國立臺灣大學電機資訊學院資訊工程學系

碩士論文

Department of Computer Science and Information Engineering

College of Electrical Engineering and Computer Science

National Taiwan University

Master’s Thesis

理解沉浸式體驗中使用者的認知狀態轉變

Living a Fantasy: Understanding the User’s Cognitive State Transitions in Immersive Experience

張庭逸

Ting-Yi Chang

指導教授: 鄭龍磻 博士

Advisor: Lung-Pan Cheng, Ph.D.

中華民國 113 年 8 月

August, 2024

doi:10.6342/NTU202402686

國立臺灣大學碩士學位論文
口試委員會審定書

MASTER'S THESIS ACCEPTANCE CERTIFICATE
NATIONAL TAIWAN UNIVERSITY

理解沉浸式體驗中使用者的認知狀態轉變

Living a Fantasy: Understanding the User's Cognitive State Transitions in Immersive Experience

本論文係張庭逸君（學號 R09922065）在國立臺灣大學資訊工程學系完成之碩士學位論文，於民國 113 年 7 月 25 日承下列考試委員審查通過及口試及格，特此證明。

The undersigned, appointed by the Department of Computer Science and Information Engineering on 25 July 2024 have examined a Master's thesis entitled above presented by CHANG, TING-YI (student ID: R09922065) candidate and hereby certify that it is worthy of acceptance.

口試委員 Oral examination committee:

程芙茵 (Jul 25, 2024 15:12 GMT+8)

（指導教授 Advisor）

系主任/所長 Director:

Acknowledgements

感謝我的指導老師鄭龍磻老師，一路上給了我探索研究方向上極大的自由，同時也給予許多指導，讓我能更了解做學術研究的心態，這真的是研究所生涯最重要的收穫。在研究過程中，即使遇到挫折與迷惘，龍哥的鼓勵和身為過來人的分享都讓我受益良多，知道這些都是必經之路，使我更坦然的面對當下的自己。在給我們這麼大自由的同時，也給了很多資源，讓我能夠專注於研究，真的很謝謝龍哥。

感謝學位考試的口試委員詹力韋老師及程芙茵老師，在口試時對這篇論文提供許多肯定。芙茵老師從認知心理學角度給予的建議，以及力韋老師基於 HCI 領域專業給予的建議，都對這篇論文的未來方向有很大的助益。感謝張永儒老師在研究前期所提供的寶貴意見，對於使用者體驗的改善有著相當重要的幫助。

感謝陳毅，從你的研究中讓我得到這篇研究最初的啟發。感謝奕碩、瑋儒，在建置實驗的實作上給予的幫助，使我能夠專注於設計，沒有後顧之憂。感謝 Cosmos（柏堯）給予的各種想法上的激盪，還有對於分析結果的幫助。感謝柏昱、宗儒、子昂，以及實驗室的大家，在研究過程中給予的想法和建議，都讓這篇論文更加完整，也謝謝大家對我在實驗時佔據了大部分實驗室空間的包容。

最後感謝我的父母和女朋友佳霓，在精神上和經濟上的支持，才能讓我走到這一步，謝謝你們。

doi:10.6342/NTU202402686

摘要

傳統上，對於使用者在沉浸式環境中主觀體驗的研究，通常集中於存在感，也就是身處於虛擬環境的感覺。我們建議將認知和感知分開，以便在現實與虛擬之間界限的日益模糊的未來，更好的理解更多樣化沉浸式體驗中使用者的主觀體驗。我們以「存在感」來描述使用者的感知體驗，並引入「認知狀態」，即指使用者相信自己正在與物理世界或虛擬世界互動的信念，來描述使用者的認知系統。我們假設使用者的認知狀態會在沉浸式體驗中轉變。我們設計了一個包含一系列高和低保真度感官線索的 30 分鐘密室逃脫體驗，來探究使用者的認知轉變及其誘發因素。通過對 36 名參與者的主題分析結果，我們概述了影響使用者認知轉變的六個因素：虛擬可信度、空間意識、感知、保真度、期望和任務順序，並提出了六條在設計沉浸式體驗上，建立使用者虛擬認知狀態的原則。

關鍵字：認知、臨場感、沉浸式體驗、虛擬實境

doi:10.6342/NTU202402686

Abstract

Traditionally, the study of users’ subjective experiences in immersive environments has typically focused on presence, which is the sense of being there. Considering the evolving forms of immersive experiences and the increasingly blurred boundaries between reality and virtuality, we suggest separating cognition and perception to better understand users’ subjective experiences. We retain “presence” to describe users’ perceptual experiences and introduce “cognitive state,” i.e., believing in either reality or virtuality, to describe the users’ cognitive systems. We hypothesize that users’ cognitive states will transition during immersive experiences. We design an experiment consisting of a series of sensory cues with high and low fidelity in a 30-minute escape room experience to probe the user’s cognitive transitions and the inducing factors. From the thematic analysis result of 36 participants, we outline the six factors affecting the user’s cognitive transitions: virtual plausibility, spatial awareness, sensation, fidelity, expectation, and task order, offering six guidelines for establishing users’ virtual cognitive state in an immersive experience.

Keywords: Cognition, Presence, Immersive Experience, Virtual Reality

doi:10.6342/NTU202402686

Contents

doi:10.6342/NTU202402686

doi:10.6342/NTU202402686

doi:10.6342/NTU202402686

List of Figures

doi:10.6342/NTU202402686

doi:10.6342/NTU202402686

List of Tables

doi:10.6342/NTU202402686

Chapter 1

Introduction

Immersive experiences are popular methods in entertainment, education, and other fields. From early movies and theaters to theme parks, escape rooms, Live Action Role Playing(LARP), and virtual reality/mixed reality(VR/MR), various types of immersive experiences have been introduced. Among these, VR/MR is likely the most attractive subject for researchers, leading to the development of numerous frameworks to better understand how users interact with VR/MR. Milgram and Kishino's reality-virtuality continuum conceptualized Mixed Reality (MR) with three dimensions: extent of world knowledge, reproduction fidelity, and extent of presence metaphor. Skarbez et al. [27] revisited the model, arguing that only "Matrix-like" virtual environments, which provide both exteroceptive and interoceptive senses through technologies like BCI, fall outside the scope of MR.

Research on users' subjective experiences in VR has traditionally focused on presence. The term "telepresence" was initially proposed in 1980 [19], and the concept of presence has since been broadly recognized as the sense of "being there." With advancements in VR technology, various models and measurement methods have been developed

doi:10.6342/NTU202402686

to gain a deeper understanding of presence.

By imagining a future with “Matrix-like” VR technology, we discovered that beyond the perceptual experiences covered by traditional presence, users’ cognition can still be influenced by other factors, such as recalling the moment they entered the virtual world. We are not the first to recognize the importance of cognition; Latoschik and Wienrich [14], as well as Schubert [25], have previously incorporated cognition into their presence models. In comparison, Hartmann and Hofer [12] and Hofer et al. [13], drawing on cognitive psychology, systematically emphasized the fundamental differences between cognition and perception, proposing a parallel processing model of cognition and presence.

In this paper, we also draw on the dual-process theory from cognitive psychology to suggest separating cognition and perception when understanding users’ subjective experiences in immersive experiences. We keep the term “presence” to describe users’ perceptual experiences, in line with Slater’s definition, defining presence as a perceptual sense of being in an environment [28, 31], while we use the term “cognitive state” to represent users’ beliefs of the world they are interacting with. Compared to Hartmann and Hofer’s concept of media awareness [12], our definition of cognitive state shifts the focus from mediated experience to the user’s subjective identification. This creates a more comprehensive space for discussing users’ subjective experiences in multimodal immersive experiences and better fits the increasingly blurred boundary between virtuality and reality in the future.

We hypothesize that cognitive state transitions occur in immersive experiences. To test our hypothesis, we conducted a user study (n=36) by designing an immersive escape room experience in which participants perceive different sensory cues with high and low

doi:10.6342/NTU202402686

fidelity and in different sequences.

Considering the dominant role of vision in the perception, we designed a two-stage immersive experience to better observe the influence of the two hypothetical factors and minimize visual interference; in the preparation stage, the participant establishes their initial cognitive state while being able to see, whereas in the experience stage, they proceed in a blind state, focusing on tactile and auditory sensory cues. This two-stage design allows us to observe the differences between having vision and being blind and also to study how other factors influence cognitive state transitions when the overpowering vision factor is removed. After the immersive experience, we conduct questionnaires and a semi-structured interview to evaluate the participants' cognitive state transitions and explore the factors contributing to their cognitive state transitions. Using thematic analysis, we identified four themes related to cognitive state: virtual scenario, spatial recognition, sensory, and expectation. We discuss the interaction between these four themes and extract six factors of cognitive state: virtual plausibility, spatial awareness, sensation, fidelity, expectation, and order of tasks. Finally, we provide seven guidelines for future MR experience design.

doi:10.6342/NTU202402686

Chapter 2

Related Work

Living a Fantasy relates to (1) VR framework, (2) Presence, (3) Dual Process Theory, and (4) Blurring Boundaries Between Reality and Virtuality

2.1 VR Framework

Virtual Reality(VR) is one of the most commonly used terms to describe the system that creates a virtual environment that allows users to immerse in it. Some models are proposed to frame VR from different perspectives. Milgram and Kishino's reality–virtuality continuum conceptualized Mixed Reality(MR) into three dimensions, extent of world knowledge, reproduction fidelity, and extent of presence metaphor [18]. Based on Milgram and Kishino's continuum, Skarbez et al. revisited the model by combining two dimensions, reproduction fidelity and extent of presence metaphor, into one dimension - immersion, and introduced Coherence as a new dimension [27]. Oppose to Milgram and Kishino' s continuum, Rauschnabel et al. separate AR and VR, using local presence as a key of AR and telepresence as a key of VR [23].

doi:10.6342/NTU202402686

Considering that even if VR technology reaches the vision of Ultimate display, the system still only controls the user's exteroceptive senses (i.e., sight, hearing, touch, smell, and taste), but not interoceptive senses (such as vestibular and proprioceptive senses), Skarbez et al. argue that any form of technology-mediated reality falls within the scope of MR, where sensory conflict remain due to the system restricted [27]. They then borrowed the concept from the movie "The Matrix" to propose a "Matrix-like" virtual environment. In "The Matrix," the consistency of exteroceptive and interoceptive senses is achieved through direct brain stimulation. They argue that such "Matrix-like" virtual environments, created through BCI (Brain-Computer Interface), are the only types of virtual environments that lie outside the scope of MR.

