SEARCH Journal of Media and Communication Research
Online ISSN: 2672-7080
© SEARCH Journal 2025
SEARCH 17(3), 53–68
DOI: 10.58946/search-17.3.P4
https://fslmjournals.taylors.edu.my/search/

Enhancing user engagement in virtual reality in Korea: A diegetic interface approach within the technology acceptance model

Stephen W. Song
University of South Florida, USA

*Donghun Chung
Kwangwoon University, Republic of Korea
donghunc@gmail.com

ABSTRACT

This study examines the impact of diegetic user interfaces on user experience in virtual reality environments, framing the investigation within the realms of the Technology Acceptance Model and Self-Determination Theory. Employing a 2 x 2 within-subjects experimental design, the research manipulates the presence of diegetic cues and the mode of score display (non-diegetic heads-up display versus diegetic integration) to assess their effects on user enjoyment. A bespoke virtual reality stimulus was created expressly for this experiment. The findings indicate that the inclusion of a diegetic user interface significantly enhances user enjoyment through a serial mediation effect. These results underscore the potential for diegetic design elements to improve virtual reality interfaces, suggesting that certain design strategies may significantly influence user satisfaction and acceptance. The study concludes with practical implications for virtual reality interface design, advocating for a strategic emphasis on diegetic elements to foster more immersive and enjoyable user experiences.

Keywords: ***Virtual reality, user experience, technology acceptance model, enjoyment, self-determination theory***

---

* Corresponding author

INTRODUCTION

The advent of virtual reality (VR) in the 1980s marked the beginning of a technological revolution, yet for decades, VR lingered on the periphery of mainstream adoption, perceived more as a technological curiosity than a transformative medium. This perception has shifted dramatically in recent years, catalysed by the global COVID-19 pandemic's social distancing mandates and the advent of accessible VR headsets like Oculus Quest and Vision Pro. These developments have expedited the migration of numerous traditionally physical activities into the virtual realm, culminating in the widespread popularization of the concept of metaverse (Song & Chung, 2021; Sparkes, 2021). Forecasts now project the VR gaming industry to escalate from a revenue of USD 8 billion in 2022 to an impressive USD 54 billion by 2033 (Research Nester, 2024).

In this burgeoning virtual landscape, the user experience within VR has emerged as a critical determinant of the medium's success (Rahman et al., 2024; Yao et al., 2024). The escalating importance of human factors in new and emerging media underscores the necessity of re-evaluating conventional design paradigms, especially within VR gaming (Kanasan, 2024). The immersive nature of VR, facilitated primarily through head-mounted displays, introduces unique challenges, including psychological discomfort (Yamaguchi, 1999), visual strain (Guo et al., 2019), and physical disorientation (Ito et al., 2021), necessitating a reassessment of user interface (UI) design principles from a human-centric perspective. By conceptualizing VR not merely as a technological innovation but as a psychological construct, we can uncover insights with far-reaching applicability beyond specific technological confines (Biocca, 1992).

This study investigates the impact of UI design on user engagement within VR, guided by the media richness theory. It posits that a well-designed UI enhances users' perceived ease of use and usefulness of the system, thereby elevating the overall enjoyment of the VR experience. To empirically test this hypothesis, a custom-designed VR whack-a-mole game was developed exclusively for this research. Through this investigation, the study aims to elucidate the influence of various UI designs on user perception and task performance within VR environments.

LITERATURE REVIEW

Virtual reality and gaming

VR is often defined as a particular technology system that includes computers, head-mounted displays (HMDs), headphones, and motion sensors (Steuer, 1992) or a world that is created using computer graphics that lets individuals to completely immerse themselves and interact with elements within the environment (Milgram & Kishino, 1994). Video games within VR using a head-mounted display (HMD) are an emerging trend (Yildirim et al., 2018) since people can easily immerse themselves in the virtual world compared to conventional video games that use two-dimensional display devices.

VR gaming has marked a significant evolution within the digital entertainment sector, transitioning from experimental beginnings to becoming a foundational element of mainstream gaming culture. Originating from the 1980s and 1990s pioneering experiments, VR gaming's journey toward immersive digital experiences took a pivotal turn with the 2016 introduction of consumer-grade VR headsets like HTC VIVE, Oculus Rift, and PlayStation VR. These devices significantly narrowed the gap between specialized experimentation and broad consumer engagement by providing immersive experiences previously unattainable with earlier, more rudimentary equipment. Notably, Oculus Quest 2 epitomizes the democratization of VR gaming, offering a standalone, high-quality experience that, by 2021, exceeded the combined sales of all its predecessors (Hector, 2021), illustrating the rapid technological advancements and increased accessibility that have transformed VR from a novelty into an integral component of digital entertainment.

The unique immersive capabilities of VR gaming offer an unparalleled sense of presence and immersion, extending its relevance beyond entertainment to encompass education, healthcare, and social interaction. These applications demonstrate VR's capacity to fundamentally alter our interaction with digital content (Rizzo et al., 2004; Slater & Sanchez-Vives, 2016), suggesting its broad utility beyond gaming. The ability of VR to simulate complex, interactive environments positions it as a critical tool for immersive experiences across various domains.

Modern VR gaming is distinguished by its ability to generate immersive 3D environments, facilitated by interactive gameplay mechanics, motion controllers, and head-mounted displays. Technological innovations have improved the fidelity and precision of VR gaming, integrating high-resolution displays, sophisticated tracking systems, haptic feedback, eye-tracking, and spatial audio (Bailenson, 2018). This immersive environment enables unique gameplay mechanics, fostering deeper engagement and social interaction through multiplayer games and virtual communities, thereby advancing creativity and innovation in game design. However, VR gaming faces challenges, including the high costs associated with quality equipment, the physical discomfort known as VR sickness (Munafo et al., 2017), and the potential exclusion of individuals unable to afford or physically engage with VR technology. Thus, overcoming these obstacles through continued innovation and enhanced accessibility is essential for VR gaming's ongoing growth.

