Factors Affecting The Acceptance Of Virtual Digital Human Technology In Live Broadcasting Settings Based On The UTAUT Model

Xingxing Lu[1], Ziyu Zhou[2], Yang Bai[2,*]

[1]Wuhan Donghu University, Wuhan, China
[2]School of Information Management, Central China Normal University, Wuhan, China

*yang.bai@ccnu.edu.cn

Abstract. Based on the UTAUT paradigm, this study aims to identify the factors that influence consumers' acceptance of virtual digital human technology in live broadcasting contexts. The study employed a random sampling approach and had a total of 351 valid respondents and the findings showed that five factors, namely, social influence, anchor anthropomorphism, interactivity, perceived amusement, and facilitating conditions, have a significant and positive influence on users' willingness and behavior to watch real-person-driven virtual digital human, whereas effort expectancy and performance expectancy have no significant effect.

Keywords: Network Live Streaming, Virtual Digital Human, Real-Person-Driven Virtual Digital Human, UTAUT Model

1 Introduction

The evolution of artificial intelligence technology has supplied virtual digital humans with more "human-like" mental traits[1], hence increasing their range of applications and activities. Virtual digital humans employ modern display technologies to create virtual images that can imitate real-world feedback. In the live broadcasting industry, the rise of virtual digital anchors has become a clear reality, boosting user engaging experiences. However, as virtual digital human technology becomes more integrated into the live broadcasting environment, viewers' attitudes and acceptance of real-person-driven virtual digital humans, as well as the variables influencing these characteristics, have yet to be thoroughly investigated and understood. In light of this, the goal of this study is to look into the factors that influence consumers' adoption of real-person-driven virtual digital human technology, in order to provide theoretical support and practical guidance for its use in live broadcasting.

J. Yin et al. (eds.), *Proceedings of the 4th International Conference on New Media Development and Modernized Education (NMDME 2024)*, Advances in Intelligent Systems Research 188,
https://doi.org/10.2991/978-94-6463-600-0_66

2 Review

The current research of Virtual Digital Human(VDH) technology focuses mostly on technological advances and industrial applications. On the technology front, researchers concentrate primarily on image and motion modeling, with the goal of strengthening the simulation effect of virtual digital humans through meticulous detail refinement. Zhaoqi proposed a series of behavioral modeling techniques for VDH, including keyframe methods, inverse kinematics methods, and motion capture methods, which provide a variety of technical paths for representing virtual digital human behaviors[2]. Furthermore, some experts believe that emotional anthropomorphism might improve the facial expressions and body movements of virtual digital humans, resulting in a greater connection with consumers[3].In terms of applications, virtual digital humans are widely used in a variety of industries. In the library industry, virtual digital humans can improve services such as intelligent Q&A and reading promotion[4]. Virtual digital humans have piqued the interest of many fans and game aficionados in the world of webcasting because to their distinct attractiveness and interactivity[5].

Real-person-driven virtual digital human technology is an innovative artificial intelligence application that uses motion capture and voice capture technologies to map real human performances in real-time onto virtual characters, resulting in virtual images with the characteristics of real human movements and voices that can interact with users in real-time. In education, virtual digital humans improve distance learning experiences by replicating teachers' facial expressions and body language[6]. In the medical area, they act as virtual doctors, offering emotional support and psychological counseling in the same way that therapists would[7]. In the realm of webcasting, the employment of real-person-driven virtual digital human technology not only enriches the content of live broadcasting, but also enhances the user's willingness to interact and sense of identity[8].

In summary, contemporary research on virtual digital human technology has achieved great achievements in technological progress and commercial applications, and has begun to look into the live broadcasting business. However, in-depth research on user acceptability of real-person-driven virtual digital human technology in live broadcasting is limited, particularly in the systematic examination of its affecting elements. This paper uses the UTAUT model to investigate the influencing factors of users' acceptance of real-life-driven virtual digital human technology using a questionnaire survey method, with the goal of providing theoretical support and practical guidance for the live broadcasting industry as well as the application of virtual digital human.

3 Research Methodology

Based on the Unified Theory of Acceptance and Use of Technology (UTAUT), this study investigated users' willingness to embrace virtual digital human technology in live broadcast scenarios using a questionnaire survey. The UTAUT model was proposed by Venkatesh et al. in 2003[9]. This model combines the strengths of eight

theoretical models and is a useful theoretical tool for evaluating technology acceptability and user behavior. In addition to the live broadcast settings, this study used a model based on the UTAUT theory, which retained the original four independent variables of performance expectancy(PE), effort expectancy(EE), social influence(SI), and facilitating conditions(FC). However, utilizing the UTAUT model alone to investigate the user's watching behavior may be insufficiently explanatory and may overlook some key factors. Based on the study's scenario, anchor anthropomorphism(AA), interactivity(IA), and perceived amusement(PA) were added to the UTAUT model to create the conceptual model for this study.

