online data is lacking. Based on the computers-are-social-actors (CASA) paradigm, we first used the methods of literature research, industry observation and expert demonstration to identify five crucial factors in the design of cultural heritage VHs: character setting, cultural clues, visual image, voice interaction, and dynamic design. Then, we selected 28 cultural heritage VHs on China's Douyin short video platform as samples and employed fuzzy-set qualitative comparative analysis (fsQCA), which combines qualitative analysis and quantitative measurement, to conduct research to explore the influence of the design factors of cultural heritage VHs on user acceptance. The findings reveal three effective configuration modes for achieving high user acceptance, namely, the "emotionally oriented" mode, the "technologically oriented" mode, and the "all-factor-coordinated" mode, in which dynamic design factors are necessary elements of all the configurations. Considering the consistency and raw coverage of the configurations, the "all-factor-coordinated" approach is the optimal way to achieve high user acceptance of cultural heritage VHs. This study provides valuable empirical evidence for better design and optimization of cultural heritage virtual humans, as well as theoretical support and practical guidance for achieving high-quality and efficient user cultural experiences in human-computer interaction environments.

INTRODUCTION

With the continuous development of digital media and the deepening of intelligent technology, virtual human (VH), as a new form of life-like entity connecting the real world and the virtual world, have gradually demonstrated the technical, commercial, communication and social value they bring Xiang et al. (2023). As a new method based on artificial intelligence technology possessing typical “anthropomorphic” characteristics Peng et al. (2019), VHs have become a link for users to enter meta-universe scenes and digital life. In the context of digitization and intelligence, the nonphysical, controllable and personalized characteristics of VHs make them an important component of information dissemination and human-computer interaction. In the field of cultural heritage, new contents, forms and channels of communication have been generated through the innovative design of digital technologies and media interaction. This not only promotes the recognition of traditional cultural heritage but also enhances users’ sense of participation in and emotional resonance with cultural heritage and moreover expands the social influence of cultural heritage. This contributes to the sustainable preservation and transmission of cultural heritage and promotes the digital transformation of cultural heritage to meet the public's demand for cultural diversity. However, as a kind of digital object, the design and application of cultural heritage VHs are affected by various factors, such as technical limitations and user perception Deng et al. (2022). For example, the virtual image’s appearance realism, interaction naturalness, and intelligence-driven degree may limit its dissemination ability. Moreover, users’ cognitive patterns, aesthetic preferences, and technological needs also affect their acceptance of cultural heritage VHs.

Therefore, this study employs the fuzzy-set qualitative

[1] School of Journalism and Communication, Sichuan International Studies University, Chongqing 400031, China.
[2] School of design, Silla University, Busan 46958, Korea.

Correspondence to: KieSu Kim, School of design, Silla University, Busan 46958, Korea.
E-mail: 99001865@sisu.edu.cn.

Keywords: virtual human; VH design; cultural heritage; user acceptance; configurations; fsQCA

comparative analysis (fsQCA) research methodology Ragin et al.(2008) under the computers-are-social-actors (CASA) paradigm Nass et al.(1994) to investigate the complex relationship between key factors in the design of cultural heritage VHs and user acceptance. The aim is to answer the following questions: What are the key factors of cultural heritage VH design that impact user acceptance? Is there an independent factor that affects users' acceptance of cultural heritage VHs? Is there a configuration of factors that influence the high user acceptance of cultural heritage VHs? If so, how can these configurations be explained? This study provides valuable empirical evidence for better design and optimization of cultural heritage virtual humans, as well as theoretical support and practical guidance for achieving high-quality and efficient user cultural experiences in human-computer interaction environments.

RATIONALE AND RESEARCH FRAMEWORK

1. Computers-are-Social-Actors (CASA) paradigm

The CASA paradigm is an important theoretical framework in the field of human–computer interaction and has been widely used to explain public or user responses to emerging technologies Xu et al. (2022), Reeves et al. (1996). The core idea of the CASA paradigm is that people perceive media such as computers, TVs, and cell phones as social actors similar to humans and form specific impressions and rules based on the social cues (e.g., text, language, voices, faces, emotions, interactions, etc.) of these media, which ultimately results in corresponding emotional attitudes and behavioral intentions Moon et al. (2000).

According to the CASA paradigm, VHs are considered social entities that interact with humans. Previous studies have achieved many important results regarding human-computer interactions between users and intelligent terminals and avatars Ho et al. (2018), Carolus et al. (2019), Edwards et al. (2019). For example, Tan and Liew Tan et al. (2020) investigated the effect of virtual character design on consumers' social needs and purchase intentions by evaluating the interaction effect between consumers and virtual product consultants in product purchases. Shank et al. (2019) investigated the individual and joint decision-making of AI and humans at the ethical and moral levels through experiments with controlled reality scenarios. Peña et al. (2022). used obedience experiments to study the behaviors and emotions exhibited by VHs when they are under mental influence and to study how customized virtual doppelgangers interact with other VHs, either consciously or unconsciously. These studies have promoted the widespread application of the CASA paradigm through experiments on users' unconscious responses Nass et al. (2000) and conscious responses Xu et al. (2022) in human–computer interactions.

However, it is difficult for experiment-based approaches to accurately identify the complex relationships of VH design factors in human-computer interactions and how these factors play a role in configuration. Therefore, based on the CASA paradigm, it is important and valuable to explore the key design factors of cultural heritage VHs from the user's perspective with the help of social media data and then to study their impact on digital communication and interaction processes. This not only expands the theoretical explanatory power of the CASA paradigm but also further explores the relational logic and operational mechanism of human–computer interactions in the field of cultural heritage.

RESEARCH FRAMEWORK

VHs are generally defined as computer-generated characters with a human appearance Thalmann et al. (2014), and the expression of VHs should be appropriate within the context of communication Vinayagamoorthy et al. (2005). Moreover, VHs should have personality traits such as helpfulness and friendliness, as well as social skills, emotional expressiveness and interactivity Alessi et al. (2000). Under the CASA paradigm, numerous social cues, such as appearance, dynamics, expression, and personality, affect user perceptions and attitudes during human–computer interactions.