In conclusion, VR gaming's evolution from experimental endeavours to a mature, widely accessible medium signifies substantial technological and conceptual progress. Despite existing challenges, the progressive development of VR technology and game design indicates a future of even more immersive and engaging experiences for users worldwide. Central to ensuring the continuous growth and broad accessibility of its benefits, the VR gaming industry must address the critical issues of cost, comfort, and inclusivity. It is important to look at user interface and media richness from a user interface and media richness perspective, as these elements are crucial in enhancing user experience and ensuring the medium's psychological impact and accessibility are maximized.

User interface and media richness

UI elements within VR can be categorized according to their diegetic relationship to the context of the environment and spatiality (Fagerholt & Lorentzon, 2009). Here, the diegetic dimension refers to the characteristics of a UI in terms of its existence in the game world. In other words, a UI is diegetic if the character within a mediated world is expected to be aware of the UI, while a UI is non-diegetic if the characters are expected to be unaware of the UI. An example of non-diegetic UI is subtitles or musical scores in movies, while a diegetic UI refers to elements such as showing the street sign in a movie scene, so both the audience and the movie characters can understand where the scene takes place.

Spatiality is another dimension of the UI which regards whether the UI element is visualized as a part of the environment of the game space or not. This means a UI element can take up a (virtually) physical space within an environment, depending on its spatiality. The diegesis and spatiality together form six types of UI within VR: heads-up display (HUD) elements (non-diegetic and non-spatial), geometric elements (non-diegetic and spatial), diegetic elements (diegetic and spatial), meta-representations (diegetic and non-spatial), meta-perception (non-diegetic or diegetic and non-spatial), and signifiers (diegetic and spatial) (Fagerholt & Lorentzon, 2009). Notably, a diegetic UI compared to a non-diegetic UI adds an overlay of information, thus enriching the mediated environment (Fagerholt & Lorentzon, 2009).

The media richness theory posits that the richness of media depends on its capability to carry nonverbal cues, provide rapid feedback, convey personality traits, and use of natural language (Daft & Lengel, 1986). According to this theory, richness of media is important since an appropriate amount of richness within media is related to its effectiveness, which leads to higher quality of outcomes (Kock, 2005). Prior research suggests that visual cues are related to less time consumption for tasks (Chen et al., 2018), or higher purchase intentions within video games (Shaouf et al., 2016).

We operationalize media richness as the joint presence of (a) multiple, task-relevant nonverbal cues (e.g., spatialized, directional indicators), (b) immediacy of feedback, and (c) contextual integration of information into the task environment according to Daft and Lengel's definition. In our manipulation, diegetic cues are in-world, spatial signages that provide immediate, direction-specific information about target location, thereby increasing richness by augmenting nonverbal cue bandwidth and feedback immediacy while maintaining context congruence (Fagerholt & Lorentzon, 2009). This contrasts with non-diegetic HUD elements, which increase information quantity, but may reduce context congruence. Therefore, the current study examines the efficacy of a HUD UI and diegetic UI, which provides varying amounts of information, by comparing its outcomes for a given task:

> H1: Richer information from UI elements is related to higher scores in a VR game.

Technology acceptance model

The technology acceptance model (TAM), as conceptualized by Davis and colleagues (1989), serves as a foundational framework for evaluating the adoption and sustained use of new information technologies. Grounded in the Theory of Planned Behaviour (TPB), TAM delineates a causal pathway from system design features through cognitive and affective responses to behavioural intention and actual system use (Ajzen & Fishbein, 2000). This model posits that those two primary cognitive beliefs—perceived ease of use and perceived usefulness—significantly influence an individual's attitude towards using a technology, which in turn, affects their behavioural intention and actual usage behaviours.

In the realm of virtual reality, particularly within the context of this study's focus on user interface design, TAM offers a nuanced lens through which to examine the factors that dictate user engagement with VR technologies. Specifically, the incorporation of diegetic user interfaces—interfaces that are integrated into the virtual environment and perceived by the user as part of the game world—presents a unique opportunity to explore how these elements affect the user's perceptions of ease of use and usefulness. Given that VR gaming's immersive nature demands interfaces that complement and enhance the user experience without breaking the sense of presence, the study investigates how diegetic and non-diegetic UIs influence these key determinants of technology acceptance.

Perceived ease of use in this context refers to the user's assessment of how intuitively they can interact with the UI elements, impacting their overall experience and task performance within the VR environment. Similarly, perceived usefulness pertains to the user's belief in the extent to which the UI contributes to achieving their goals within the game, such as improving task efficiency or enhancing the gameplay experience. Previous research across various domains, including e-commerce (Henderson & Divett, 2003), artificial intelligence (Kim et al., 2021), and video gaming (Sánchez-Mena et al., 2017; Wang & Goh, 2017), has underscored the mediating role of these cognitive beliefs in shaping users' attitudes and behaviours toward technology use.

Therefore, by applying TAM to the study of diegetic versus non-diegetic UIs in VR gaming, the research aims to elucidate how the informational richness and integration of UI elements influence perceptions of ease of use and usefulness, subsequently affecting user engagement and satisfaction. This exploration is pivotal in advancing our understanding of effective UI design in VR, promoting more immersive and enjoyable user experiences through informed interface strategies.