2.2 Presence

The term *telepresence* was proposed in 1980 [19]; after that, the concept of presence has been broadly used to describe the subjective experience in the context of VR since the 1990s [3, 37, 43]. Presence is conceptualized as a perceptual illusion [15, 16, 30], indicating the sense of "being there" in a mediated virtual environment [26, 33, 37].

As more dimensions of meaning are added to the concept of presence, terms have been introduced to differentiate sub-concepts, such as *spatial presence(telepresence)* [19, 25, 42] and *social presence(copresence)* [15, 21]. To clarify the concept of presence and its constitution, different presence frameworks have been proposed: Wirth et al. [42] proposed a theoretical two-step process model of spatial presence. In the first step, users construct *spatial situation model(SSM)*, a mental model of the mediated virtual environment. In the second step, users decide whether to accept SSM as their *primary egocentric ref-*

doi:10.6342/NTU202402686

erence frame(PERF), and if they do so, they present in the mediated virtual environment. Slater [29] introduces *Place Illusion(PI)* and *Plausibility Illusion(Psi)* as two orthogonal illusions that constitute presence. PI(some called spatial presence or telepresence) indicates the illusion of ”being there”, and Psi(some called coherence) indicates the illusion that what happening in the virtual environment is really happening. Skarbez et al. [26] reviewed existing models and categorized the factors in these models, and proposed a conceptual model based on *Place Illusion*, *Plausibility Illusion*, and *Social Presence Illusion*. Social Presence Illusion(some called social presence or copresence) denotes the illusion of being together with other real people in a virtual environment. Zahorik and Jenison [45] introduce an alternative view of presence, linking it to successfully supported action in an environment, whether virtual or real. Usoh et al. assess the effectiveness of the WS [43] and SUS questionnaires [32, 39] in distinguishing real from virtual experiences and concluded that while their utility for comparing experiences across different environments is doubtful, such questionnaires may be useful when subjects experience the same type of environment.

2.3 Dual Process Theory

The two types of thinking, one fast and intuitive and the other slow and deliberate, have long existed across various domains. The dual-process theory related to higher cognitive processes originated in the 1970s [41]. Dual-process theory(also known as dual-system theory) in cognitive and social psychology separates two processing modes, known as Type 1 processing and Type 2 processing [8, 35]. Type 1 processing is fast, automatic, and non-conscious, characterized by heuristic and intuitive processes. In contrast, Type 2 processing is slow, controlled, and conscious, involving analytical and reflective thinking.

doi:10.6342/NTU202402686

Two different models explain how these two types of processing operate during decision-making: parallel-competitive and default-interventionist. Parallel-competitive model assumes the two types of processing operate independently to deliver responses [34, 35]. In this model, Type 1 and Type 2 processing are regarded as distinct, independent processing that do not inherently influence each other; instead, they engage in competition when conflicts arise to deliver a response. Default-interventionist model assumes that in most cases Type 1 processing produces default intuitive responses, but in certain situations, Type 2 processing will be invoked to intervene in the default response [9, 10].

2.4 Blurring Boundaries Between Reality and Virtuality

Theoretical and technological developments both indicate that as people's imaginations about the future expand, the boundaries between reality and virtuality become increasingly blurred.

Mixed-fantasy framework proposes a three-anchor framework by adding imaginary reality as the third dimension to Milgram and Kishino's real-virtual continuum [36]. Imaginary Reality Gaming [1] provides a practical example of applying imaginary reality in games, showcasing the potential for new types of experiences. Building on imaginary reality, Reality Rifts [6] further explores how to guide and stimulate users' imagination, detailing the design space for employing imaginary reality. Programmable matters [11] envision a future where media modules are more indistinguishable from the physical world in sensory terms. This vision is supported by advancements in technology; responsive materials [44] and microscale swarm robots [7] offer precise methods for controlling physical properties. DreamCeption [40] points out the possibility of lucid dreams as a form of

doi:10.6342/NTU202402686

“Matrix-like” virtual environment.

On a philosophical level, the postmodern view also suggests that the boundary between the virtual and the real is increasingly blurred. This perspective introduces the concept of hyperreality to describe the collapse of distinctions between constructs such as place, body, and distinguishable material and intellectual entities [20, 24]. As these distinctions erode, the clear boundaries between reality and virtuality dissolve, leading to a new paradigm where the two realms are intertwined and indistinguishable.

doi:10.6342/NTU202402686

Chapter 3

From Presence to Cognitive State

In this section, we will present how and why we suggest distinguishing cognitive state from presence in the discussion of users' subjective experience, and the potential to extend the understanding of users' subjective experience by exploring users' cognitive state transition in MR experience.

Presence, as we presented previously, is widely accepted as a perceptual illusion of being in a mediated virtual environment [16, 26, 30]. Although the concept of presence effectively captures the subjective state of users in current MR interactions, in our subsequent thought experiment exploring the future singularity of VR technology, we discovered a space of the user's subjective experience that falls beyond the scope of presence. We imagine that even if VR systems achieve "Matrix-like" VR in the future, there are still factors beyond presence that influence users' belief in whether they are interacting with the physical or virtual world. Two possible factors not included in the concept of presence are: 1. When users recall the moment they entered the VR system and 2. When someone informs them they are in VR and they believe it. We consider that this belief in interacting with the physical or virtual world is a state belonging to the user's cognition.

doi:10.6342/NTU202402686

Some previous literature has attempted to incorporate cognition into the discussion on presence. In the model proposed by Latoschik and Wienrich [14], cognition, as one of the foundational layers (alongside sensation and perception layer), influences congruence and plausibility, thereby triggering the emergence of presence. Schubert [25] conceptualizes spatial presence as a conscious cognitive feeling, representing a functional outcome from users' unconscious cognitive processes.

Drawing on cognitive psychology, some literature emphasizes the fundamental differences between cognition and presence when exploring their relationship. Adapting a dual-process psychological logic(mentioned in section 2.3), Hofer et al. [13] consider spatial presence as a lower-order, intuitive, nonconscious cognitive processing (Type 1 processing) and plausibility as a higher-order, reflective, conscious cognitive processing (Type 2 processing). Hartmann and Hofer [12] adopt the dual-process theory and propose a processing model of users' subjective experience in mediated experience, which consists of two parallel processes: presence and media awareness. They consider presence as Type 1 processing, and media awareness as Type 2 processing which denotes the users' belief that what they perceive is from human-made technology. They adopt the parallel-competitive model within the dual-process theory, asserting that presence and media awareness operate independently without influencing each other.

In our model, we acknowledge presence as the Type 1 processing in the dual-process theory; however, concerning Type 2 processing, our emphasis lies in the users' belief of the world that they are interacting with, which we call cognitive state. The cognitive state we propose shares similarities with the concept of media awareness presented by Hartman and Hofer [12] in that both can be used to describe the user's awareness of being in a virtual environment. However, the cognitive state can involve the states of interacting with the

doi:10.6342/NTU202402686

physical world, interacting with the virtual world, or simultaneously interacting with both. Moreover, the concept of cognitive state focuses on the user's subjective identification with the virtual world. This expansion shifts the focus from whether users are interacting with a mediated environment to their subjective experiences (without judgment), which could better accommodate the increasingly blurred boundary between reality and virtuality in the future.

The roles of presence and cognitive state in our model are closer to the default-interventionist model rather than the parallel-competition model. Cognitive state, as a Type 2 processing, may interact with presence, which contradicts the parallel-competition model. While there is no direct evidence, empirical research in psychology suggesting interrelations between perception and cognition [5] may serve as indirect evidence indicating that cognitive state and presence could influence each other.

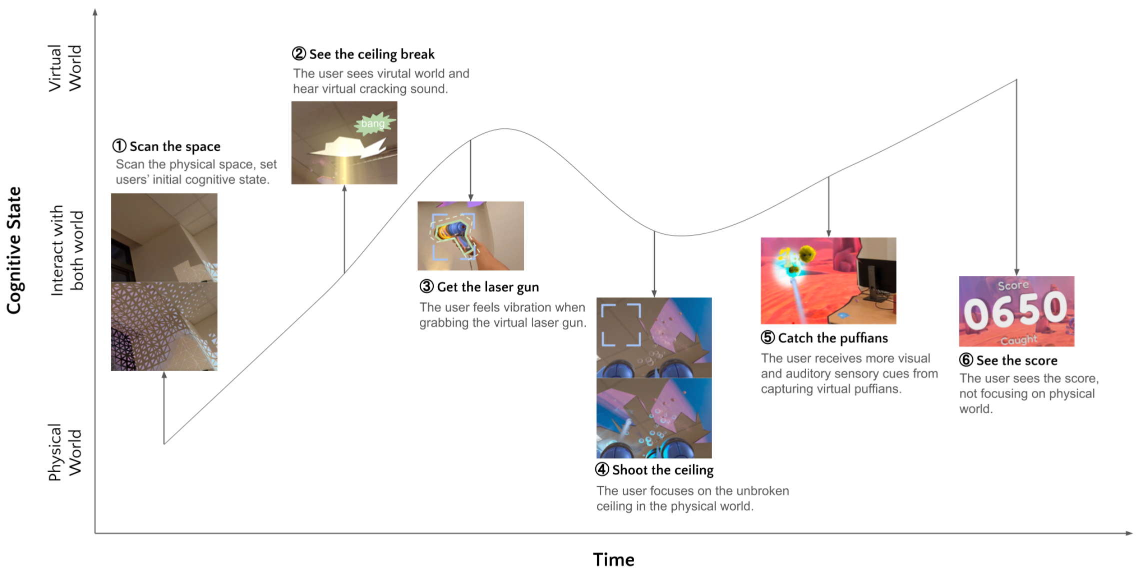

Figure 3.1: Event that affects the transition of the cognitive state during an experience. Taking "First Encounters" [17] as an example, we identify six events within the experience that could potentially trigger cognitive state transition, and mark the differences in sensory cues that users receive when these transitions occur.