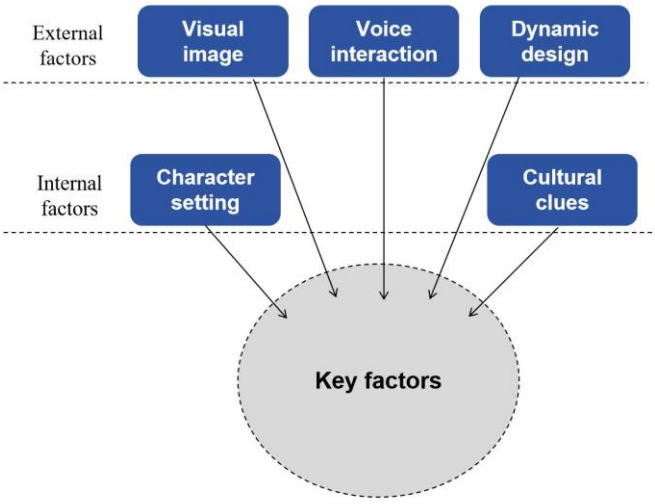
(1) Social cues: key factors in the design of cultural heritage VHs

Several established studies have categorized social cues into computers, virtual agents, etc. Nass et al. (2004) argued that social cues include language, voice, face, emotion, interactivity, engagement, and autonomy. Wuenderlich and Paluch Wuenderlich et al. (2017) studied users' perceptions of the authenticity of artificial intelligence (AI) and virtual agents and identified two different types of cues—agent-related and communication-related cues. Liew and Tan Liew et al. (2021) systematically reviewed AI experiments with expertise cues by examining the academic literature from 2005-2021 and proposed a classification of social cues (i.e., the DASSCI model) that encompasses demographics, appearance, social prestige, professionalism, communication style, and information quality Feine et al. (2019). further integrated the social cues of conversational assistants and categorized them into four first-level dimensions—visual, auditory, verbal cues, and intangible cues—for a total of 48 social cues. Zhang et al. (2010) categorized social cues into profile cues, verbal cues, and content cues based on computer-mediated communication. Xiang et al. (2023) used four dimensions of social cues—namely, appearance, persona, interactivity, and the technological base—to study the relationship between

VHs and user acceptance on the social media platform Xiao-Hongshu. These results provide a reference and inspiration for the multifactor acquisition of cultural heritage VHs studied in this paper.

To realize more comprehensive cultural preservation, inheritance, education, learning, and consumer transformation in the digital intelligence era, cultural heritage has gradually designed and developed VHs to apply to different scenarios. From simple 2D characters to complex 3D avatars, cultural heritage VH interactions have had a significant impact on user engagement, immersion, and learning efficiency Machidon et al. (2018). Based on the CASA paradigm and social cues theory, combined with previous research and industry observation, we propose five key factors in the design of cultural heritage VHs, namely, character setting, cultural cues, visual image, voice interaction, and dynamic design (Figure 1). To validate these findings, we recruited a group of three experts to review the proposed factors. The expert panel members are listed in Appendix A. All three had industry experience, and two were academics. All the participants were experts in digital product design and interaction design management, leading teams to use multiple independent media and complete multiple digital interaction design projects.

Figure 1: Five key factors of cultural heritage VH design



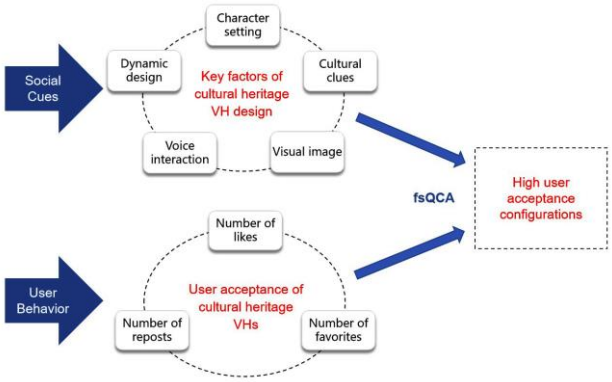
As shown in Figure 1, the five factors are categorized as internal or external factors. Among the internal factors, character setting is an attribute factor for identifying identity, and cultural clues are an endogenous factor for growth; from the external factors, visual image is the most intuitive image factor, voice interaction is a direct language factor, and dynamic design is a dynamic factor with a sense of stimulation.

(2) User behavior: user acceptance of cultural heritage VHs

User acceptance refers to how the presence of "others" affects people's cognition, emotion and behavior. In the field of human-computer interaction, user acceptance refers to people's attitudes and opinions about the social attributes of intelligence, such as robots and avatars, as well as their willingness to interact with them as social participants Krägeloh et al. (2019).

One of the key features of VHs regarding user acceptance is social engagement. With the emergence of cultural heritage VHs on social media platforms, they have accumulated a large number of followers and become influential social actors in the CASA paradigm. Current research focuses on the characteristics and interactivity of a single virtual object. Yang and Lee Yang et al. (2019) developed a perceptual model of a virtual assistant device to explain the intentions of potential customers when using the device. Yang and Zhao Yang et al. (2023) explored the logic of the symbolic construction of the virtual human "Liu Yexi", a cultural intellectual property, to study how to generate emotional resonance and feedback from users. In addition, some studies have focused on the technical aspects of entity virtualization Chittaro et al. (2004). studied how online museums use virtual tour guides to help users navigate, search, learn, and interact in virtual environments Kim et al. (2016). constructed cultural heritage virtual environments on mobile platforms; systematically researched the technologies, algorithms, and models of avatars and scenarios; and evaluated user acceptance based on questionnaires. However, these studies mainly relied on experiments or questionnaires with relatively limited collection and interpretation of big data, ignoring the broader social relations in human-computer communication and interaction. Therefore, using social media data to study user acceptance of influential cultural heritage VHs is a worthwhile research path.

Figure 2: Research framework for user acceptance of cultural heritage VHs



We take the key factors of cultural heritage VH design as the antecedent variable; user acceptance of cultural heritage VHs on social media as the outcome variable; and the number of likes, favorites, and reposts of cultural heritage VHs' accounts (including personal accounts and institutional accounts) as the specific measure of user acceptance to explore how to achieve high user acceptance of cultural heritage VHs. See Fig. 2 for the research framework.