H2: Richer information from UI elements is related to higher ease of use.
H3: Richer information from UI elements is related to higher perceived usefulness.
H4: Ease of use predicts perceived usefulness.

Self-determination theory of enjoyment

Enjoyment is thought to be a central construct for entertainment, including video games. While different approaches to conceptualizing enjoyment exist, satisfaction of intrinsic needs is known to be one of the prominent frameworks for understanding the multi-faceted nature of enjoyment (Deci & Ryan, 2000; Tamborini et al., 2011).

The self-determination theory (SDT) is a theory of human motivation and satisfaction of needs, which leads to more enjoyment of an activity (Deci & Ryan, 2000). Among different types of needs, SDT categorizes two different types of needs, extrinsic and intrinsic needs. Extrinsic needs refer to those that are externally rewarded, such as compliments from others, scoring high on exams, and earning more money. Intrinsic needs comprise three basic needs which are autonomy, competence, and relatedness. SDT asserts that satisfaction of intrinsic needs is paramount to the enjoyment of an activity and psychological well-being. Autonomy refers to the feeling that an activity is internally derived and not influenced by outside forces. Competence is defined as a perception that an individual is capable and effective. Relatedness refers to a desire to be connected to other social beings (Deci & Ryan, 2000).

Understanding media enjoyment from the SDT perspective is particularly interesting since the intrinsic need satisfaction is related not only to enjoyment, but also motivation, which predicts future behaviour (e.g., repeated behaviour, and intention to use) in relation to the activity. According to the concept of SDT, previous investigations that employed enjoyment as need satisfaction using SDT found a significant relationship between elements of communication and enjoyment for different contexts such as video games (Tamborini et al., 2011), and social media use (Reinecke et al., 2014). Considering that enjoyment is considered as a form of attitude that affects volitional and spontaneous behavioural outcomes (Nabi & Krcmar, 2004), enjoyment as need satisfaction is employed as an effective response (i.e., attitude for the UI) within the TAM framework. Therefore, the following hypotheses are proposed:

H5: Richer information from UI elements is related to higher enjoyment.
H6: Ease of use predicts enjoyment.
H7: Perceived usefulness predicts enjoyment.

Together, we propose a theoretical framework based on TAM, as shown in Figure 1.

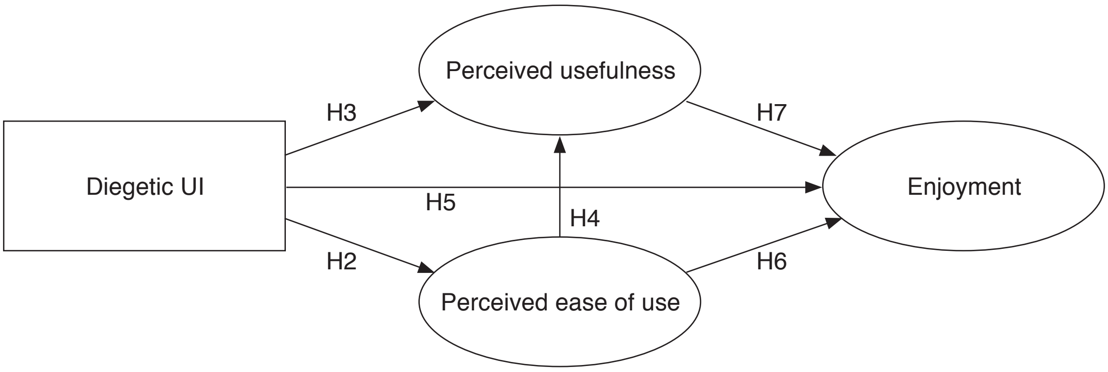

Figure 1. Theoretical model

METHODS

Participants

A total of 25 participants were recruited from a private university in South Korea. To prevent possible confounds due to prior experience with VR, only those with a minimal amount of experience with VR (less than an hour of prior VR use) were recruited. The age of participants that were recruited in this study ranged from 20 to 25 (M = 21.64, SD = 1.36). The majority of participants were male (n = 15; 60%).

Stimulus material

A whack-a-mole video game within VR was created specifically for this study. The game was made using Unity, and Oculus Integration Asset from Unity Asset Store was implemented to port a version for using HMD. Within the virtual environment, a donut-shaped table was placed in the centre of a large room. The outer diameter of the top of the table was designed to be approximately three meters if it existed in reality. The width of the tabletop (i.e., the diameter difference between the inner rim and the outer rim of the donut-shaped desktop) was approximately one meter. On the tabletop, eight holes for targets to appear for the whack-a-mole game were placed using equal spacing. The participant's avatar was located within the hollow centre of the table from the beginning to the end of the virtual experience. This meant that there was minimal need for the participant to walk around during the virtual experience, however, they had to turn around to locate a target in case it appeared from a hole behind the avatar.

Body tracking was implemented to follow the participant's body movement. Head tracking was the primary source to mirror the participant's overall location within the virtual environment. Hand tracking was also applied to follow either the left- or right-hand controller, depicted as a hammer within VR. The task within VR was to hit as many targets as possible by swinging the hammer against the target within a given time, which is typical for a whack-a-mole game. A target randomly appeared from one of the eight holes on the table. Only one target appeared at a time. The target disappeared when the user successfully whacked it with the hammer. If the user failed to hit the target within two seconds, the target automatically disappeared and a new target appeared from one of the holes.