Figure 3.1 illustrates how cognitive state transitions during an immersive experience.

doi:10.6342/NTU202402686

Due to the substantial interaction between the player and both the physical and virtual environments in this game, we take "First Encounters" [17] as the example to explain how a user's cognitive state might transition. We identified six events in the game where we felt cognitive state transitions: scan the space, see the ceiling break, get the laser gun, shoot the ceiling, catch the puffians, and see the score. After reviewing these events, we highlight the difference in sensory cues within each event and the sequence of events, considering them as factors that influence the transition of cognitive state. Firstly, we believe that the source of sensory cues (physical or virtual) has a significant impact on changing our cognitive states, and the main criterion for determining the source of sensory cues is their fidelity. Moreover, whether the user actively or passively receives these sensory cues can also influence their cognitive state. Compared Event 2 and Event 4 in Figure 3.1, the main difference, besides the laser gun, is that the user actively cause the ceiling break in Event 4. Therefore, the feeling of interaction with this virtual event(ceiling break) is stronger, potentially influencing the user's cognitive state to transition more toward the virtual world. Secondly, the assumption about the impact of the sequence of events comes from comparing Event 3 and Event 6. These two events essentially involve interacting only with virtual content. However, the cognitive state at Event 2 leans more towards reality, while the cognitive state at Event 5 leans more towards the virtual, which may result in the cognitive state of Event 6 being closer to the virtual.

To sum up, we take the differences in sensory cues—fidelity, active or passive reception, and the sequence of tasks, as three primary independent variables in our experiment, which will be further discussed in 4.

doi:10.6342/NTU202402686

Chapter 4

Method

Extending from the thought experiment, we initially aimed to compare how participants with different initial cognitive states would experience cognitive state transitions when receiving identical sensory cues.

4.1 Development Process of Experiment Design

Considering the dominant role of vision in the human sensory system [22, 38], we designed a two-stage experience: the preparation stage and the experience stage. In the preparation stage, the participant establishes their initial cognitive state while being able to see, whereas in the experience stage, they proceed in a blind state, focusing on tactile and auditory sensory cues. This two-stage design allows us to observe the differences between having vision and being blind and also to study how other factors influence cognitive state transitions when the overpowering vision factor is removed.

In the preparation stage, we envisioned interactions in the following different scenarios to establish various initial cognitive states: virtual world, physical world, reality with

doi:10.6342/NTU202402686

Figure 4.1: The virtual scene and the physical space for the preparation stage of the original design.

virtual items, and virtual world with real objects. Considering the execution difficulty and the number of control variables, we chose the two extremes of cognitive states: the virtual world and the physical world. In the initial experiment's preparation stage, one group of participants interacted using a VR headset, while the other group performed the same interaction in reality. As shown in Figure 4.1, in the preparation stage, we replicated the virtual environment into the physical room.

In the experience stage, participants will be required to complete several tasks, during which both groups will receive the same variety of auditory and tactile sensory cues.

During the implementation of the experience, we discovered significant challenges in creating a blind situation for the reality group in the experience stage. We tried two methods: creating an environment with no light and having participants wear blindfolds between the preparation stage and the experience state. The former was difficult to implement and made it hard to observe participants. The latter disrupted the experience too much according to participant feedback, and it introduced more differences between the reality and VR groups. Both methods failed to achieve the original objectives of our experimental design. To reduce the complexity, we removed the reality group and split the VR group into two groups. Each group experienced tasks in different sequences to test

doi:10.6342/NTU202402686

whether the order of tasks would affect cognitive state transitions.

We refined our experimental design to focus solely on the three independent variables mentioned in chapter 3, fidelity, active or passive reception, and sequence of tasks. We streamlined the process by eliminating the reality group, ensuring greater consistency and control over variables. Next, we detail the experimental procedures that were used to systematically investigate the impact of different sensory cues and the sequence of tasks on cognitive state transitions.

4.2 Apparatus

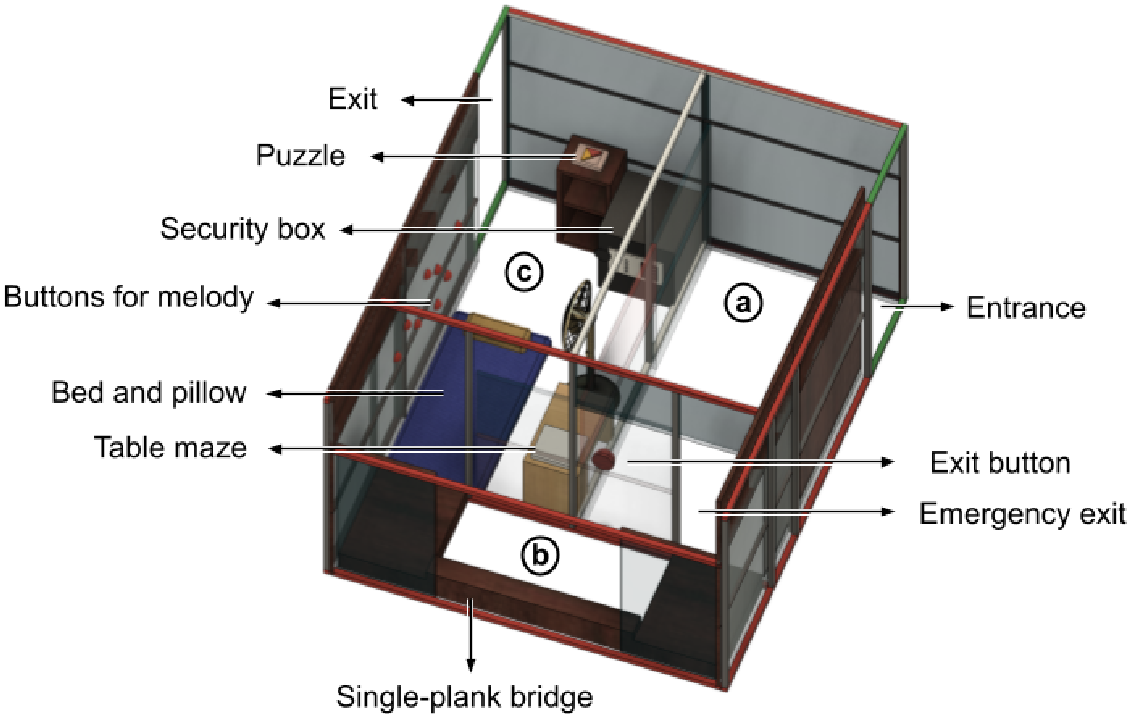

Figure 4.2: The 3D model of our study space design. Preparation and Situation 1 are conducted in room a, Situation 2 in room b, and Situation 3 in room c.

In Figure 4.2, the 3D model shows the design of our study's setup. We build a 3.6m × 4.5m × 2.7m space consisting of 3 rooms by Laminated Veneer Lumbers(LVL) and Plywood. Preparation and Situation 1 are conducted within room a, while Situation 2 and

doi:10.6342/NTU202402686

3 respectfully took place in rooms b and c.

Our VR system comprised an HTC Vive Pro headset, 4 steamVR lighthouses with participants wearing bHaptics Tactsuit [2]. In order to record participants' actions for further discussion, we set 2 GoPro Hero 8 at the ceiling of the space and on participants' haptic vest, and two staff recorded videos on their phones. To provide auditory sensory cues, we installed a Bluetooth speaker on the ceiling and another behind the wall with buttons in Room c. These speakers played ambient music and sound effects. A partici-pant wears transparency-mode wireless Bluetooth headphones, which allow them to hear environmental sounds and communicate with the instructor and the agent system. Props such as the table maze and the security box in Figure 4.3 were made using laser cutting and 3D printing. We used Arduino boards to connect sensors and actuators, such as 15 buttons, to our computer system to control the stimuli throughout the entire experience.

4.3 Procedure

Our final experimental design consisted of a 30-minute escape room immersive experience followed by a 30-minute questionnaire and semi-structured interview segment to test our hypothesis about participants' cognitive state transitions.

In the escape room experience, a participant plays the role of a recruit for a secret agent organization, undergoing training and testing before joining the organization. To correspond to different scenarios, we further divided the original experience stage; thus, the final experiment consists of four stages: Preparation, Situation 1, Situation 2, and Situation 3. Continuing from the previous subsection, we divided the participants into two groups. Both groups established an initial VR cognitive state but performed the tasks

doi:10.6342/NTU202402686

in different sequences in Preparation and Situation 3, as shown in Figure 4.3. In these various tasks, participants received different levels of sensory cues, which we will detail in section 4.4.

Before Preparation

Before entering the experimental space, the participant is required to wear the equipment provided by the secret agent organization, including a VR headset, wireless Bluetooth earphones, a haptic vest, and working gloves. After that, they are guided into the experimental space in a blind state.

Preparation

At the start of the experiment, in a cloakroom modified by the special agent organization, the participant sees their instructor. Under the instructor's guidance, the participant completes several tasks to develop the skills needed to become an agent.

The purpose of this stage is to establish the participant's initial VR cognitive state through these immersive interactions.

Situation 1

After completing the training, the instructor informs the participant that they need to pass a final test—stealing a document from a millionaire's room—to join the agent organization. To avoid detection, the participant must complete the task with the power cut-off, following the instructor's guidance and with the assistance of the agent system. After explaining the mission, the instructor leaves to cut off the power and then directs

doi:10.6342/NTU202402686

the participant to find the button that opens the emergency exit leading to the billionaire's room.

In this stage, we cut off the power to the VR headset, placing the participant in a blind situation for the subsequent stages. The participant will explore the space they previously observed but without vision, since Preparation and Situation 1 share the same room(Figure 4.2a), which participants fully observed during the preparation stage.

Situation 2

After exiting through the emergency exit, the participant finds themselves on the balcony of a tall building, as explained by the instructor via Bluetooth wireless earphones. Following the spatial descriptions provided by the instructor, the participant located the single-plank bridge and proceeded along it to the balcony outside the billionaire's room.

In this stage, the participant traverses through an new space(Figure 4.2b) following the instructor's command without vision. The instructor describes the appearance and layout of the space to guide participants to go across the single-plank bridge.