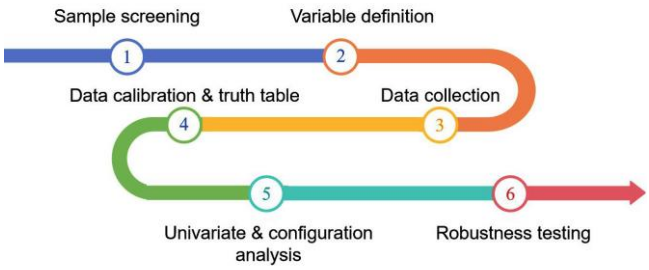
RESEARCH DESIGN

1. Research methodology

Qualitative comparative analysis (QCA) is a research method that combines both qualitative analysis and quantitative measurement to analyze the complex multivariate causal relationships between independent variables and dependent variables. This method uses multiple specific cases as qualitative indicators and has been found to be effective in small- and medium-sized sample studies Zhu et al. (2021). We used fsQCA Nass et al. (1994) to investigate the group relationships that affect the user acceptance of cultural heritage VHs. Compared to clear-set qualitative comparative analysis (csQCA), fsQCA allows variables to take values between 0 and 1 (0 represents "completely unaffiliated" and 1 represents "completely affiliated"). We assigned values to the condition variables using trichotomies of 0, 0.5, and 1, while the outcome variables were normalized to the data; therefore, the fsQCA method was chosen for the analysis in this study.

There are six operational steps (Figure. 3). (1) screening an appropriate number of sample cases based on the research questions; (2) defining the variables and designing scientific and objective measurements and assignments; (3) collecting data as needed; (4) data calibration and constructing truth tables; (5) Single-variable necessity analysis and group analysis; and (6) robustness testing to observe subset relationships as well as changes in consistency and coverage.

Figure 3: Steps in implementing the QCA research methodology



2. Data acquisition

To avoid interference from the media platform's user size, ecology and circles on the research results, we chose the social media platform Douyin (the Chinese domestic version of TikTok) as the platform for obtaining the research sample. Douyin is a social media platform Kaye et al. (2021) with short videos as the main form of content and currently has approximately 700 million active users. Douyin is not only a content viewing platform but also a social interaction space where users can interact, comment, like and share content with others Yang et al. (2021), Wang et al. (2020). Users on Douyin, especially young users, interact with each other on the platform and are prone to forming relevant interesting layers, which provides a rich dataset sample for user acceptance studies.

In selecting the VHs as samples, we first conducted keyword searches on Douyin. Hashtags such as "#cultural heritage virtual humans", "#cultural heritage digital person", "#virtual digital intelligence," "#metaverse virtual person", and "#metaverse digital person" were initially screened. Due to the small number of cultural heritage VHs with separate accounts, institutional accounts posting videos of cultural heritage VHs were also included in the initial screening. Then, further manual screening was conducted by combining the definitions of VH concepts and social cues from the previous literature review. Statistical analyses were conducted based on multiple data points, such as number of followers, posting activity, and interaction frequency, and an optimal account was reserved for VHs with overlapping identities and information. Finally, given that this study focused on cultural communication and human-computer interaction, samples intended for personal learning or product customization were cleansed. After several rounds of screening, we ultimately selected 28 cultural heritage VHs on the Douyin platform as research samples (Table 1). The 28 cultural heritage VHs are presented in Appendix B.

Since 16 VHs were not separate accounts, we could not separately count the number of followers for VHs of institutional accounts. Therefore, the number of followers is not included in the variable statistics. We utilized the number of likes, favorites, and reposts of short videos posted by each account from January 1, 2022, to September 20, 2023, as specific measures. It should be noted that all short videos of independent accounts were included, while only short videos with the hashtag "#VH's name" were counted for institutional accounts.

3. Variable design and assignment

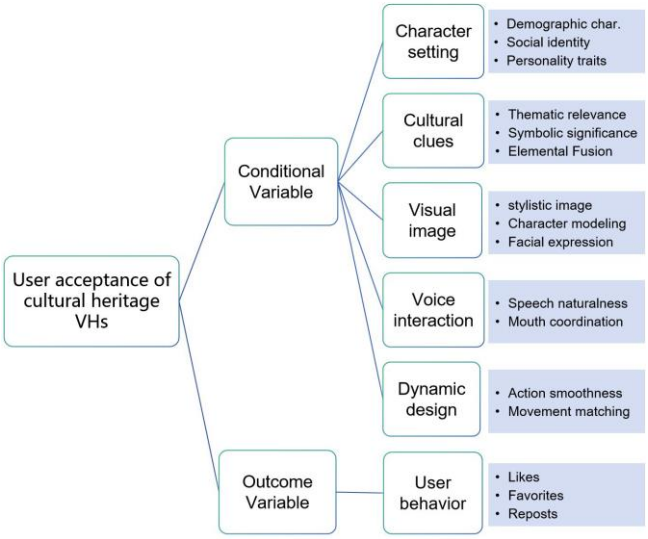
Generally, 4-7 conditional variables are considered the most appropriate range for fsQCA Yang et al. (2021).