For cue treatment within the diegetic UI factor, four signages that displayed the location of a target were placed within the virtual environment and around the table. The signage was approximately 1.5 meters in height. The display panel on top of the signage displayed four kinds of arrows: left, right, down, and circle. The shape and direction of the arrows changed depending on the current location of the target, to indicate that the target is on the left, right, below, or on the exact opposite side of the table at any moment. As for no cue treatment, the signage element was absent throughout the virtual task.

For HUD treatment within non-diegetic UI factor, HUD displayed a mini map of the table and indicated where the target appeared, while also presenting the current score and time left. For the diegetic display treatment, the aforementioned information was displayed on one of the four walls of the room in the backdrop. The virtual environment is shown in Figures 2, 3, 4, and 5.

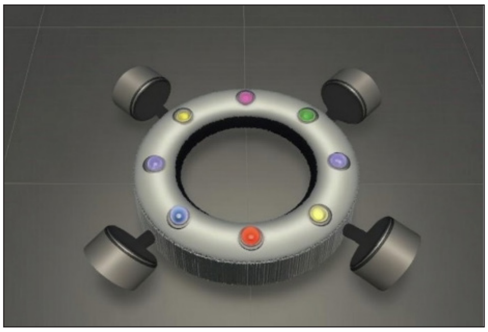

Figure 2. Top-view for the table and signage within VR

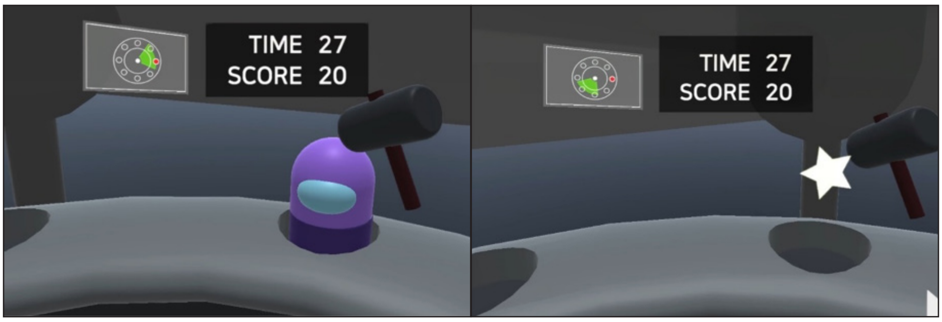

Figure 3. Interaction between target and the hammer on the tabletop

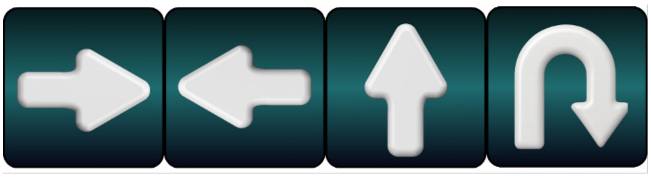

Figure 4. Arrows on the signage for the cue condition of diegetic UI factor

Figure 5. Score displays: non-diegetic HUD (left) and diegetic display (right)

Procedure

A 2 (diegetic UI: cue vs. no cue) x 2 (score display: non-diegetic HUD vs. diegetic display) within-subjects factorial design laboratory experiment was conducted to test the hypotheses that were proposed in this study. Upon entering the experimental room that was assigned for the experiment, the participants were asked to complete a consent form and a pre-survey questionnaire. The pre-survey questionnaire consisted of demographic information questions (i.e., age and sex). Once the participants completed the pre-survey, they were provided with background information on the experimental task. Participants were informed that the task was to play a whack-a-mole game within VR, with varying UI designs. Then, the participants underwent a tutorial session of the game, which lasted for 30 seconds.

Then, all participants experienced all four of the treatment conditions in **random order** to control for potential order effects. Each treatment session lasted for 100 seconds. After each session, participants were instructed to answer a brief post-exposure survey questionnaire that asked about their experience during the game. Participants were asked to rest for 2 minutes before starting the next session, a measure implemented to mitigate potential fatigue due to repeated VR exposure. After finishing all four sessions and post-exposure surveys, participants were dismissed. The experiment approximately took 30 minutes per participant.

Measures

Perceived ease of use measures consisted of three items, which were originally “intuitive controls” measures from player experience of need (PENS) inventory (Ryan et al., 2006). Perceived usefulness was an excerpt from a previous study on using TAM, which consisted of three items. Autonomy and competence items, which formatively constructed the concept of enjoyment, were adopted from the intrinsic motivation inventory (Plant & Ryan, 1985) which was modified for video game context in previous research (Tamborini et al., 2011). Enjoyment, conceptualized under SDT as the satisfaction of intrinsic needs, was therefore defined by these two distinct, yet related, reflective constructs. Autonomy refers to the feeling that an activity is internally derived, while competence is defined as a perception that an individual is capable and effective. Each of the intrinsic need constructs consisted of three measurement items. All measurement items were assessed on a 5-point Likert-type scale which ranged from "1 = Strongly disagree" to "5 = Strongly agree", and the score was operationalized by counting the number of successful hits on targets for each session. The reliability scores of the measures are demonstrated in Table 1 (Appendix).

FINDINGS AND DISCUSSION

A series of repeated measures ANOVAs were first conducted to investigate the relationship between the factors and dependent variables. Diegetic UI had a significant effect on score, $F(2, 24) = 4.14$, $p = .05$, $\eta^2 = .15$, while the effect of score display, $F(2, 24) = 2.55$, $p = .12$, $\eta^2 = .10$ and the interaction effect of the two factors, $F(2, 24) = .21$, $p = .65$, $\eta^2 = .01$ were not significant. Score was higher when the cue was displayed (M = 31.88, SD = 7.69) compared to when there was no cue (M = 30.02, SD = 8.54). Thus, H1 was partially supported.