Situation 3

After reaching the balcony outside the billionaire's room, the instructor reiterates the participant's mission and the usage of the agent system. The participant then uses the agent system to query the mission and follows the system's directions to complete a series of tasks sequentially, ultimately stealing the documents and safely leaving the billionaire's room.

In this stage, the participant needs to explore the new space (Figure 4.2c) without any

doi:10.6342/NTU202402686

prior spatial hints and without vision. They must complete five tasks sequentially, each accompanied by different auditory and tactile sensory cues.

4.4 Fidelity of Sensory Cues and Tasks

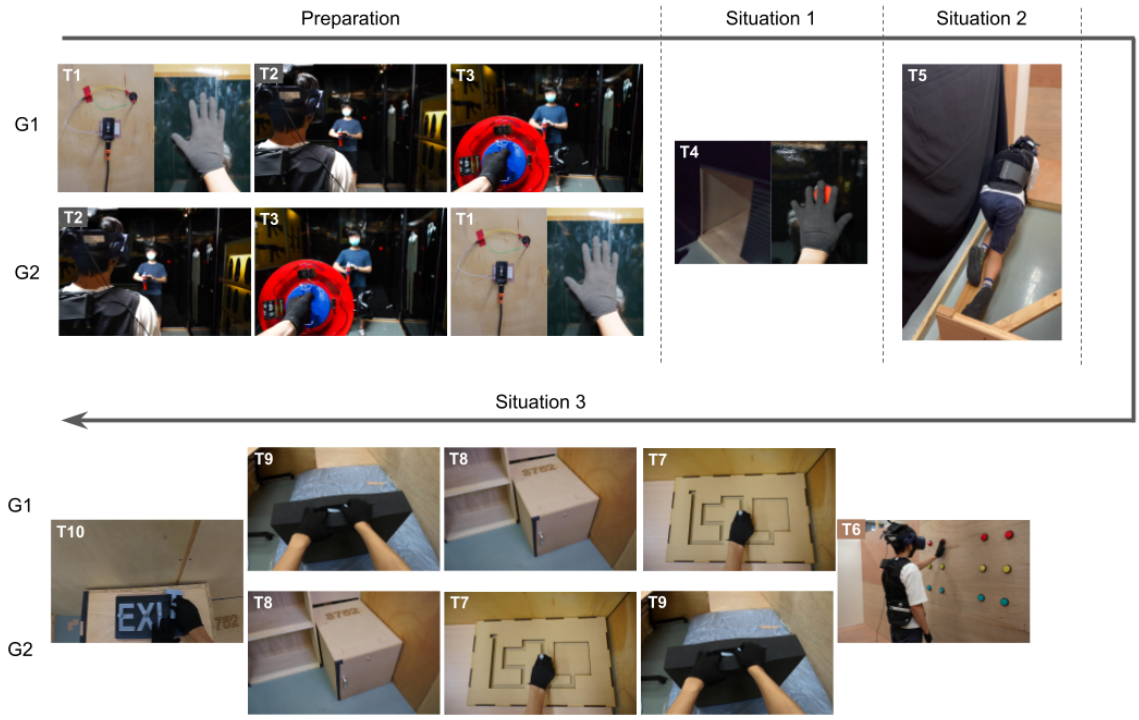

Figure 4.3: All tasks participants performed during the experience and the different sequences for the two groups. (T1) logging in the agent system, (T2) blocking bullets by a shield, (T3) hit by bullets, (T4) pressing the exit button to open the door, (T5) passing through single-plank bridge, (T6) pressing buttons to form melody, (T7) solving the table maze, (T8) finding out the password to open the security box, (T9) looking for letters, (T10) solving puzzle.

Haptic Cues		Auditory Cues	
H1	Shot by bullets	A1	Background music (speaker)
H2	Vibrations of the shield	A2	Wind sound (speaker)
H3	Vibrations through the wall	A3	Sound of agent system (headphones)
H4	Press buttons	A4	Instructions from instructor (direct voice)
H5	Pass the single-plank bridge	A5	Instructions from instructor (headphones)
H6	Feel the soft material of the bed	A6	Clicking when pressing buttons
H7	Control objects to navigate through a maze	A7	The sound of a power outage (headphones)
H8	Touch the emerging code	A8	The sound of an iron door opening
H9	Open the wooden box	A9	The music triggered by pressing the button (speaker)
H10	Retrieve letters from inside the pillow	A10	The background noise of a fan
H11	Assemble 3D-printed letters into the recessed slots	A11	The sound of bullets hitting the vest
		A12	The sound of gunfire

Table 4.1: All haptic and auditory cues in our MR experience. "H" stands for haptic cues, while "A" stands for auditory cues in the article.

doi:10.6342/NTU202402686

In our study, we distinguish the auditory and tactile cues as either high or low fidelity. Each task incorporate different types of sensory cues. High fidelity signifies participants perceive the effect that actually happen in the experience; for example, participants are actually hit by bullets and they hear the sound of bullets making contact with their vest. Low fidelity represents that participants perceive what happen in the MR experience through a form of simulation; for instance, participants open the exit by pressing the exit button but heard the sound of a metal door being opened from a speaker. This sound is considered a low-fidelity cue in our study. In the entire immersive experience, there are high and low-fidelity sensory cues in both auditory and haptic sensory cues. Moreover, we classify auditory cues into signal and ambient sound and haptic cues into active and passive cues. The former is distinguished by whether they are triggered by specific events. The latter is differentiated by whether they can control the level of interaction, content, and sensation.

Figure 4.3 illustrate all the tasks and the sequence in which the two groups of participants execute these tasks. Table 4.1 lists all haptic and auditory cues in our MR experience. Table 4.2 shows the relation between tasks and both auditory and haptic sensory cues, categorizing these sensory cues as high or low fidelity, signal/ambient, active/passive in our experiment.

Logging in the agent system. Because participants perceive vibration from the wall(H3), which is made by vibration motors and controlled by our system, it is low fidelity and active. The background music(A1), sound of the agent system(A3), and instructions from the instructor(A4) are classified as high fidelity because, in the experience, there are indeed a speaker playing music, an agent system, and an instructor speaking.

Blocking bullets by a shield. Participants perceive vibration through the vibration

doi:10.6342/NTU202402686

Haptic		
	Active (Controlled by system)	Passive (Controlled by user)
High Fidelity	H1, H8	H4, H5, H6, H7, H11
Low Fidelity	H2, H3	H9, H10

Auditory		
	Signal	Ambient
High Fidelity	A3, A4, A5, A6, A11, A12	A1, A10
Low Fidelity	A7, A8, A9	A2

	Tasks	Haptic cues	auditory cues
T1	*Logging in agent system	H3	A3, A4, A1
T2	*Blocking bullets by a shield.	H2	A4, A12, A1
T3	*Hit by bullets	H1	A4, A11, A12, A1
T4	Press exit button to open the door	H4	A7, A8, A5, A1
T5	Passing through single-plank bridge	H5	A2
T6	*Pressing buttons to form melody	H4, H6	A9, A3, A6, A11, A12, A1, A10
T7	*Solving the table maze	H7	A3, A1, A10
T8	*Finding out the password to open the security box	H9, H8	A3, A1, A10
T9	*Searching for letters for the puzzle	H10, H6	A3, A1, A10
T10	Solving the puzzle to complete all drills	H11	A3, A1, A10

Table 4.2: The categorization of the sensory cues as high/low fidelity, signal/ambient, active/passive. The codes refers to the sensory cues in Table 4.1. And, both the auditory and haptic sensory cues in each task in the MR experience. Sensory cues are ordered in low to high. (*) denotes that the order of these tasks is altered in the Alternatively group. ”H” stands for haptic cues, while ”A” refers to auditory cues in the article.

motor we attach on the shield(H2), which is low fidelity and active. The sound of gunfire(A12) is high fidelity because it is made by a Nerf gun firing.

Hit by bullets. Participants are actually hit by a Nerf gun bullets(H1), so it is high fidelity and active, since it is controlled by our system. The sound of bullets impacting the vest (A11) is also classified as high fidelity since it is the actual sound of the bullet hitting the vest.

Pressing exit button to open the door. Participants press the physical button(H4) we attach, therefore, it is high fidelity and passive, since participants have control over how they press the buttons themselves. The sound of the power outage(A7) and door opening(A8) are low fidelity because they are simulated.

Passing through single-plank bridge and bed. Participants walk or crawl on the physical single-plank bridge(H5) we built and could control how to pass the bridge. Therefore, it is classified as high fidelity and passive. Meanwhile, they heard of the simulated ambient sound of wind(A2).

doi:10.6342/NTU202402686

Pressing buttons to form melody. They press physical buttons(H4), which make high fidelity clicking sound from the buttons(A6), while the sound from the speaker simulate piano sound(A9), which is a low-fidelity auditory cue. While pressing the button, participants will stand on the bed(H6) we placed in Figure 4.2 room c, which provides a real soft material feeling. The background noise(A10) of a fan is from a fan we place in the room, so it is a high-fidelity ambient sound.

Solving the table maze. Since participants controlled the object to navigate in a table maze(H7), the haptic cue is categorized as high fidelity and passive.

Finding out the password to open the security box. Opening the wooden box(H9) is classified as low fidelity and passive because the wooden box is a simulation to a metal security box. Touching the emerging code of the box(H8) is high fidelity and active since our system control the code they touched.

Searching for letters for the puzzle. We 3D-printed a letter and hid it inside a pillow(H10), which is simulated by a sponge, making it a low-fidelity passive haptic cue.

Solving the puzzle. Participants placed the letters they collected in the 3D-printed puzzle(H11) classified as a high-fidelity passive haptic cue.