3.1 Assumption Model

Although the UTAUT model incorporated behavioral intention and use behavior, different researchers used somewhat different dependent variables in their investigations. For example, Zhou[10] utilized behavioral intention as the dependent variable; Devaraj et al.[11] used usage frequency as the dependent variable; and Lin et al.[12] kept both behavioral intention and use behavior as dependent variables. This study used the UTAUT model's two dependent variables, behavioral intention and behavior, to determine the level of acceptance of virtual digital human technology by examining viewing intention and viewing behavior when watching virtual digital humans live in live broadcasting settings. Accordingly, the dependent variables in this study were willingness to watch and viewing behaviors, while the independent variables were based on the UTAUT model, which included PE, EE, SI, FC, PA, IA, and AA. Figure 1 depicts the user acceptability model for virtual digital human technology in live broadcasting circumstances, which was developed in this paper.

3.2 The Design and Implementation of a Questionnaire

The model hypotheses led to the selection of 28 elements. The items in the questionnaire were designed using a Likert scale, with number alternatives ranging from 1 to 5 reflecting five levels of sentiment, from "strongly disagree" to "strongly agree." Additionally, to prevent respondents from randomly answering, this section included an attention check item to select valid questionnaires in subsequent processing. The questionnaire for this study was created using the online research tool "Wenjuanxing." Prior to the formal distribution of the questionnaire, a small-scale sample test was conducted to check the scale's reasonableness, and 27 questionnaires were distributed. Following the test, the descriptions of questions with a low level of cognitive differentiation by individual users were revised and improved, and 29 questionnaires were reprinted. The analytical results indicated that the obtained data's reliability and validity were excellent. Then, between July 7, 2023 and July 27, 2023, a total of 552 surveys were officially provided, with 351 valid questionnaires obtained after eliminating invalid questionnaires filled out incorrectly for attention detection and invalid questionnaires with all of the identical alternatives.

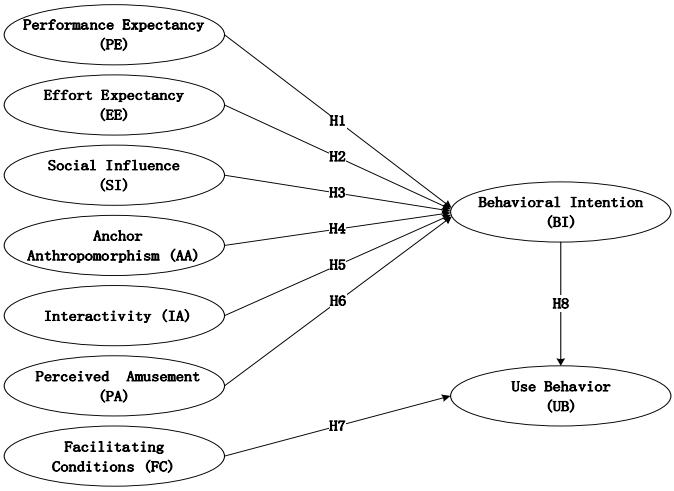

Fig. 1. The initial model of user acceptance of virtual digital human technology in the live broadcasting settings

4 Result

The article used AMOS 24.0 to analyze the questionnaire data. First, a preliminary model is built using the theoretical framework, with variables defined and path linkages established. After importing the questionnaire data, Cronbach's coefficient, Bartlett's sphericity, and AVE values were used to assess its reliability and validity. A standard partial variance test was then used to determine the correlation between the variables. Following that, a model fit test was performed to determine model usability using the first SEM fit values. Indeed, based on these indicators, the questionnaire design was found to be very reliable through two pre-tests of the questionnaire data and subsequent questionnaire changes. Finally, the initial model fit and path coefficient calculations were carried out to check the SEM path coefficients and significant routes of the final model.

4.1 Descriptive Statistical Analysis of the Questionnaire

The obtained sample data underwent statistical analysis, which included gender, age, education, and the average amount of time spent watching live broadcasts per day. The results are shown in Table 1.