The effect of diegetic UI was significant on perceived ease of use, $F(2, 24) = 6.14$, $p = .02$, $\eta^2 = .20$, perceived usefulness, $F(2, 24) = 10.03$, $p < .01$, $\eta^2 = .30$, and competence, $F(2, 24) = 5.91$, $p = .02$, $\eta^2 = .20$, but not autonomy $F(2, 24) = .53$, $p = .48$, $\eta^2 = .02$. Perceived ease of use was higher when cue was present (M = 4.47, SD = .66) compared to no cue (M = 4.02, SD = .92). Perceived usefulness was also higher when the cue was present (M =4.57, SD = .61) compared to no cue (M = 4.06, SD = .94). Competence was higher with the cue (M = 3.85, SD =1.00) compared to no cue (M = 3.56, SD = 1.13).

The main effect of score displays, as well as the interaction effect was not significant on any of the perceived ease of use, perceived usefulness, competence, or autonomy. The results of ANOVAs as well as comparison of means and standard deviations are demonstrated in Tables 2 and 3 (Appendix), respectively.

Table 2. Repeated measures ANOVA results

DV	Source	df	F	p	η^2
Score	Diegetic	(1, 24)	4.14	.05	.15
	Spatial	(1, 24)	2.55	.12	.10
	Diegetic*Spatial	(1, 24)	0.21	.65	.01
Perceived ease of use	Diegetic	(1, 24)	6.14	.02	.20
	Spatial	(1, 24)	3.37	.08	.12
	Diegetic*Spatial	(1, 24)	0.45	.51	.02
Perceived usefulness	Diegetic	(1, 24)	10.03	<.01	.29
	Spatial	(1, 24)	0.43	.52	.02
	Diegetic*Spatial	(1, 24)	2.97	.10	.11
Autonomy	Diegetic	(1, 24)	0.53	.48	.02
	Spatial	(1, 24)	0.92	.35	.04
	Diegetic*Spatial	(1, 24)	0.16	.69	.01

Table 2. (con't)

DV	Source	df	F	p	η^2
Competence	Diegetic	(1, 24)	5.91	.02	.20
	Spatial	(1, 24)	1.78	.20	.07
	Diegetic*Spatial	(1, 24)	0.67	.42	.03

Note. Diegetic = diegetic UI factor, Spatial = score display factor

Since ANOVA results revealed that diegetic UI had significant main effects on most dependent variables while the main effects of score display or the interaction effects were not significant, subsequent analyses were conducted focusing on diegetic UI. Thus, among the four combinations of treatments, those with the same treatments for diegetic UI were averaged for subsequent analyses.

First, the direct and mediation effect between diegetic UI and the dependent variables were tested using MEMORE (Montoya & Hayes, 2017), a script for SPSS that conducts mediation analysis for two-condition within-subjects design data. There was a significant effect of diegetic UI on perceived ease of use (β = . 447, p = . 020) and perceived usefulness (β = . 507, . p = . 042). Therefore, H2 and H3 were supported. However, diegetic UI did not significantly predict enjoyment (β = .190, p = . 092). Therefore, H5 was not supported. The result showed that the mediation effect of the path "diegetic UI → perceived ease of use → perceived usefulness", "diegetic UI → perceived usefulness → enjoyment", and "diegetic UI → perceived ease of use → enjoyment" was significant. Details of the mediation analysis are presented in Table 4.

Table 4. Mediation analysis results

Mediation paths	β^a	SE	Lower CI^b	Upper CI^b
Diegetic UI → PEoU → PU	. 300	. 142	. 065	. 626
Diegetic UI → PU → Enjoyment	. 212	. 093	. 027	. 394
Diegetic UI → PEoU → Enjoyment	. 112	. 069	. 005	*. 264*

Note. PEoU = perceived ease of use, PU = perceived usefulness, a = Unstandardized path coefficient, b = 95% interval

Next, PLS-SEM (partial least square structural equation modelling) was employed to test the effect of diegetic UI as a design element within TAM. PLS-SEM was chosen instead of Covariance Based Structural Equation Modelling (CB-SEM) since PLS-SEM is more applicable to experiments with smaller sample sizes and is suitable for exploratory studies (Hair et al., 2011). The experimental design included two within-subjects factors, and only diegetic UI had a significant effect on dependent variables. Thus, the subsequent PLS-SEM analysis incorporated the difference in the dependent variable values between the two diegetic UI treatment conditions (cue vs. no cue), following the analysis methods employed in previous studies by Shin et al. (2021).

Model fit indexes that are conventionally used in CB-SEM are covariate-based, but the Chi-square-based model fits are largely inapplicable to PLS-SEM (Hair et al., 2019). Instead, the following indicators were tested to see if they were within the recommended threshold ranges (Hair et al., 2019). Most indicator loadings were above the ideal range (≥ .708) and some were above the acceptable range (≥ .500), as shown in Table 5 (Appendix).

The collinearity of the indicators was assessed using variance inflation factor (VIF), and most were within the ideal range (< 3) while some were within the acceptable range (< 5), which is demonstrated in Table 6 (Appendix).

The reliability of the reflective latent variables was tested using Cronbach's alpha (≥ .70) and composite reliability (CR; ≥ .80) and were above the ideal range. The convergent validity of latent variables was assessed using average variance extracted (AVE; ≥ .50), and all four reflective latent variables were above the recommended threshold, as shown in Table 1. The discriminant validity (heterotrait-monotrait ratio; HTMT) of reflective latent variables were below the recommended range (< .85) as shown in Table 7 (Appendix).