To avoid misunderstanding, we further explain why the speaker sounds in A1 are classified as high fidelity, while those in A2 and A9 are classified as low fidelity. Background music in a completely realistic experience would naturally be played through speakers, making A1 a high fidelity cue. However, the sounds of a door opening and the sound trigger by buttons would not typically come from speakers in a real-world scenario, thus A2 and A9 are categorized as simulated low fidelity cues. The classification of auditory cues from the headphones (A3, A5, A7) into different categories follows a similar rationale.

doi:10.6342/NTU202402686

4.5 Questionnaire and Interview

Building on Skarbez et al.'s suggestion to use single-item measures to simply assess presence and Bouchard et al.'s validation of the effectiveness of single-item measures, we designed a 7-point Likert scale question to measure participants' cognitive states. In the questionnaire, for each of the ten tasks, two questions were asked: (1) Where do you believe the entire experience up to this point took place? (1: real world, 7:virtual world), and (2) Where do you believe this task took place? (1: real world, 7:virtual world). After the questionnaire, we conducted a semi-structured interview to ask more general questions during the experience. The interview focused on three topics: sensory cues, fidelity, and expectation.

4.6 Participants

We recruited 36 participants (18 male, 18 female), ages 20 to 39 (mean=24.22, SD = 3.46).

For participants' prior experience with VR, 1 weekly, 5 monthly, 5 about once half a year, and 20 used VR less than 5 times, 6 never. In order to understand if different types of cues in different orders will affect participants' cognitive state transition, we separated participants into Continuously (G1) and Alternatively (G2) groups, as Figure 4.3, by how we order the high and low-fidelity sensory cues in Preparation and Situation 3.

doi:10.6342/NTU202402686

4.7 Data Analysis

For data analysis, we conducted an inductive thematic analysis [4]. During the analysis, we utilized the advantages of Atlas.ti[1] in connecting and comparing codes. The preliminary code list was generated by selecting segments from interview transcripts that were helpful in our understanding of participants' cognitive state transitions during the experimental process. Subsequently, the research team iteratively reviewed all codes, making necessary revisions to individual codes or entire code groups as needed. When revisions occurred, the research team cross-referenced other codes to examine their applicability. Throughout this process, we established a more rigorous structure for code groups and, based on this framework, segmented codes into appropriate themes. Finally, through group discussions, we further refined the themes, ensuring consensus among research team members on the identified theme.

[1] https://atlasti.com

doi:10.6342/NTU202402686

Chapter 5

Results

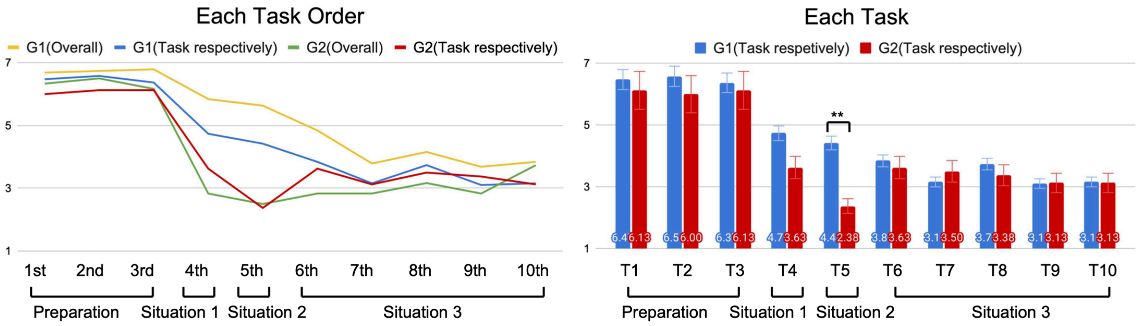

Figure 5.1: Left: The average cognitive state scores for G1 and G2 task sequences (see Figure 4.3). Right: The average cognitive state scores and standard deviations for G1 and G2 across each task. "Overall" and "task respectively" correspond to the two questions in section 4.5. (**) denotes p<.05.

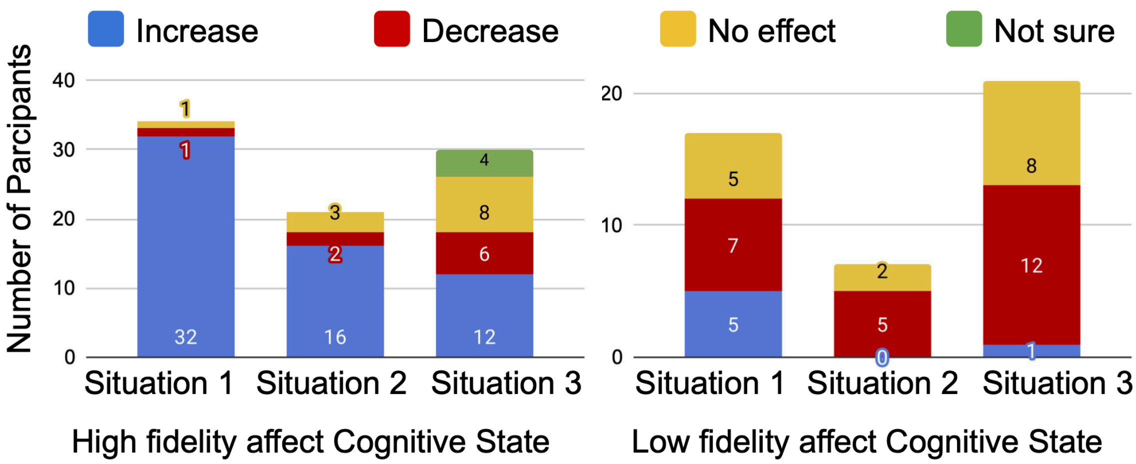

Figure 5.2: The number of participants report that whether high or low fidelity sensory cues affect their cognitive state during the interview.

doi:10.6342/NTU202402686

Figure 5.1(left) illustrates the average cognitive state rating on a 7-point Likert-scale across different task orders, differentiating between the Continuously (G1) and Alternatively (G2) groups, and whether participants rated the entire experience so far or only the specific task they were engaged in(see section 4.5). The score of cognitive state gradually decreases(lean to the physical world) as the number of tasks being completed in both Continuously(G1) and Alternatively(G2) groups, while the score for the Alternatively group(G2) decreased more than the Continuously group(G1) in the 4th task, which is finding the exit button and press it.

Figure 5.1(right) shows the average cognitive state score on a 7-point Likert-scale across different tasks, differentiating between the Continuously (G1) and Alternatively (G2) groups. The rating in the task "Passing through single-plank bridge and bed" shows a significant difference ($p<.05$) between the Continuously (G1) and Alternatively (G2) groups.

Figure 5.2 represents the number of participants reporting how high and low-fidelity sensory cues affected their cognitive state. We exclude participants who have not been mentioned in this regard.

5.1 Virtual Scenario

We consider the establishment of a reliable virtual scenario as a crucial factor in shaping the cognitive state, as P26 stated, "*As long as the virtual scenario is sufficiently well-constructed, I feel like I'm inside VR.*" Ten participants (P11, P12, P13, P14, P15, P18, P20, P26, P30, P36) reported that experiencing alignment with the virtual scenario make their cognitive state transition to the virtual world. In contrast, eight participants(P1,

doi:10.6342/NTU202402686

P2, P4, P8, P25, P27, P33, P35) reported that when something inconsistent with the context appeared or occurred, their cognitive state transitioned to the physical world. P8 also reported that virtual scenario contributes to maintaining the cognitive state after the deprivation of vision: "*The preceding conditions create a feeling as if I'm already situated within that environment, so to some extent, I've already been convinced that I'm inside a virtual reality. Later on, even without seeing that thing, I still have an inner sense of persuasion.*"

We have identified three factors conducive to creating a robust virtual scenario.

5.1.1 Impact of sensory cues on virtual scenario

In terms of the sensory aspects, four participants (P12, P17, P26, P33) reported that receiving instructions through headphones enhanced their immersion in the virtual scenario. P26's statement highlights that sensory cues improve the virtual scenario because they originate from the virtual environment: "*I accept instructions because I'm inside the virtual scenario, so receiving instructions reminds me that I'm inside thevirtual scenario.*" P35's statement emphasizes that simultaneously perceived sensory cues from both the virtual world and the real world enhanced the virtual scenario: "*After I traced that shape (the code) and answered correctly, the safe door opened, and the agent system said, 'Mission accomplished.' I thought the whole thing went pretty smoothly, with the physical and virtual aspects blending in seamlessly. So, I felt like I could really get into that scenario.*" Three participants (P13, P15, P18) also reported that high-fidelity sensory cues enhanced the virtual scenario (see subsection 5.3.5). Moreover, even low-fidelity sensory cues were noted to transition the virtual scenario to the virtual world as long as they align with the virtual scenario. When asked what made them particularly feel like it was VR,

doi:10.6342/NTU202402686

P20 replied, "*The shield [...] because holding something in my hand that I don't know what it is, probably not a real shield, but it makes me feel like it fits the current scene.*"

5.1.2 Impact of emotions on virtual scenario

Four participants (P6, P8, P12, P35) reported that emotions, such as tension, anxiety, and anticipation, could enhance their immersion in the virtual scenario. P12 mentioned, "*(While crossing the single-plank bridge) I did try to let go of the handrail, and that's when I felt nervous, which made me more immersed in this virtual scenario.*" P17 reported an interesting observation that strong emotions might lead to detachment of virtual scenario, "*Similar to watching a horror movie, this scenario makes me feel really scared, indicating that I'm deeply immersed in it. In such cases, I need to start telling myself that it's all fake, using my cognition to help me detach from that scenario.*"

5.1.3 Impact of cognitive allocation on virtual scenario

P6, P26, and P35 reported that less focus on virtual scenario could lead to a detachment on virtual scenario. P26 stated, "*When I'm solving the table maze, I tend to focus more on what I'm currently touching, and I'm less immersed in the virtual world's environment.*" P35 stated, "*When I allocate a significant portion of my cognitive resources to solving puzzles, I may also experience a slight detachment from the VR environment.*"

5.1.4 The interaction between virtual scenario and fidelity

It is interesting that P27 reported virtual scenario can supply sensory fidelity 5.3.5: "*[...] and the safe, even if you actually put an ordinary box there, but you've given me such*

doi:10.6342/NTU202402686

a scenario that I'm in a rich man's room, and this is his safe. Even if the box feels fake, my mind will fill this gap with the scenario, and I'll feel like it's quite realistic." When we asked if high-fidelity sensory cues influence their cognitive state, P29's statement revealed that if high-fidelity sensory cues align with the virtual scenario, it makes cognitive state transition to the virtual world, and vice versa: "*If we're talking about music, I feel like it affects my cognitive state because its atmosphere gives me a sense of immersion in the rich man's room. For example, when I'm doing a thrilling task and there's background music that makes me feel really tense. However, if it's about pressing buttons, the real sound of someone knocking on that wooden board, while also realistic, doesn't make me think about how it integrates with the story, environment, and task. Instead, it makes me feel like there's someone outside controlling the game.*" Virtual scenarios also affect how high-fidelity sensory cues transition cognitive state, and we explain this in more detail in 5.3.5.