Table 1: Samples of cultural heritage VHs on the Douyin platform

S/N	Name	Character	Account	Attribution
F01	Yaoyao Wen	Chinese Cultural Relics Exchange Center's first virtual spokesman for cultural heritage	48635230006	Independent account
F02	Xiaoya Qin	Virtual spokesman of Chinese Qinqiang Art Festival	71185847769	
F03	Xiaomei Su	Virtual digital human of Song culture	26683919718	
F04	Tianyu	Digital human representing Dunhuang Culture	tianyu_002354	
F05	Xiaoyu Gu	Digital human representing Song Dynasty culture	54737828707	
F06	Jingyue	Chinese Aerospace Museum's first aerospace virtual human	cascjingyue01	
F07	Xiaozhu	Chinese traditional culture virtual human	81102156163	
F08	Xuedun Boy	Virtual communicator of Tibetan culture	73930877245	
F09	Yanhuan Ling	virtual human representing the National Library of ancient books	lingyanhuan01	
F10	Lynn	Digital human in the west wing of Suzhou Museum	74165566811	
F11	Youzhi Qin	Oriental aesthetics of Chinese style virtual human	dyjcdgnjfcbgehj	
F12	Jiuli	Virtual spokesman for traditional festivals	93123607843	
F13	Wenwen Ai	First virtual staff of the National Museum of China	chnmuseum	Institutional account
F14	Gujin Tong	Virtual staff of the National Museum of China	chnmuseum	
F15	Jiayao	Cultural ambassador of digital Dunhuang	dhyiy1944	
F16	Meixiang Ling	China's first virtual human for the Greater Bay Area Folk Museums	nanduNshipin	
F17	Zhao Ban	Virtual staff of Guangzhou Qiqiao Cultural Museum	86461617700	
F18	Qinqin	virtual human of Guangzhou National Edition Museum	86461617700	
F19	Dongpo Su	Su Dongpo digital man of Chinese Bookstore	zhonghuashuju	
F20	Douyin Boy	First virtual staff of China Aerospace Museum	Douyin_guanfang	
F21	Feifei Fan	Digital guardian of intangible cultural heritage	frontop	
F22	Xiaoqin	First virtual human of Cantonese Opera Culture	XiaoQin115	
F23	Qingniao	Virtual referrer for Shanxi culture and tourism	shanxitourism	
F24	Xiaoqiao	Virtual man at the Museum of Rare Musical Instruments of Intangible Cultural Heritage	H66186688	
F25	Zichuan Li	Virtual anchor for traditional wood spinning techniques	15932762132li	
F26	Zijian Li	Virtual anchor for traditional wood spinning techniques	15932762132li	
F27	Jingjing Lou	Virtual promoter of Fujian Hakka Tulou	yjtl	
F28	Young Bai Li	A virtual image of the Tang Dynasty poet Bai Li during his youth	36995104251	
*Data as of September 20, 2023				

Based on the CASA paradigm and social cues theory, we combined previous related research and industry observation; focused on the theoretical analysis and variable indicator design of Xiang et al. (2023), Zuo et al. (2019), and Wen and Zuo Wen et al. (2019); proposed five conditional variables, character setting, cultural clues, visual image, voice interaction, and dynamic design; and constructed specific measurement indicators of these conditional variables. Moreover, with user acceptance as the outcome variable, the number of likes, favorites, and reposts were used as specific measurement indicators. A panel of three experts was also invited to deliberate on the settings, from variables to measurement indicators (see Appendix A). All three experts had industry experience, and two were academics. All the participants were experts in digital product design and interaction design management, leading teams to use multiple independent media and complete multiple digital interaction design projects. Finally, the variables and measurement indicators are shown in Figure. 4.

Figure 4: Variables and measurement indicators



First, in accordance with fsQCA, the specific measures of the condition variables were assigned values of 0, 0.5, and 1, and the specific measures of the outcome variables were normalized according to the number of likes, favorites, and reposts of the sample data. The CRITIC method was subsequently used in the Spsspro platform to internally weight the conditions and outcome variables. The CRITIC method is designed to comprehensively calculate weights by evaluating the variability and conflicts between indicators, and a larger weight indicates that the indicator contains more information and is relatively more important Wang et al. (2003). Finally, the variable design and assignment methods were obtained. The variable design and assignment methods are detailed in Appendix C.

(1) Explanation of Conditional Variables

One conditional variable is the character setting. A persona is the specific personality trait and behavior of an avatar, and different persona settings evoke different experiences and emotional connections from users Yang et al. (2023). Character setting is an internal factor that tends to constitute a conscious response of the user and contains specific measures such as demographic characteristics, social identity and personality traits. The demographic characteristics include a virtual human's age, sex, ethnicity, occupation, family, marriage, education, income, beliefs, and geographic information. Social identity refers to whether the virtual person has a clear position, such as belonging to educational, entertainment, or service types. Personality traits are unique, recognizable personality factors that a virtual person possesses.

The second conditional variable concerns cultural cues. By constructing perceptible cultural factors and providing meaningful connections with culture Sylaiou et al. (2022), it meets the diverse needs of users for cultural heritage protection, tourism, education, and consumer transformation. Cultural cues are an internal factor that tends to generate conscious reactions from users, with thematic relevance, symbolic significance, and elemental fusion as the specific measurement indices. Thematic relevance refers to the relationship of the communicated theme and content to the specific cultural heritage site. Symbolic significance focuses on the degree of observability and perceivability of cultural symbols used in avatar images and scenes. Elemental fusion indicates how well the skin, clothing, and makeup elements of the avatar image match and integrate with the specific cultural heritage.

The third conditional variable is visual image. Through three-dimensional expression simulation technology, the realism and real-time sense of facial expression and character modeling can be greatly improved Yao et al. (2019). A visual image is an external factor that tends to constitute an unconscious response of the user, and stylistic images, character modeling, and facial expressions are the specific measurement indices. Stylistic images are divided into three categories: cartoon, realistic and ultrarealistic images. A cartoon style inspires greater user favoritism, realism heightens the perception of character personality, and the appearance of ultrarealistism increases user trust Ruhland et al. (2015). Character modeling refers to a virtual person's on-camera image and includes head-oriented modeling, half-body-oriented modeling and full-body-oriented modeling. Facial expressions mainly consist of technologically generated character expressions or facial movements, which can be judged by the diversity and degree of realism of facial expressions through blinking, smiling, crying, etc.

The fourth conditional variable is voice interaction. The key feature that attracts users and creates a sense of immersion and participation is the emotional capacity of the VH Pfeiffer et al. (2011), and a naturally smooth, mouth-synchronized and emotional voice is more conducive to promoting user immersion Karuzaki et al. (2021). Voice interaction is an external factor that tends to favor the unconscious response of the user, with speech naturalness and mouth coordination as specific measures. Speech naturalness includes VHs' emotional fullness, comfort, and anthropomorphism. Mouth coordination represents whether a VH's mouth matches its voice when it speaks, and studies have shown that a poorly coordinated mouth will elicit negative feedback from users Debarba et al. (2020).