Table 1. Results of demographic analysis of the sample

Variants	Options	Frequencies	%
Gender	Male	162	46.2%
	Female	189	53.8%
Age	Under 25	107	30.5%

	26-35	131	37.3%
	36-50	76	21.7%
	Over 50	37	10.5%
Education	High school and below	50	14.2%
	Junior college students	82	23.4%
	Undergraduate students	140	39.9%
	Graduate students and above	79	22.5%
Average length of time spent watching live broadcasts per day	Less than 10 minutes	104	29.6%
	10-20 minutes	89	25.4%
	21-30 minutes	79	22.5%
	31-60 minutes	48	13.7%
	More than 1 hour	31	8.8%

Table 1 shows that males and females are relatively balanced with 46.2% and 53.8% respectively. In terms of age, young individuals under the age of 25 and adults aged 26 to 35 make up the majority of respondents, accounting for 30.5% and 37.3% of the total, respectively. In terms of education, undergraduates have the biggest share (39.9%), followed by postgraduates (22.5%), indicating that respondents have a higher level of education. In terms of watching live broadcasts, the majority of respondents watch them for more than 10 minutes per day, with the greatest percentage of 21-30 minutes at 22.5%, demonstrating that respondents are interested in and engaged with live broadcast content.

4.2 Questionnaire Reliability and Validity Testing

According to reliability analysis, Cronbach's alpha values for all variables range from 0.8 to 1, with an overall Cronbach's alpha value of 0.978, indicating that the scales used in this study have high internal consistency and reliability. In terms of validity, the total KMO score is 0.971 (more than 0.8), suggesting that the results are extremely significant. Meanwhile, Bartlett's sphericity test yields a significance level of less than 0.01, supporting the data's importance. Furthermore, discrimination validity was assessed, and the AVE square root values of the variables demonstrated strong overall validity distinction, which was acceptable.

Common method biases were also assessed to ensure that the model fit test could be done. The results in Table 2 show that RMSEA fell by 0.007 (<0.05), SRMR decreased by 0.004 (<0.05), CFI increased by 0.008 (<0.1), and TLI increased by 0.008 (<0.1), indicating that there are no significant common technique biases and that it passes the test.

Table 2. Results of demographic analysis of the sample

Fit Index	RMSEA	SRMR	CFI	TLI
Master Model	0.051	0.025	0.973	0.967
Model After Adding Potential Method Factors	0.044	0.021	0.981	0.975
Variations	-0.007	-0.004	0.008	0.008

4.3 Hypothesis Test Results

The model's path coefficient analysis was performed using AMOS 24.0 (S.E. stands for standard error), and it was discovered that six of the eight paths had significant and acceptable path coefficients, while two did not. Table 3 shows the outcomes of the hypothesis tests.

Table 3. Hypothesis test results

hypothesis	Hypothetical contents	path coefficients	P	Result
H1	Performance expectations positively influence users' willingness to watch real person-driven virtual digital human anchors	-0.079	0.223	Unsupported
H2	Effort expectations positively influence Users' willingness to watch real person-driven virtual digital human anchors	-0.157	0.162	Unsupported
H3	SI positively influence users' willingness to watch real person-driven virtual digital human anchors	0.214	0.033	Supported
H4	AA positively influence users' willingness to watch real person-driven virtual digital human anchors	0.352	***	Supported
H5	IA positively influence users' willingness to watch real person-driven virtual digital human anchors	0.352	0.033	Supported
H6	PA positively influence users' willingness to watch real person-driven virtual digital human anchors	0.279	0.038	Supported
H7	FC positively influence users' willingness to watch real person-driven virtual digital human anchors	0.503	***	Supported
H8	Behavioral intention positively influence users' willingness to watch real person-driven virtual digital human anchors	0.566	***	Supported

("***" means Sig.<0.001)

5 Discussion

The empirical research results reveal that social influence affects users via herd mentality, indicating the strength of media and social contexts in altering users' interests. Anchor anthropomorphism as a basic element of real-person-driven virtual digital human, with its high level of realism and adaptability, makes it impossible for consumers to discern between the virtual and the real[13], hence improving the watching experience. Perceived amusement is closely related to users' attraction to live material, therefore fascinating and innovative live situations become critical for user retention. The identification of these key components is consistent with previous research, which emphasizes the interaction of technological, psychological, and social factors in the process of user acceptance of new technologies.