5.1.5 Unusual experiences

While P8 and P21 indicated that the experience should be sense-making, 12 of 36 participants reported unusual experiences and virtual scenario made their cognitive state transition to the virtual world. P27 commented, "*I feel like VR is meant to bring something other than everyday life. For example, I cannot be a secret agent in real life, but I can experience it in a game. Furthermore, I won't get shot by bullets in my daily life, but this (study) allows me to experience it with a shield. Probably the things that I cannot do in everyday life would be interesting for me.*" When being asked what experience can be called VR, P9 answered, "It usually involves scenarios that are quite different from real life, often presenting situations and settings we wouldn't encounter in our everyday lives.

doi:10.6342/NTU202402686

These scenarios can range from fantasy to magical and so on." On the other hand, P27 and P31 reported that when engaging in activities that are closely connected to everyday life, the cognitive state tends to lean to the real world: "*because I can tell it(pressing buttons) is real, and these actions are more common in real life, it doesn't feel as much like being in VR.*"(P27)

5.2 Spatial Awareness

In this theme, we observed that spatial awareness affects cognitive state. The theme has 2 sub-themes: developing spatial awareness of the virtual environment and connection to the physical world.

5.2.1 Developing spatial awareness of virtual environment

P10 and P15 pointed out that developing spatial awareness of the virtual environment influences the cognitive state. P10 remarked, "*I think that perceiving these (realistic) elements would make it more like being in VR because they make you feel more like you exist in a new space, rather than just an empty room with nothing.*" P33 indicated the relationship between the virtual scenario and the importance of establishing the virtual environment's structure: "*[...] It means that to what extent one needs to construct the complete scene [depends on the virtual scenario]. If I have a harder time completing the task because I can't imagine myself in a wealthy person's room, the overall sense of immersion may be stronger. However, if the connection between the process of achieving my goals and the story isn't as deep, then the realism might be relatively weaker.*" P13 highlighted the importance of the sense of "being there", saying, "*In the VR experience,*

doi:10.6342/NTU202402686

what one sees is not actually what one sees in reality. Therefore, in terms of perception, it relies quite heavily on hearing every sound or picking up every clue, and also on the sense of touch to give oneself a genuine sense of being in the scene, and to know that one is truly inside the created environment."

P9 and P12 further indicated that exploring the environment makes the cognitive state transition toward virtual. P12 explained, "*I felt being in VR during the exploration phase when I was searching for mission locations. I felt detached from VR when I was solving missions. When I was exploring, I disconnected from the reality of the environment and became deeply absorbed in the imaginary scenarios.*" However, P25's cognitive state transition toward the physical world after the exploration: "*Once I had the width of the space I was in mind, that thing (the VR headset) was just like a blindfold to me. It was as if it somewhat pulled me back to reality. Just like if you had explored a space and you didn't have a virtual experience visually, it made me feel as if I were wearing a blindfold and being brought back to the real world.*" After people develop spatial awareness in the virtual environment, P9 suggested giving people the capability of entering/exiting the virtual space at will: "*Only entering and exiting doesn't quite like a VR. I would say if I could walk into and out of this virtual space by my willingness, I would feel more like in VR.*"

5.2.2 Connection to real world

P29, P32, P33, and P35 mentioned that their cognitive state transition to the physical world when they can imagine what the actual environment looks like. P32 reported, "*[...] as I progress further, I rely more on touch and hearing. While using these senses, I can't visualize what this virtual room looks like; I can only recognize it through touch. But*

doi:10.6342/NTU202402686

when relying on touch, I directly associate objects with their real-world counterparts. For example, a bed is just a bed. I can't imagine a new space with that; it's like trying to picture what this room looks like in real life. After solving more tasks that required me to touch, I started to form a clearer image of what this room might be like in the real world. As a result, my overall sense of VR gradually diminishes." P33 echoed that establishing the real environment's structure makes the cognitive state transition to the physical world: "*I can imagine that if I take off the headset right now and open my eyes, I am in a space like this, with some scenery. So, without any specific cues, lights, or signals, it doesn't seem like being in a wealthy person's room or a game scene, not at all. Instead, it's more about figuring out what kind of scene I am in (in real life).*" P12 indicated the connection between objects and the real world makes the cognitive state transition to the physical world: "*Let's start with the second area. In the second area, the vision is deprived, and it solely relies on touch. The sense of touch is quite strong, and it quickly brings me back to reality because these objects have many connections to the real world, which allows me to identify them instantly.*"

5.3 Sensory

5.3.1 Vision

Three participants (P4, P23, P32) emphasized that vision is the primary factor of their cognitive state. Certainly, the deprivation of vision does result in a transition to the physical world cognitive state, but the process of the transition is gradual rather than instantaneous. P32 "*After removing vision, the sense of VR gradually diminishes. It's not an immediate return to real life, but rather, over time, I feel more and more like they are*

doi:10.6342/NTU202402686

back in the real world." P35's statement indicated that virtual scenario helps maintain VR cognitive state in the transition process: "*It's somewhat like a gradual detachment(from VR) [...] My mind still recalls the previous tasks and scenarios [...], but because the attention gradually shifts to pure haptic, the (VR) cognitive state gradually declines.*" P26 referred that other sensory cues must be strong enough to make them feel like in VR: "*Because there's no visual support, it's not easy to feel like I'm in a virtual space through other senses alone. During this time, the feedback from my other senses needs to be strong enough for me to feel like I'm in VR.*"

5.3.2 Other cues

Participants reported that more and consistent sensory cues result in a transition to their virtual cognitive state. Not surprisingly, participants reported that "*more sensory input would make it feel more like VR.*"(P30) and the consistency of different sensory make their cognitive state lean to VR. P25 said, "*When my body's five sensory perceptions are consistent, it makes me feel as though I am inside that (VR) environment.*" P18's statement connect the inconsistency with unreality: "*When visual, haptic, and auditory cues don't match, it can create an unreality, causing me to slightly disconnect from the VR environment.*"

5.3.3 Providing sensory cues through devices

P9, P16, P20 reported that noticing the sensory cues are simulated can make cognitive state transition to VR; P9 described more precisely that the key is noticing the devices as intermediates between VR and real world: "*When I can get feedback through the device,*

doi:10.6342/NTU202402686

I feel like I'm someone immersed in the VR world because I'm using a VR device to get feedback." Furthermore, simultaneously perceiving sensory cues from devices and from real objects is reported as a factor of transition cognitive state to VR since it could create a sense of interaction between the virtual environment and the real world. P26 stated, "*I feel like I'm in VR when I press the buttons on the wall to produce sounds I hear within the constructed environment. I thought that the sound was coming from the virtual world, but when I pressed the button on the wall, it was relatively real; so, there was a feeling of interaction between the two worlds.*" However, four participants(P14, P17, P29, P32) reported an instant transition to the physical world cognitive state when they felt artificial manipulation during the experience. P14 stated "Suddenly pressing that button (which triggered a real hitting sound), and the sound of hitting the wall not appearing in time, I became aware that I was in a physical space with staff nearby. At that moment, it didn't feel like I was in VR."

5.3.4 Creating interactability

P9, P13, P26, P33, P36 mentioned the interactability enhances cognitive state. When we asked what is the main reason for being in VR, P13 responded, "*the interactability (is the main reason); it's like there's an interaction between my virtual and real experiences. The virtual scenario and the tactile sensations in reality interact with me (enhances cognitive state).*" "*I felt like being in VR in the first room I saw. The ability to interact with things like shields and other objects added to the feeling of being in VR. Also, when getting shot and receiving other feedback, I feel like interaction and feedback make me feel more immersed,*" P36 said, where the shield and getting shot are categorized as the active feedback in our study. On the other hand, three participants(P26, P33, P36) reported that the

doi:10.6342/NTU202402686

interaction with virtual objects also makes cognitive state transition to VR. "*(The feeling of being in VR is intense) when I began to realize that I could interact with some objects inside the screens. When I picked up the shield, it felt like, 'Oh, I'm entering the game.' The feeling was especially intense when I interacted with the system's voice,*" (P33).

5.3.5 Fidelity

P13, P15, and P18 state that high-fidelity sensory cues bring them into the virtual scenario, and subsequently improve their VR cognitive state. P15 stated, "*I felt a wind coming from the side, making it seem like the bridge was outdoors. Because the scenario was set in a high building, the presence of the wind felt entirely natural made me more accepting of the environment's setting.*"

However, some participants reported that after the deprivation of vision, high-fidelity sensory cues reduce the score of cognitive state; "*It feels less like VR, almost too real, like something that could be experienced without VR.*", P5 reported. P33 stated, "*When there are no lighting cues, all the other (realistic) signals doesn't make me imagine what the rich man's room or the game scene would look like, not at all; instead, I'm contemplating what kind of setting (in the physical world) I'm in.*" P22 also states, "*Whether the scene I imagine in my mind aligns with my other sensory perceptions will have the greatest impact within VR or not.*" Furthermore, P15 gave a clear description of the situation in which they believe high fidelity can make cognitive state, "*the VR environment is a highly fantastical place, then this sense of realism might be a bit of a drawback. If it's a very realistic VR space, then it could be considered a positive aspect.*" P24's statement also shows that high-fidelity sensory cues are connected to real life so they make cognitive state lean to the physical world:"*In my perception, VR and reality exist at two ends of a spectrum. The*

doi:10.6342/NTU202402686

realism makes me feel closer to real life, which, in turn, makes VR seem more distant to me."