The fifth conditional variable is dynamic design. The dynamic capture of human gestures, bodies, expressions, mouths, and movements is used to synchronize VHs to simulate emotions and expressiveness Liu et al. (2020) to increase users' willingness to participate and interact. Dynamic design is also an external factor that favors unconscious user responses, with action smoothness and movement matching as specific measurements. Action smoothness indicates that the VH maintains a smooth movement path when executing actions and produces smooth and realistic responses. Movement matching refers to the degree of adaptation between the VH's actions, behaviors, and responses, as well as the context and the user's expectations.

(2) Interpretation of outcome variables

The primary outcome variable is user acceptance, which indicates the extent to which users pay attention to and accept cultural heritage VHs. This is comprehensively reflected through the specific measurement indicators of likes, favorites and reposts. The number of likes reflects the degree of user favoritism and acceptance of the content and information, the number of favorites reflects the value perception and personal preference of the content and information, and the number of reposts indicates the degree to which users recommend the content. During the data processing, the volume of published works was normalized to the mean to balance the sample sizes of different data. Finally, the data of these three specific measures were normalized to form the outcome variable of user acceptance of cultural heritage VHs.

DATA ANALYSIS

1. Data calibration

Given the different data types of conditions and outcome variables, we first used the direct method for data calibration in the fsQCA analysis. The three key points of full affiliation, crossover point and full nonaffiliation of the variables were set to 80%, 50% and 20%, respectively Papamitsiou et al. (2020). After calibration, the raw data were converted into fuzzy-set affiliation scores between 0 and 1. To avoid the group attribution problem that may occur when the variable affiliation is 0.5, we changed the value of 0.5 for affiliation to 0.501 Du et al. (2022).

2. Necessity and sufficiency analysis

Based on the QCA analysis framework, we first conducted univariate analysis through fsQCA3.0 software (Table 2). When judging the relationship between sufficiency and necessity for the condition variable and the outcome variable, we mainly observe the consistency index. When consistency > 0.8, the condition variable can be regarded as a sufficient condition for the outcome variable; when consistency > 0.9, the condition variable is a necessary condition for the outcome variable. In addition, we used coverage to determine the strength of the explanation of the outcome variable by the conditioning variable. A larger coverage indicates that the condition variable is more effective at explaining the outcome variable Nass et al. (1994).

As shown in Table 2, among the outcome variables with high user acceptance, although each variable has relatively good coverage (approximately 56%-71%), the consistency of the individual condition variables is less than 0.8. Obviously, no single factor constitutes a necessary or sufficient condition for achieving high user acceptance. Next, we continued to explore possible configurations among the condition variables.

3. Configurational effects analysis

Subsequently, we constructed a truth table for each conditional variable to analyze the combinations of conditional variables that affect the user acceptance of cultural heritage VHs. When analyzing the truth table, a case frequency threshold of 1 was set Rihoux et al. (2008), and both the original consistency threshold and the PRI consistency threshold were 0.8 Fiss et al. (2011), Zhang et al. (2020). After obtaining the complex solution, intermediate solution and parsimonious solution, the intermediate solution was used to generate the configuration, and the parsimonious solution was used to assist the intermediate solution in distinguishing the core conditions and marginal conditions of the configuration. In this case, the core conditions are those with a significant impact on the outcome, and the marginal conditions are those with a smaller but supporting effect on the outcome. In the configurational effects analysis of high user acceptance, three paths that met the specified conditions were generated (Table 3).

Among the three configurations with high user acceptance, the highest consistency was found in C2

Table 2: Necessity and Sufficiency Analysis of Single Conditional Variables

Condition Variable	High User Acceptance	
	Consistency	Coverage
Character setting	0.735238	0.639753
Cultural clues	0.780172	0.650488
Visual image	0.608202	0.569797
Voice interaction	0.640942	0.563602
Dynamic design	0.720955	0.710102

Table 3: High user acceptance configurations

Condition Variables	Configurations		
	C1	C2	C3
Character setting	●	⊗	●
Cultural clues	●	•	●
Visual image	⊗	⊗	•
Voice interaction	⊗	●	•
Dynamic design	●	●	●
Consistency	0.931034	0.984387	0.880126
Raw coverage	0.202957	0.10532	0.491189
Solution consistency		0.885772	
Solution coverage		0.6224	
●core condition, ●marginal condition, ⊗missing core condition, ⊗missing marginal condition			

(approximately 98.44%), followed by C1 (approximately 93.1%), and finally C3 (approximately 88.01%). The consistency of C1 and C2 was more than 90%, and the consistency of C3 was more than 85%, which indicated that all three configurations showed a strong correlation between the fulfillment of the conditional variables and the outcome variables. Thus, these three configurations are valid combination modes.

According to the core and marginal attributes of the influencing factors, the relationships among the three configurations are shown in Figure. 5.

As shown in Figure 5, C1 shows a combination of high-character setting, high cultural clues, and high dynamic design as core conditions, supplemented by nonhigh visual image and nonhigh voice interaction as marginal conditions.

C2 demonstrated a combination of high voice interaction, high dynamic design, and nonhigh visual image as the core conditions, supplemented by high cultural clues and a nonhigh character setting as the marginal conditions. C3 reflects a combination of a high-character setting, high cultural clues, and high dynamic design as the core conditions, supplemented by high visual image and high voice interaction as the marginal conditions. In addition, all three configurations specify dynamic design as a core condition, suggesting that dynamic design is an essential factor for high user acceptance.

According to the raw coverage data, the proportional relationships of the three configurations are shown in Figure. 6.