The reliability of a formative construct is usually evaluated through its theoretical support (Moody et al., 2014). In the current study, enjoyment was introduced in the model as a formative construct, which was formative of two reflective constructs, autonomy and competence. Since autonomy and competence were theoretically formulated and empirically tested as distinct subconstructs of satisfaction of intrinsic needs (i.e., enjoyment) (Deci & Ryan, 2000) and have been applied in the video game context before (Tamborini et al., 2011), we concluded that enjoyment was indeed a formative construct of autonomy and competence.

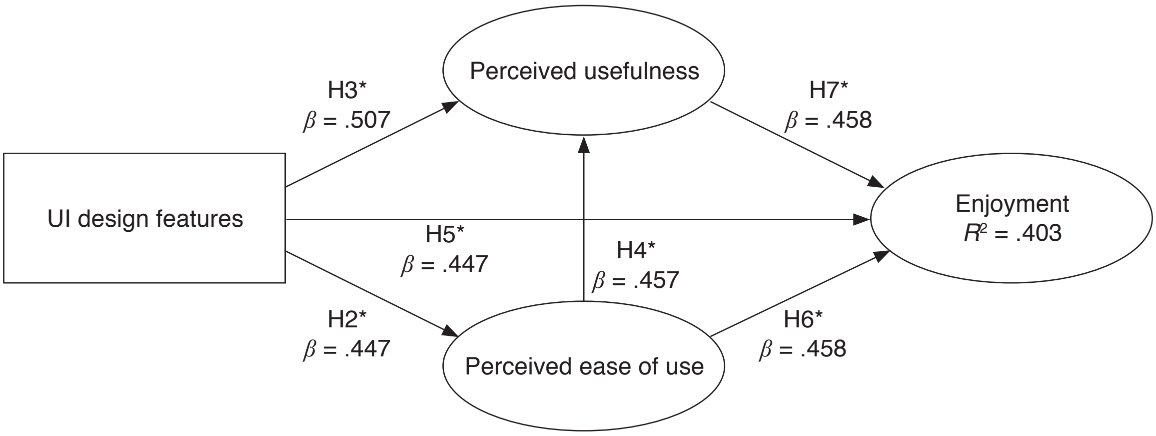

Figure 6. Demonstration of results

Since the indicators of the PLS-SEM model were within the recommended range, the relationships between latent constructs were investigated. Here, the path from perceived ease of use to perceived usefulness was significant ($\beta = .457$, $p = .033$). Therefore, H4 was supported. However, enjoyment was not significantly predicted by perceived ease of use ($\beta = .278$, $p = .070$). Therefore, H6 was not supported. The path from perceived usefulness to enjoyment was significant as well ($\beta = .458$, $p = .019$). Thus, H7 was supported. Together, the model explained the difference in enjoyment due to diegetic UI moderately ($R^2 = .403$). The final SEM result of the current study is demonstrated in Figure 6.

This study investigated the effect of UI design based on its diegetic and spatial dimensions within TAM. The results mostly supported our predictions. Diegetic UI is overall effective in improving perceived ease of use, perceived usefulness, and enjoyment. The mediation effect result suggests that while diegetic UI does not have a direct relationship to the enjoyment of a virtual experience, it is serially mediated by perceived ease of use and perceived usefulness. This result identifies the mechanism or pathway through which the design element successfully influences user affect, emphasizing the critical role of cognitive assessment as defined by TAM. This result should be interpreted with caution because the significant mediation reflects model-consistent associations rather than strict causality.

This suggests that the introduction of additional information was indeed helpful for the user to better enjoy the virtual experience. However, the effect of HUD UI was not significant on any of the usability indicators. This does not necessarily mean that HUD UI is less meaningful than diegetic UI, rather it suggests that HUD UI may have been ineffective in our implementation of HUD in the context that was specific to the current study.

It is possible to argue that the result of the current study is merely because the diegetic UI functioned as a guidance system, which was useful and easy to use for users. However, this finding offers a critical distinction from previous research. Previous studies have found that typical guidance within a videogame undermines enjoyment since it harms the autonomy of the user (Ryan et al., 2006). Our investigation showed that the introduction of diegetic UI cues did not significantly undermine autonomy, yet enjoyment was ultimately improved via cognitive acceptance (PEOU and PU).

This strongly suggests that the diegetic UI functioned as a legitimate improvement on user experience, effectively delivering instructive information without the psychological cost of reduced self-determination. This contrasts sharply with TAM-based VR research that has struggled to implement guidance without negative intrinsic motivational consequences. Prior VR-TAM work typically emphasizes presence, perceived usefulness, and ease of use as predictors of intention under HUD-dominant or system-level UI assumptions. Our findings instead show that an in-world (diegetic) cue—not simply more on-screen information—improves performance and perceived ease of use, with downstream associations to usefulness and enjoyment. This pattern highlights context congruence as a key design lever, beyond merely increasing information density.

The findings provide specific theoretical implications for VR UX design, too. Future VR interfaces should adopt diegetic design principles that integrate information seamlessly into the virtual world. This approach ensures that information is readily accessible, improving the user's perceived competence and ease of use, thereby boosting overall enjoyment. The key design principle derived is that designers must implement instructive elements in a manner that preserves the user's sense of autonomy and immersion, which the diegetic approach successfully achieved in this study.

CONCLUSION

In conclusion, the findings in the current study suggest that certain types of UI might be more effective in conveying information and thus enhancing the experience within interactive VR environments. This study contributes to the virtual reality and TAM literature, and provides insight into the industry, considering that diegetic UI design may lead to a better perception of commercially designed VR games.