On the other hand, three participants(P14, P22, P32) reported that low-fidelity sensory cues enhance cognitive state since they match their knowledge of VR technology. With regard to the instructor in Preparation, P14 added that it is unreal hence enhancing cognitive state: "*Apart from the fact that he breathes, the instructor doesn't quite seem like a real person, like his posture or his reaction when seeing us. [...] More like in VR because based on my current understanding of VR, it would have these characteristics in techniques. The characters wouldn't be that realistic.*" P12 also commented that making sensory cues unreal can be a good way to break the connection with the real world: "*For instance, English letters are supposed to transform so that I wouldn't be able to quickly connect them back to being just letters and return to reality. If the (sensory cues) are something that can't be connected to the real world, that's when my virtual imagination takes over more.*"

5.4 Expectation

The theme has 2 sub-themes: sensory cues match/mismatch expectation and task order.

5.4.1 Sensory cues match/mismatch expectation

P7, P9, P16, P23, and P35 mentioned that their cognitive state transition to virtual if the sensory cues they perceive match or even exceed their expectations, and vice versa. P7 reported, "*For example, I didn't anticipate that it (the experience) could provide a*

doi:10.6342/NTU202402686

tactile sensation, which enhances my feeling of being in VR. It should offer various sensory cues to me. If something should appear but doesn't, like my hand, it would diminish the sense of being in VR." P35 echoed this statement, saying, "*The second thing (that affects my cognitive state) is whether what I actually touch matches what I imagine. It's the alignment of these two; if they match, I also feel like it's quite authentic in that way.*" P16 further indicated that the expectation comes from the virtual scenario when expressing what breaks the immersion, noting, "*He (the trainer in VR) looks very fierce, but the bullets he shot are very soft. And because you got hit, you can tell that it's something like a sponge or similar material. It's quite different from what you'd expect from the muzzle of the gun.*"

5.4.2 Task order

The sequence of tasks also impacts the cognitive state. P1, P10, P16, and P31 reported that they compared tasks, leading their cognitive state transition to the physical world: " *Because the single-plank bridge left such a deep impression on me. [Researcher: You mean because pressing the button did not leave an impression as deep as the single-plank bridge to you, so the score is lower?] Yes*" (P1). P16 indicated that the matches of the expectation also raise the bar for what is expected with saying, "*When previous tasks meet expectations, any deviation after can lead to disappointment.*"

doi:10.6342/NTU202402686

Chapter 6

Discussion

From the results of the thematic analysis, we found two topics contribute to establishing VR cognitive state that corresponds to the two dimensions of Slater's perceptual presence model [29, 31], virtual scenario to Psi, and spatial awareness to PI. Moreover, besides virtual scenario and spatial awareness, we extract the other four topics as factors of cognitive state: sensation, fidelity, the source of cues, and expectation.

6.1 Virtual Plausibility

In our work, we introduce Plausibility Illusion(Psi) [31] to the concept of our cognition state in an MR system and define two terms: physical plausibility and virtual plausibility. Psi occurs when users believe "what is apparently happening is really happening", leading to three possible cognitive situations: 1) Physical plausibility, indicating the belief that the event is occurring in the real world; 2) Virtual plausibility, signifying the belief of it is happening in the virtual world; 3) Both physical and virtual plausibility simultaneously occurring. These two terms expand the discussion of Psi. For instance, in the task

doi:10.6342/NTU202402686

of pressing buttons reported by P27 (subsection 5.1.5), Psi occurred as users believed they actually pressed the button. In this case, physical plausibility occurred, but virtual plausibility failed since pressing buttons is not plausible in the given virtual scenario. P35's statement in subsection 5.1.1 and P13's statement in subsection 5.3.4 also indicate that virtual plausibility and physical plausibility can simultaneously occur. We believe that the occurrence of physical plausibility leads to a transition in the cognitive state toward the physical world, while the occurrence of virtual plausibility leads to a transition in the cognitive state toward the virtual world. We focus more on virtual plausibility since how virtual plausibility success and failure differ through different MR systems. We consider building and maintaining a reliable virtual scenario as a key factor of virtual plausibility. In a system with a reality-simulating scenario, the mechanism of virtual plausibility and physical plausibility are close; while in a system with a fantasy scenario, these two plausibility might differ a lot.

From section 5.1, we observed the occurrence of virtual plausibility brings a transition of cognitive state to VR, and vice versa. We also observed three factors that contribute to forming virtual plausibility and maintaining a virtual scenario:

6.1.1 Unusual experience

The observations in (subsection 5.1.5) revealed that unusual experiences make cognitive state transition to VR, while common tasks close to everyday life make cognitive state lean to the physical world. Moreover, building a fantasy virtual scenario makes unusual experiences more likely to produce virtual plausibility. Therefore, we suggest designers to design a more fantasy virtual scenario with more unusual experiences.

doi:10.6342/NTU202402686

6.1.2 Believing sensory cues come from virtual world

From subsection 5.2.2, we observed that by perceiving sensory cues that trigger their imagination of the physical world, participants build a strong physical plausibility instead of virtual plausibility. P9 and P26's statements in subsection 5.3.3 demonstrate a strong virtual plausibility is formed when they feel the sensory cues are provided by devices. On the other hand, we observed a key factor of not believing sensory cues are from the virtual world is the awareness of artificial manipulation in subsection 5.3.3. Therefore, we suggest designers let users notice sensory cues given by devices and controlled by the system to form their virtual plausibility.

6.1.3 Allocate more users' attention to virtual scenario

From subsection 5.1.3, we observed a detachment due to a long time focusing on a single task. Moreover, from subsection 5.1.1, instructions that form virtual plausibility help re-focusing on the virtual scenario, and enhance the virtual scenario. Therefore, we recommend designers focus on keeping users' attention on the virtual scenario.

6.2 Spatial Awareness

From section 5.2, we observed that not only Place Illusion (Pi) [31] but also spatial awareness are the factors influencing cognitive state. Similar to what we did for Psi in section 6.1. We also define virtual spatial awareness and physical spatial awareness. When people develop virtual spatial awareness, their cognitive state transition to the virtual world, and vice versa. Similarly, if people feel like they are in the virtual environment

doi:10.6342/NTU202402686

(Pi), their cognitive state transitions to the virtual world as well. We note that Pi is different from spatial awareness. While Pi typically refers to the sense of "being there", developing spatial awareness involves constructing a clear structure of the environment. However, the development of spatial awareness is usually accompanied by Pi, as people tend to have the sense of being in a place after having a clear structure of the space in mind. Since both Pi and Psi are factors affecting cognitive state, we acknowledge that the sense of presence is a factor that influences cognitive state.

From P33's quote, we can also observe that the importance of virtual spatial awareness to the cognitive state depends on the virtual scenario. Specifically, if the experience is story-based, it is highly recommended to assist users in developing their virtual spatial awareness. One useful method suggested by participants is exploring the scene as the exploration helps people construct the structure of the virtual environment. Although P25's cognitive state transition to the physical world after the exploration task, this could be attributed to focusing on one task for too long (as discussed in subsection 5.1.3). In summary, to help people develop their virtual spatial awareness, we suggest immersive experience designers incorporate more explorative tasks into their designs.

6.3 Sensation

From the overall experience score in Figure 5.1(left) exhibits a drop after vision deprivation and the results from subsection 5.3.1, vision cues are demonstrated as the key factor of cognitive state. From subsection 5.3.2, we observed that rich and consistent sensory cues make cognitive state transition to VR. According to subsection 5.3.4, we recommend that instead of merely providing sensory feedback, users should be encour-

doi:10.6342/NTU202402686

aged to actively engage with and receive sensory stimuli through interaction. However, from subsection 5.2.2, and subsection 5.1.1, we consider the importance of sensory cues to align them with the virtual scenario and spatial awareness.

6.4 Fidelity

Figure 5.2 and the participants' statement in subsection 5.3.5 demonstrate that in most situations, high-fidelity sensory cues could make cognitive state transition to VR. However, it can also be observed that the participants who reported an increase in fidelity enhancing their cognitive state decreased by more than half from Situation 1 to Situation 3. In subsection 5.3.5, the results demonstrate the reason of high-fidelity sensory cues result in a transition to the physical world of cognitive state including triggering the imagination of physical space, not matching the virtual scenario, and connecting to the physical world. On the other hand, we extract two situations in which low fidelity can make cognitive state transition to VR: 1) alignment with participants knowledge of current technology and 2) breaking the connection with the physical world. From these findings, we consider that how high and low-fidelity sensory cues affect cognitive state is highly related to other discussion topics.

Nonetheless, we have discerned two constraints within our work concerning the discussion of fidelity. First, from the interview, we found that participants' perception of sensory fidelity did not completely align with our design. For example, P5 mentioned that "*the shield is more realistic (compared to getting hit themselves)*" while blocking bullets with the shield is low-fidelity in our design; P9 mentioned that "*[...]pillow, puzzle, and safes - all of these felt very real*" while pillow in our design is low-fidelity. Moreover,

doi:10.6342/NTU202402686

some sensory cues reported by participants are not on our list, such as the shape and the size of the safe, the haptic sensory cues from the shield itself, and the lack of mouth movement while the instructor was talking. Second, we found some participants confused real with reasonable. When we asked if there was anything unreal in room c, P14 replied "*the position of the bed*", while the position of the bed is less relevant to fidelity. Therefore, we consider that attempting to control the fidelity of all sensory cues within an experience is unrealistic and inefficient. However, if there is a need for a deeper understanding of the relationship between fidelity and cognitive state, more validated methods for evaluating fidelity should be introduced.

6.5 Expectation

From 5.4, we recognize that whether sensory cues match people's expectations has a significant impact on their cognitive state. When the sensory cues align with or even exceed what people expect to perceive, their cognitive state transitions to the virtual world, and vice versa. As people usually derive their expectations from the virtual scenario, it is crucial for designers to ensure that they can provide corresponding cues based on the virtual scenario. While providing stronger cues requires additional effort, designers could also consider revising their virtual scenario to better align with the expected cues.