Figure 5: The composition relationship of the three configurations

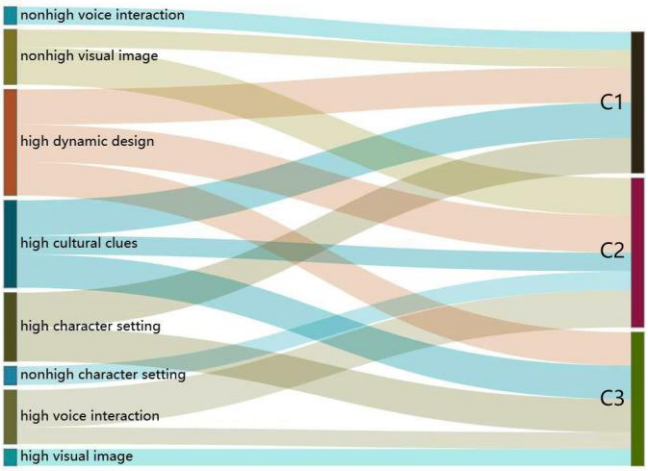
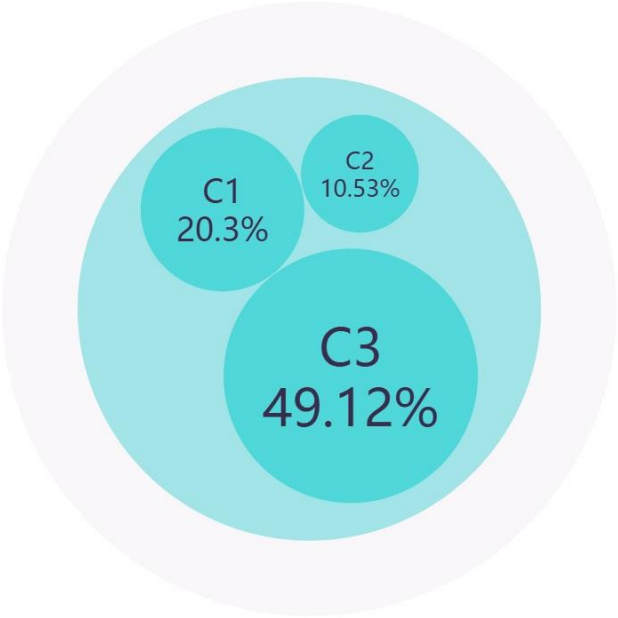


Figure 6: The proportional relationships of the three configurations



As shown in Figure 6, the raw coverage of C1 was approximately 20.3%, indicating that this configuration can explain a certain but limited number of samples. The raw coverage of C2 was approximately 10.53%, which was relatively low, indicating that the sample size that could be explained was relatively small. The raw coverage of C3 is approximately 49.12%, suggesting that the C3 configuration can explain approximately half of the samples and indicating this mode's strong significance in implementing outcome variables.

4. Robustness tests

To verify the robustness of the findings, the following tests were conducted in this study for the high user acceptance configurations Fiss et al. (2011).

The first test involved changing the calibration anchors. Under the conditions of satisfying the case frequency threshold of 1 and the original consistency threshold and PRI consistency threshold of 0.8, the fully affiliated and fully unaffiliated anchors of the variables were changed twice—first to 85% and 15%, and then to 75% and 25%, respectively—and the crossover point was unchanged at 50%. These conditions produced identical configurations to those in Table 3, with minor changes in the consistency and coverage of the solutions. In the second test, we changed the raw consistency threshold. Changing the raw consistency thresholds to 0.85 and 0.75 produced exactly the same configurations as in Table 3, with no change in the consistency or coverage of the solutions. Furthermore, the PRI consistency threshold can be changed. Reducing the PRI consistency threshold to 0.75 produced exactly the same configurations as in Table 3, and the consistency and coverage of the solutions remained unchanged.

Therefore, the robustness of the results obtained in this study is proven by the above robustness tests.

DISCUSSION & CONCLUSION

The factors influencing cultural heritage VHs and the configurational effects of high user acceptance are discussed as follows.

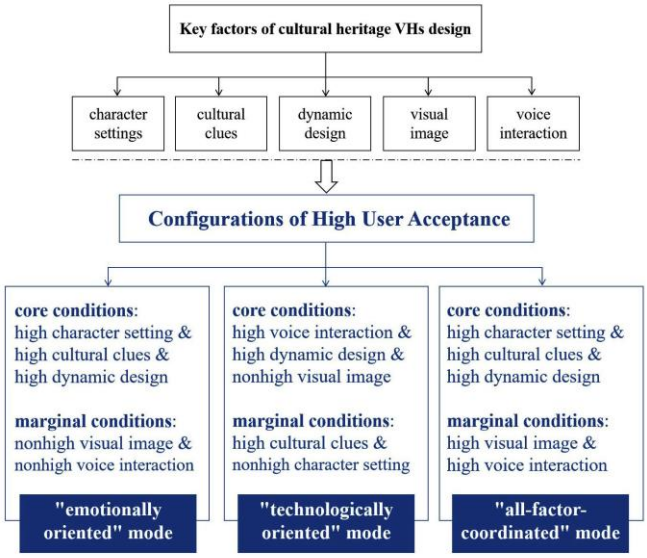
First, we identify five key factors of cultural heritage VH design through literature research, industry research, and expert discussions, including character settings, cultural clues, dynamic design, visual image, and voice interaction. These key factors together affect the experience and acceptance of Douyin users toward cultural heritage VHs. Among them, character setting is an attribute factor for identifying identity, cultural clues are an endogenous factor with growth, visual image is the most intuitive image factor, voice interaction is a direct language factor, and dynamic design is a dynamic factor with a sense of stimulation.

Next, the findings of the necessity and sufficiency analyses regarding the single influential factors indicate that while individual factors cannot be necessary or sufficient for high user acceptance, dynamic design factors are central in all high user acceptance configurations. This finding indicates that anthropomorphic movements and high matching

dynamics of VHs can strongly attract users and trigger strong emotional resonance. However, although dynamic design plays an extremely critical and important role, it is still not a necessary or sufficient condition for generating high user acceptance on its own. Even if the technical input and simulation are excellent, users may not be able to gain full recognition and acceptance independently. The user acceptance of cultural heritage VHs is driven by a combination of influential factors.

Furthermore, three configurations were found to high user acceptance of cultural heritage VHs (Fig. 7).