There are several limitations to this study from which future directions of research can be derived. First, the small sample size and the use of PLS-SEM limit the statistical power and certainty of the findings. Additionally, the sample consisted of homogeneous

Korean undergraduates, limiting the cross-cultural validity and generalizability of our results. While we controlled for prior VR familiarity, we did not control for general prior gaming experience or analyse potential gender differences. Thus, future research should attempt to use a between-subjects design to replicate this study, incorporating a larger and more diverse sample to identify these individual differences. In addition, while the effect of UI was conceptualized to be due to media richness, this aspect was not empirically tested. Thus, future studies should extend the current finding by clarifying the psychological underpinnings of diegetic UI and exploring why exactly diegetic UI affects user's experience within VR.

Finally, while relatedness within SDT posits that the social dimension of intrinsic needs is important for enjoyment, the current study was incapable of testing this aspect since our stimuli did not have any social components. Therefore, future studies should extend current findings by adding social activity within VR.

Open Access: This article is distributed under the terms of the Creative Commons Attribution License (CC-BY 4.0) which permits any use, distribution and reproduction in any medium, provided the original author(s) and the source are credited.

References

Ajzen, I., & Fishbein, M. (2000). Attitudes and the attitude-behavior relation: reasoned and automatic processes. *European Review of Social Psychology, 11*(1), 1–33.

Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. WW Norton & Company.

Biocca, F. (1992). Communication within virtual reality: Creating a space for research. *Journal of Communication, 42*(4), 5–22. https://doi.org/10.1111/j.1460-2466.1992.tb00810.x

Chen, T., Wu, Y. S., & Zhu, K. (2018). Investigating different modalities of directional cues for multi-task visual-searching scenario in virtual reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (pp. 1–5). Association for Computing Machinery.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5), 554–571.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management Science, 35*(8), 982–1003.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.2307/1449618

Fagerholt, E., & Lorentzon, M. (2009). *Beyond the HUD-user interfaces for increased player immersion in FPS games* [Master's thesis]. Chalmers University of Technology, Sweden.

Guo, J., Weng, D., Zhang, Z., Liu, Y., Duh, H. B. L., & Wang, Y. (2019). Subjective and objective evaluation of visual fatigue caused by continuous and discontinuous use of HMDs. *Journal of the Society for Information Display, 27*(2), 108–119.

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139–152. https://doi.org/10.2753/MTP1069-6679190202

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24.

Hector, H. (2021, April 1). Oculus Quest 2 sales figures prove VR has finally gone mainstream. *TechRadar*. https://www.techradar.com/news/oculus-quest-2-sales-figures-prove-vr-has-finally-gone-mainstream

Henderson, R., & Divett, M. J. (2003). Perceived usefulness, ease of use and electronic supermarket use. *International Journal of Human-Computer Studies, 59*(3), 383–395. https://doi.org/10.1016/S1071-5819(03)00079-X

Ito, K., Tada, M., Ujike, H., & Hyodo, K. (2021). Effects of the weight and balance of head-mounted displays on physical load. *Applied Sciences, 11*(15), 6802.

Kanasan, M. (2024). The virtual playground: Exploring action video games among adolescents. *SEARCH Journal of Media and Communication Research, 16*(4), 89–103. doi: 10.58946/search-Special Issue.ICOMS2023.P7

Kim, J., Merrill Jr, K., & Collins, C. (2021). AI as a friend or assistant: The mediating role of perceived usefulness in social AI vs. functional AI. *Telematics and Informatics, 64*, 101694. https://doi.org/10.1016/j.tele.2021.101694

Kock, N. (2005). Media richness or media naturalness? The evolution of our biological communication apparatus and its influence on our behavior toward E-communication tools. *IEEE Transactions on Professional Communication, 48*(2), 117–130. https://doi.org/10.1109/TPC.2005.849649

Milgram, P., & Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems, 77*(12), 1321–1329.

Montoya, A. K., & Hayes, A. F. (2017). Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychological Methods, 22*(1), 6 –27. https://doi.org/10.1037/met0000086.

Moody, G. D., Galletta, D. F., & Lowry, P. B. (2014). When trust and distrust collide online: The engenderment and role of consumer ambivalence in online consumer behavior. *Electronic Commerce Research and Applications, 13*(4), 266–282. https://doi.org/10.1016/j.elerap.2014.05.001

Munafo, J., Diedrick, M., & Stoffregen, T. A. (2017). The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental Brain Research, 235*, 889–901.

Nabi, R. L., & Krcmar, M. (2004). Conceptualizing media enjoyment as attitude: Implications for mass media effects research. *Communication Theory, 14*(4), 288–310.

Plant, R. W., & Ryan, R. M. (1985). Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: An investigation of internally controlling styles. *Journal of Personality, 53*(3), 435–449.

Rahman, A. A., Murad, K., & Adnan, W. H. (2024). The acceptance and use of virtual reality on learning attitude among junior boy scouts in Kuala Lumpur, Malaysia. *SEARCH Journal of Media and Communication Research, 16*(4), 123–136. doi: 10.58946/search-Special Issue.ICOMS2023.P9

Reinecke, L., Vorderer, P., & Knop, K. (2014). Entertainment 2.0? The role of intrinsic and extrinsic need satisfaction for the enjoyment of Facebook use. *Journal of Communication, 64*(3), 417–438.