Although performance expectancy and effort expectancy are widely recognized as important components in the UTAUT paradigm, the findings of this study indicate that their impact on adoption of real-person-driven virtual digital human technology is negligible. On the one hand, the performance expectancy in this paper focuses on and measures the extent to which virtual human anchors can provide cognitive value to users, whereas real-person-driven virtual digital human anchors are more widely used in areas with high interaction, such as entertainment live broadcasting and game live broadcasting, and the "real people" behind them provide users with a more vivid and flexible viewing experience via real-time control. As a result, in terms of cognitive value, it is influenced by the pan-entertainment environment and has a weaker link. According to certain research, users are more interested in the entertainment provided by live anchors during live broadcasting[14]. They are more concerned with the calm sensation of physical performance and the emotional empathy experience provided by live anchors, and they rarely consider the sense of identity in values provided by them. As a result, consumers have a low performance expectancy for the cognitive value provided by real-person-driven virtual digital human anchors. On the other side, low effort expectancy may be due to the convenience of live broadcasting and users' familiarity with the technology. As technology advances and becomes more popular, consumers become more skilled with live streaming technology, which means that the advanced nature of virtual digital human technology makes it difficult for consumers to experience operational difficulties while watching live programming, and there is no need to learn and waste energy on convenience of use, thus effort expectancy is minimal.

6 Conclusion

Based on the UTAUT paradigm, this study uses a questionnaire survey to investigate the factors influencing users' acceptance of real-person-driven virtual digital human technology in live broadcasting scenarios. The study's findings show that five aspects, namely social influence, anchor anthropomorphism, interactivity, perceived amusement, and facilitating conditions, have a significant and beneficial effect on users' viewing intention and behavior. These aspects include not only technological facilitating settings, but also psychological views and social implications, demonstrating that user acceptance is a multidimensional and complicated process. Concurrently, empirical investigation reveals the possible use of real-person-driven virtual digital human technology in live broadcasting. More research should be done on the aspects that influence the experience of specific user groups in segmented live broadcasting settings in order to completely comprehend the individualized application strategies of real-person-driven virtual digital human technology in the live broadcasting area. This will help to build a good and sustainable virtual digital human live broadcasting ecosystem that meets the needs of users while also boosting the growth of the live broadcasting sector.

Acknowledgment

This research was partly supported by Wuhan Donghu University Youth Foundation Program(2024dhsk004), and the National Social Science Foundation of China (23CTQ019).

References

1. Stein J P, Appel M, Jost A, et al.: Matter over mind? How the acceptance of digital entities depends on their appearance, mental prowess, and the interaction between both. International Journal of Human-Computer Studies, 142, 102463 (2020).
2. Wang Zhaoqi: Study on Synthesis of Virtual Human. Journal of Graduate School, Academia Sinica 02,89-98 (2000).
3. Ham J, Li S, Looi J, et al.: Virtual humans as social actors: Investigating user perceptions of virtual humans' emotional expression on social media. Computers in Human Behavior 155, 108161 (2024).
4. Noh, Y.: A study on the discussion on Library 5.0 and the generation of Library 1.0 to Library 5.0. Journal of Librarianship and Information Science 55(4), 889-905 (2023).
5. Zhao, X.: Research on Factors Influencing the Rewarding Behavior of Virtual Anchors' Fans. In International Conference on Economic Management and Green Development. Springer, Heidelberg, 544-554 (2023).
6. Deng, X., Zhou, G., Xiao, B., et al.: Effectiveness evaluation of digital virtual simulation application in teaching of gross anatomy. Annals of Anatomy-Anatomischer Anzeiger 218, 276-282 (2018).
7. Cerasa A, Gaggioli A, Marino F, et al.: The promise of the metaverse in mental health: the new era of MEDverse. Heliyon 8(11): e11762 (2022).
8. Stein J P, Linda Breves P, Anders N.: Parasocial interactions with real and virtual influencers: The role of perceived similarity and human-likeness. New Media & Society 26(6), 3433-3453 (2024).
9. Venkatesh V, Morris M G, Davis G B, et al.: User acceptance of information technology: Toward a unified view. MIS Quarterly 27(3), 425-478 (2003).
10. Zhou T, Lu Y B, Wang B.: Integrating TTF and UTAUT to explain mobile banking user adoption. Computers in Human Behavior 26, 760-767 (2010).
11. Devaraj S, Easley R F, Crant J M.: How does personality matter Relating the five-factor model to technology acceptance and use. Information Systems Research 19(1), 93-105 (2008).
12. Lin J M ,Hung S, Chen C.: Fostering the determinants of knowledge sharing in professional virtual communities. Computers in Human Behavior 25(4), 929-939 (2009).
13. Tao, F., & Dang, Y.: Virtuality, simulation and fake: The technical development and philosophical criticism of virtual anchors. Prometeica-Revista de Filosofía y Ciencias (Especial), 36-51 (2022).
14. Zhang S, Liu H, Mei L, et al.: Predicting viewer's watching behavior and live streaming content change for anchor recommendation. Applied Intelligence, 1-16(2022).