Figure 7: Configurations of high user acceptance of cultural heritage VHs



The first configuration takes high character setting, high cultural clues, and high dynamic design as core conditions, supplemented by nonhigh visual image and nonhigh voice interaction as marginal conditions. This mode emphasizes emotional orientation. From the perspective of this "emotionally oriented" mode, through in-depth characterization and cultural transformation of cultural heritage VHs, avatars establish good emotional connections with users. The second configuration combines high voice interaction, high dynamic design, and nonhigh visual image as the core conditions, supplemented by high cultural clues and nonhigh character setting as the marginal conditions. The focus of this model is on technology orientation. This "technologically oriented" mode emphasizes voice interaction and action experience with strong intelligence and high fidelity, which may have a decisive influence on the cognitive attitude of users. The third configuration reflects a combination of high character setting, high cultural clues, and high dynamic design as the core conditions, supplemented by high visual image and high voice interaction as the marginal conditions. This model reflects the coordination of all factors. This "all-factor-coordinated" mode reflects the balanced development layout of all the influencing factors. Considering the consistency and raw coverage of the configuration, the "all-factor-coordinated" approach is the optimal combination for enhancing the user acceptance of cultural heritage VHs. This combination path is both prioritized and comprehensive, resulting in effective interactivity and communication power. Although characterization and cultural elements are the main cues used to attract the public, images and speech are still needed to enhance the quality of interaction and information dissemination. These configurations help differentiate and understand the multiple ways in which cultural heritage VHs are socially acceptable in different contexts and help developers and operators create intellectual property for cultural heritage VHs. It should be noted that all the configurations specify dynamic design as a core condition, indicating that dynamic design is crucial and a necessary factor for all high user acceptance models.

In future research, the issue of user acceptance of cultural heritage VHs can be explored in several ways. First, using diverse samples is a key approach. Based on platform tonality and user scale, Douyin was chosen as the data collection platform in this study, which limited the sample size to a certain extent. Future research can consider expanding the sample to more media platforms and application scenarios. In addition, additional international and cross-cultural indicators can be included to compare the user acceptance of cultural heritage VHs in different countries and cultural contexts. This approach will help us achieve a more comprehensive understanding of cultural heritage VHs in terms of digital intelligence. The second suggestion is that the scope of research variables and specific measurement indicators are also important research directions in the quantitative study of user acceptance of cultural heritage VHs. Under the CASA paradigm, this paper proposes variables and specific measurement indicators based on the literature and industry observation. However, due to the limitations of research focus and data availability, this study does not take into account additional factors such as capital examination, brand influence, or ethical issues. In the future, with an increase in relevant data, researchers can perform more comprehensive and accurate assessments of user acceptance of cultural heritage VHs with the help of more comprehensive indicators and measurement methods. This approach will help to better understand the impact of various factors on the acceptance of VHs and how to further improve the user acceptance of VHs. The third

suggestion is that the social response mechanism of users to intelligent and virtual technologies should be a research area worthy of attention. This includes both conscious and unconscious responses during human-computer interaction. Researchers can use a variety of methods, such as questionnaires, eye-tracking, and cognitive experiments, to capture the conscious behavior of users. Neuroscience methods and emotion recognition techniques can also be applied to test the potential unconscious responses of users. These responses will contribute to a more comprehensive understanding of how the general public interacts with cultural heritage VHs, providing additional insights and guidance to improve the interaction process and enhance the user experience.

DECLARATIONS

Acknowledgements

We thank AJE (https://www.aje.cn/) for its linguistic assistance during the preparation of this manuscript.

Ethical Approval

The authors declare the integrity of the scientific record. All authors, read, reviewed, edited and approved the manuscript.

Funding

This research is supported by Sichuan International Studies University Research Fund, China (No. sisu2019017).

Availability of data and materials

Not applicable.

Competing interests

The authors declare no competing interests.

REFERENCES

1.Xiang AL, Li TZ, Ma MY, et al. 2023. How to improve the public acceptance of virtual human: a Qualitative Comparative Analysis based on 36 cases. Journalism & Communication Review. 76(02): 26-42.

2.Peng L. 2019.Digital survival of people in the intelligent era -- separable "virtual entities", "digital components" and "embodiment" that will not disappear. Shanghai Journalism Review. (12): 4-12.

3.Deng Z. 2022.Meta Human: media promotion under identity reorganization and relationship empowerment. Study and Practice. (09): 129-135.

4.Ragin C. 2008. Redesigning social inquiry: Fuzzy sets and beyond. Chicago: University of Chicago Press.

5.Nass C, Steuer J, Tauber E R. 1994.Computers are social actors. Proceedings of the SIGCHI conference on Human factors in computing systems. 72-78.

6.Xu K, Chen X, Huang L. 2022.Deep mind in social responses to technologies: A new approach to explaining the Computers are Social Actors phenomena. Computers in Human Behavior. 134: 107321.

7.Reeves B, Nass C. 1996. The media equation: How people treat computers, television, and new media like real people. Cambridge, UK. 10(10).

8.Moon Y. 2000. Intimate exchanges: Using computers to elicit self-disclosure from consumers. Journal of consumer research. 26(4): 323-339.

9.Ho A, Hancock J, Miner A S. 2018 Aug. Psychological, Relational, and emotional effects of self-disclosure after conversations with a chatbot. J Commun. 68(4):712-733.

10.Carolus A, Binder J F, Muench R, et al. 2019.Smartphones as digital companions: Characterizing the relationship between users and their phones. New Media & Society.21(4): 914-938.

11.Edwards C, Edwards A, Stoll B, et al. 2019. Evaluations of an artificial intelligence instructor's voice: Social Identity Theory in human-robot interactions. Computers in Human Behavior. 90(1): 357-362.

12.Tan S M, Liew T W. 2020. designing embodied virtual agents as product specialists in a multi-product category E-commerce: The roles of source credibility and social presence. International Journal of Human–Computer Interaction. 36(12): 1136-1149.

13.Shank D B, DeSanti A, Maninger T. 2019. When are artificial intelligence versus human agents faulted for wrongdoing? Moral attributions after individual and joint decisions. Information, Communication & Society. 22(5): 648-663.

14.Peña J, Craig M, Baumhardt H. 2022.The effects of avatar customization and virtual human mind perception: A test using Milgram's paradigm. New Media & Society. 14614448221127258.

15.Nass C, Moon Y. 2000. Machines and mindlessness: Social responses to computers. Journal of social issues. 56(1): 81-103.

16.Thalmann D, Maïm B, Maïm J. 2014.Geometric issues in reconstruction of virtual heritage involving large populations. 3D Research Challenges in Cultural Heritage: A Roadmap in Digital Heritage Preservation. 78-92.