Research Nester. (2024, February 1). Virtual reality in gaming market revenue to hit USD 54 billion by 2033 says Research Nester. *GlobeNewswire*. https://www.globenewswire.com/en/news-release/2024/02/01/2821791/0/en/Virtual-Reality-in-Gaming-Market-revenue-to-hit-USD-54-Billion-by-2033-says-Research-Nester.html

Rizzo, A. A., Schultheis, M., Kerns, K. A., & Mateer, C. (2004). Analysis of assets for virtual reality applications in neuropsychology. *Neuropsychological Rehabilitation, 14*(1–2), 207–239.

Ryan, R. M., Rigby, C. S., & Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion, 30*(4), 344–360.

Sánchez-Mena, A., Martí-Parreño, J., & Aldás-Manzano, J. (2017). The effect of age on teachers' intention to use educational video games: A TAM approach. *Electronic Journal of e-Learning, 15*(4), 355–366.

Shaouf, A., Lü, K., & Li, X. (2016). The effect of web advertising visual design on online purchase intention: An examination across gender. *Computers in Human Behavior, 60*, 622–634.

Shin, M., Lee, S., Song, S. W., & Chung, D. (2021). Enhancement of perceived body ownership in virtual reality-based teleoperation may backfire in the execution of high-risk tasks. *Computers in Human Behavior. 115*, 106605. https://doi.org/10.1016/j.chb.2020.106605

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI, 3*, 74.

Song, W., & Chung, D. (2021). Explication and rational conceptualization of Metaverse. *Informatization Policy, 28*(3), 3–22. doi:10.22693/NIAIP.2021.28.3.003

Sparkes, M. (2021). What is a metaverse. *New Scientist, 251*(3348) (p.18).

Steuer, J. (1992). Defining virtual reality: Dimensions determining telepresence. *Journal of Communication, 42*(4), 73–93. https://doi.org/doi:10.1111/j.1460-2466.1992.tb00812.x

Tamborini, R., Grizzard, M., David Bowman, N., Reinecke, L., Lewis, R. J., & Eden, A. (2011). Media enjoyment as need satisfaction: The contribution of hedonic and nonhedonic needs. *Journal of Communication, 61*(6), 1025–1042. https://doi.org/10.1111/j.1460-2466.2011.01593.x

Wang, X., & Goh, D. H.-L. (2017). Video game acceptance: A meta-analysis of the extended technology acceptance model. *Cyberpsychology, Behavior, and Social Networking, 20*(11), 662–671.

Yamaguchi, T. (1999). Physiological studies of human fatigue by a virtual reality system. *Presence: Teleoperators & Virtual Environments, 8*(1), 112–124.

Yao, L., Tugiman, N., & Sharipudin, M. N. S. (2024). Virtual human influencers in live streaming commerce on social media platforms: Exploring parasocial interactions with consumers in China. *SEARCH Journal of Media and Communication Research, 16*(4), 47–59.

Yildirim, C., Carroll, M., Hufnal, D., Johnson, T., & Pericles, S. (2018, August). Video game user experience: to VR, or not to VR?. In *2018 IEEE Games, Entertainment, Media Conference (GEM)* (pp. 1–9). IEEE.

Stephen W. Song (Dr)

is an assistant professor in the Department of Journalism and Digital Communication at University of South Florida, USA. He earned his PhD in Mass Communications at Syracuse University, NY, USA. His research interest is in the application of psychological theories on new and emerging communication technologies.

Donghun Chung (Dr)

received his PhD in Communication from Michigan State University, USA. He is a professor in the School of Media and Communication at Kwangwoon University, Seoul, South Korea, and his research examines human experience, with an emphasis on cognitive science, human–computer interaction, and the diffusion of innovation.

Appendix

Table 1. Cronbach's Alpha (α), CR, and AVE of latent constructs.

Reflective Construct	Cronbach's α	CR	AVE
Autonomy	.776	.872	.696
Competence	.719	.845	.648
Perceived ease of use	.912	.944	.849
Perceived usefulness	.886	.929	.814

Note. CR = Composite reliability, AVE = Average variance extracted

Table 3. Means and standard deviations

	Cue & HUD		No cue & HUD		Cue & No HUD		No cue & No HUD	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Score	31.44	7.84	28.96	8.95	32.32	7.69	31.08	8.15
Perceived ease of use	4.60	.66	4.08	.90	4.33	.65	3.96	.95
Perceived usefulness	4.65	.60	4.03	.96	4.45	.61	4.09	.94
Autonomy	3.69	.79	3.64	.84	3.8	.76	3.68	.86
Competence	3.76	.96	3.53	1.12	3.95	1.04	3.59	1.15

Table 5. Factor loadings and p-values

Item	Factor loading	Item	Factor loading
AUTO1	.897**	PEOU1	.932**
AUTO2	.882**	PEOU2	.889**
AUTO3	.711**	PEOU3	.942**
COMP1	.894**	PU1	.948**
COMP2	.827**	PU2	.824**
COMP3	.678**	PU3	.929**

Note. AUTO = autonomy, COMP = competence PEoU = perceived ease of use, U = perceived usefulness, $^{**}p < .01$.

Table 6. Variance inflation factors for items

Item	VIF	Item	VIF
AUTO1	2.572	PEOU1	3.865
AUTO2	2.348	PEOU2	3.066
AUTO3	1.285	PEOU3	2.835
COMP1	2.423	PU1	3.930
COMP2	2.185	PU2	1.930
COMP3	1.188	PU3	3.513

Note. AUTO = autonomy, COMP = competence PEoU = perceived ease of use, U = perceived usefulness.

Table 7. Heterotrait-monotrait ratio for reflective constructs

Reflective Construct	1	2	3	4
1. Autonomy				
2. Competence	.835			
3. Perceived ease of use	.339	.725		
4. Perceived usefulness	.467	.812	.494	