6.6 Order of Tasks

From the quantitative result, we observed that the overall cognitive score of G1 is higher than G2 in Figure 5.1(left). It is probably because the first task (log-in agent system) for G1 is vision-oriented, which helps people in G1 develop a good spatial awareness

doi:10.6342/NTU202402686

at the beginning. In contrast, people in G2 got vision-deprived in task 4 right after they experienced the vision-oriented task, making them feel relatively more deprived. This also explains why G2 has greater drops between tasks 3 and 4 for both the overall and task respectively scores. The order effect is also reported by participants in the theme 5.4.2, as the bar of the expectation will be raised if the sensory cues fulfill the previous expectation. In other words, if the system's capabilities cannot perfectly provide appropriate sensory feedback that matches the growing expectations, the cognitive state will tend to transition toward the physical world at some point. We recommend designers control the sequence of tasks to ensure that this transition occurs in less critical moments while simultaneously lowering user expectations. In contrast, there is no report on whether the mismatch of the expectation could lower the expectation of the next task. We consider this as a possible order effect since people usually compare their current experience with the previous one and regard it as one of our future work.

6.7 Summary of Guidelines

To summarize the discussion in this section, we provide several guidelines for establishing a user's virtual cognitive state.

- Design a fantasy virtual scenario with more unusual experiences
- Let users notice sensory cues given by devices and controlled by the system
- Make users allocate more attention to the virtual scenario
- Design the process of exploring the virtual space in a way that closely aligns with the virtual scenario
- Consider adjusting virtual scenario to match sensory cues

doi:10.6342/NTU202402686

- Carefully selecting the moment to allow the user's cognitive state to transition to reality, thereby lowering expectations

6.8 Study Limitations

Visual Sensory Dominance

A significant limitation of our study is the enduring dominance of visual sensory input, as mentioned in 5.3.1. Despite our attempts to deprive users of their visual sense, it becomes evident that once we remove their visual input, the users' cognitive state significantly diminishes. While certain factors may help maintain or enhance cognitive states, their influence pales in comparison to the impact of visual sensory cues. Some studies have also explored how vision can assist in deceiving perception when it comes to less realistic sensory cues, highlighting the predominant role of vision as the most influential factor in cognitive states.

Sensory Overload

Another limitation we encountered in our study is the inclusion of an excessive number of sensory cues within each task. This abundance of sensory cues can lead to a multitude of interrelated factors influencing one another. Participants were exposed to numerous sensory cues simultaneously, making it challenging to pinpoint the specific influence of each cue on their cognitive state. Future investigations might benefit from a more focused examination of individual sensory cues to gain deeper insights into their contributions to cognitive state.

doi:10.6342/NTU202402686

Chapter 7

Conclusion

Drawing on the dual-process theory from cognitive psychology, we propose separating cognitive state and presence to better address the limitations of the traditional concept of presence in the face of diverse immersive experiences of the future. We hypothesize there are cognitive state transitions within an immersive experience. Through a user study involving an immersive escape room experience, we observed cognitive state transitions influenced by different sensory cues. Using thematic analysis, we outlined the six key factors affecting these transitions across four themes. We then give seven guidelines for future experience design after understanding the factors of cognitive state.

doi:10.6342/NTU202402686

References

[1] P. Baudisch, H. Pohl, S. Reinicke, E. Wittmers, P. Lühne, M. Knaust, S. Köhler, P. Schmidt, and C. Holz. Imaginary reality gaming: ball games without a ball. In Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology, UIST '13, page 405–410, New York, NY, USA, 2013. Association for Computing Machinery.

[2] bHaptics. bhaptics tactsuit x40, 2021.

[3] F. Biocca. The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments [1]. Journal of Computer-Mediated Communication, 3(2):JCMC324, 09 1997.

[4] V. Braun and V. Clarke. Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2):77–101, 2006.

[5] A. Cahen and M. Tacca. Linking perception and cognition. Frontiers in Psychology, 4, 2013.

[6] L.-P. Cheng, Y. Chen, Y.-H. Peng, and C. Holz. Reality rifts: Wonder-ful interfaces by disrupting perceptual causality. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23, New York, NY, USA, 2023. Association for Computing Machinery.

doi:10.6342/NTU202402686

[7] A. R. Cheraghi, S. Shahzad, and K. Graffi. Past, present, and future of swarm robotics. In Intelligent Systems and Applications: Proceedings of the 2021 Intelligent Systems Conference (IntelliSys) Volume 3, pages 190–233. Springer, 2022.

[8] J. S. B. Evans. In two minds: dual-process accounts of reasoning. Trends in cognitive sciences, 7(10):454–459, 2003.

[9] J. S. B. T. Evans. On the resolution of conflict in dual process theories of reasoning. Thinking & Reasoning, 13(4):321–339, 2007.

[10] J. S. B. T. Evans and K. E. Stanovich. Dual-process theories of higher cognition: Advancing the debate. Perspectives on Psychological Science, 8(3):223–241, 2013. PMID: 26172965.

[11] S. Goldstein, J. Campbell, and T. Mowry. Programmable matter. Computer, 38(6):99–101, 2005.

[12] T. Hartmann and M. Hofer. I know it is not real (and that matters) media awareness vs. presence in a parallel processing account of the vr experience. Frontiers in Virtual Reality, 3, 2022.

[13] M. Hofer, T. Hartmann, A. Eden, R. Ratan, and L. Hahn. The role of plausibility in the experience of spatial presence in virtual environments. Frontiers in Virtual Reality, 1, 2020.

[14] M. E. Latoschik and C. Wienrich. Congruence and plausibility, not presence: Pivotal conditions for xr experiences and effects, a novel approach. Frontiers in Virtual Reality, 3, 2022.

doi:10.6342/NTU202402686

[15] K. M. Lee. Presence, Explicated. Communication Theory, 14(1):27–50, 01 2006.

[16] M. Lombard and T. Ditton. At the Heart of It All: The Concept of Presence. Journal of Computer-Mediated Communication, 3(2):JCMC321, 09 1997.

[17] Meta. First enounters, 2023.

[18] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. IEICE TRANSACTIONS on Information and Systems, 77(12):1321–1329, 1994.

[19] M. Minsky. Telepresence. 1980.

[20] M. Nunes. Jean baudrillard in cyberspace: Internet, virtuality, and postmodernity. Style, pages 314–327, 1995.

[21] C. S. Oh, J. N. Bailenson, and G. F. Welch. A systematic review of social presence: Definition, antecedents, and implications. Frontiers in Robotics and AI, 5, 2018.

[22] M. I. Posner, M. J. Nissen, and R. M. Klein. Visual dominance: an information-processing account of its origins and significance. Psychological review, 83(2):157, 1976.

[23] P. A. Rauschnabel, R. Felix, C. Hinsch, H. Shahab, and F. Alt. What is xr? towards a framework for augmented and virtual reality. Computers in Human Behavior, 133:107289, 2022.

[24] L. Ropolyi. Virtuality and reality—toward a representation ontology. Philosophies, 1(1):40–54, 2015.

[25] T. W. Schubert. A New Conception of Spatial Presence: Once Again, with Feeling. Communication Theory, 19(2):161–187, 04 2009.

doi:10.6342/NTU202402686

[26] R. Skarbez, F. P. Brooks, Jr., and M. C. Whitton. A survey of presence and related concepts. ACM Comput. Surv., 50(6), nov 2017.

[27] R. Skarbez, M. Smith, and M. C. Whitton. Revisiting milgram and kishino's reality-virtuality continuum. Frontiers in Virtual Reality, 2, 2021.

[28] M. Slater. A note on presence terminology. Presence connect, 3(3):1–5, 2003.

[29] M. Slater. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1535):3549–3557, 2009.

[30] M. Slater. Immersion and the illusion of presence in virtual reality. British Journal of Psychology, 109(3):431–433, 2018.

[31] M. Slater, D. Banakou, A. Beacco, J. Gallego, F. Macia-Varela, and R. Oliva. A separate reality: An update on place illusion and plausibility in virtual reality. Frontiers in Virtual Reality, 3, 2022.

[32] M. Slater, A. Steed, J. McCarthy, and F. Maringelli. The influence of body movement on subjective presence in virtual environments. Human Factors, 40(3):469–477, 1998. PMID: 9849105.

[33] M. Slater and S. Wilbur. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments, 6(6):603–616, 12 1997.

[34] S. A. Sloman. The empirical case for two systems of reasoning. Psychological bulletin, 119(1):3, 1996.

doi:10.6342/NTU202402686

[35] E. R. Smith and J. DeCoster. Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. Personality and social psychology review, 4(2):108–131, 2000.

[36] C. Stapleton, C. Hughes, M. Moshell, P. Micikevicius, and M. Altman. Applying mixed reality to entertainment. Computer, 35(12):122–124, 2002.

[37] J. Steuer, F. Biocca, M. R. Levy, et al. Defining virtual reality: Dimensions determining telepresence. Communication in the age of virtual reality, 33:37–39, 1995.

[38] P. Touzalin-Chretien, S. Ehrler, and A. Dufour. Dominance of Vision over Proprioception on Motor Programming: Evidence from ERP. Cerebral Cortex, 20(8):2007–2016, 12 2009.

[39] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks Jr. Walking> walking-in-place> flying, in virtual environments. In Proceedings of the 26th annual conference on Computer graphics and interactive techniques, pages 359–364, 1999.

[40] P.-Y. C. Wang, R. Rajesh, N. Y. X. Lee, A. S. S. Loose, N. Overdevest, N. Semertzidis, and F. Mueller. Dreamception: Towards understanding the design of targeted lucid dream mediation. In Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems, CHI EA ’24, New York, NY, USA, 2024. Association for Computing Machinery.

[41] P. Wason and J. Evans. Dual processes in reasoning? Cognition, 3(2):141–154, 1974.

[42] W. Wirth, T. Hartmann, S. Böcking, P. Vorderer, C. Klimmt, H. Schramm, T. Saari, J. Laarni, N. Ravaja, F. R. Gouveia, F. Biocca, A. Sacau, L. Jäncke, T. Baumgartner,

doi:10.6342/NTU202402686

and P. Jäncke. A process model of the formation of spatial presence experiences. Media Psychology, 9(3):493–525, 2007.

[43] B. G. Witmer and M. J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoperators and Virtual Environments, 7(3):225–240, 06 1998.

[44] X. Xia, C. M. Spadaccini, and J. R. Greer. Responsive materials architected in space and time. Nature Reviews Materials, 7(9):683–701, 2022.

[45] P. Zahorik and R. L. Jenison. Presence as being-in-the-world. Presence, 7(1):78–89, 1998.

doi:10.6342/NTU202402686