17.Vinayagamoorthy V, Steed A, Slater M. 2005.Building characters: Lessons drawn from virtual environments. Proceedings of Toward Social Mechanisms of Android Science: A CogSci 2005 Workshop. Stressa, Italy: COGSCI 2005. 119-126.

18.Alessi N E, Huang M P. 2000.Evolution of the virtual human: From term to potential application in psychiatry. CyberPsychology & Behavior. 3(3): 321-326.

19.Nass C. 2004.Etiquette equality: exhibitions and expectations of computer politeness. Communications of the ACM. 47(4): 35-37.

20Wuenderlich N V, Paluch S. 2017. A nice and friendly chat with a bot: User perceptions of AI-based service agents. ICIS 2017 Proceedings. 11.

21.Liew T W, Tan S M. 2021.social cues and implications for designing expert and competent artificial agents: A systematic review. Telematics and Informatics. 65: 101721.

22.Feine J, Gnewuch U, Morana S, et al. 2019. A taxonomy of social cues for conversational agents. International Journal of Human-Computer Studies. 132: 138-161.

23.Zhang F. 2010.Illusion and Reality. Beijing: China Social Sciences Press.

24.Machidon O M, Duguleana M, Carrozzino M. 2018. virtual humans in cultural heritage ICT applications: A review. Journal of Cultural Heritage. 33: 249-260.

25.Krägeloh C U, Bharatharaj J, Sasthan Kutty S K, et al. 2019. Questionnaires to measure acceptability of social robots: a critical review. Robotics. 8(4): 88.

26.Yang H, Lee H. 2019.Understanding user behavior of virtual personal assistant devices. Information Systems and e-Business Management. 17: 65-87.

27.Yang PP, Zhao JG. 2023.Fusion of reality and reality and dimensionality enhancement of scenes: Research on the IP construction of the virtual human "Liu Yexi" from the perspective of the metaverse. Broadcasting Realm. (4): 53-56.

28.Chittaro L, Ieronutti L, Ranon R. 2004.Navigating 3D Virtual Environments by Following Embodied Agents: a Proposal and its Informal Evaluation on a Virtual Museum Application. PsychNology Journal. 2(1): 24-42.

29.Kim L C, Lam T K, Chee C Y. 2016. A multi-modal virtual walkthrough of the virtual past and present based on panoramic view, crowd simulation and acoustic heritage on mobile platform. International Journal of Computer and Information Engineering. 10(10): 1869-1879.

30.Zhu H, Zhang MX. 2021. Influential factors and configurational effects of user privacy setting behavior in mobile app. Information Science. 39(7): 54-62.

31.Kaye D B V, Chen X, Zeng J. 2021.The co-evolution of two Chinese mobile short video apps: Parallel platformization of Douyin and TikTok. Mobile Media & Communication. 9(2): 229-253.

32.Yang Y, Ha L. 2021.Why people use TikTok (Douyin) and how their purchase intentions are affected by social media influencers in China: A uses and gratifications and parasocial relationship perspective. Journal of Interactive Advertising. 21(3): 297-305.

33.Wang Y. 2020.Humor and camera view on mobile short-form video apps influence user experience and technology-adoption intent, an example of TikTok (DouYin). Computers in Human Behavior. 110: 106373.

34.Zuo B, Wen FF, Song JJ, et al. 2019.The characteristics, dimensions and psychological effect of social categorization. Advances in Psychological Science. 27(1).

35.Wen FF, Zuo B. 2019.Concepts,cues and influential mechanisms of social categorization. Journal of Psychological Science. 42(02): 395-401.

36.Wang K, Song HZ. 2003. Comparative analysis of three objective weighting methods. Technoeconomics & Management Research. (06): 48-49.

37.Yang MY Zhu L. 2023.Media presentation of virtual human gender roles from the perspective of the metaverse. Youth Journalist. (10): 85-86.

38.Sylaiou S, Fidas C. 2022. virtual humans in Museums and Cultural Heritage Sites. Applied Sciences. 12(19): 9913.

39.Yao SM, Li WH, Li WQ, et al. 2019. Facial expression simulation technology for virtual avatar. Journal of Graphics. 40(03): 525-531.

40.Ruhland K, Zibrek K, McDonnell R. 2015.Perception of personality through eye gaze of realistic and cartoon models. Proceedings of the ACM SIGGRAPH Symposium on Applied Perception. 19-23.

41.Pfeiffer T, Liguda C, Wachsmuth I, et al. 2011.Living with a virtual agent: Seven years with an embodied conversational agent at the Heinz Nixdorf MuseumsForum. Proceedings of the Re-Thinking Technology in Museums 2011-Emerging Experiences.

42.Karuzaki E, Partarakis N, Patsiouras N, et al. 2021.Realistic virtual humans for cultural heritage applications. Heritage. 4(4): 4148-4171.

43.Debarba H G, Chagué S, Charbonnier C. 2020.On the plausibility of virtual body animation features in virtual reality. IEEE Transactions on Visualization and Computer Graphics. 28(4): 1880-1893.

44.Liu TT, Liu Z, Xu HH, et al. 2020.Modeling virtual humans' emotions based on cognitive appraisal theory of emotion: A review. Journal of Psychological Science. 43(01):53-59.

45.Papamitsiou Z, Pappas I O, Sharma K, et al. 2020.Utilizing multimodal data through fsQCA to explain engagement in adaptive learning. IEEE Transactions on Learning Technologies. 13(4): 689-703.

46.Du YZ, Liu QC, Chen KW, et al. 2022.Ecosystem of doing business, total factor productivity and multiple patterns of high-quality development of Chinese cities: A configuration analysis based on complex systems view. Journal of Management World. 38(09): 127-145.

47.Rihoux B, Ragin C C. 2008.Configurational comparative methods: Qualitative comparative analysis (QCA) and related techniques. Sage Publications.

48.Fiss P C. 2011.Building better causal theories: A fuzzy set approach to typologies in organization research. Academy of Management Journal. 54(2): 393-420.

49.Zhang M, Lan HL, Chen WH, et al. 2020. Research on the antecedent configuration and performance of strategic change. Management World. 36(09): 168-186.